

**JOURNAL OF THE
SOUTHEAST ASIAN
LINGUISTICS SOCIETY**

Volume 4.1

May 2011

JSEALS

Journal of the Southeast Asian Linguistics Society

Managing Editor:

Paul Sidwell (Pacific Linguistics, Canberra)

Editorial Advisory Board:

Mark Alves (USA)

George Bedell (Thailand)

Marc Brunelle (Canada)

Gerard Diffloth (Cambodia)

Marlys Macken (USA)

Brian Migliazza (USA)

Keralapura Nagaraja (India)

Peter Norquest (USA)

Amara Prasithrathsint (Thailand)

Martha Ratliff (USA)

Sophana Srichampa (Thailand)

Justin Watkins (UK)

JSEALS is the peer-reviewed journal of the Southeast Asian Linguistics Society, and is devoted to publishing research on the languages of mainland and insular Southeast Asia. It is an electronic journal, distributed freely by Pacific Linguistics (www.pacling.com) and the JSEALS website (jseals.org).

JSEALS was formally established by decision of the SEALS 17 meeting, held at the University of Maryland in September 2007. It supersedes the Conference Proceedings, previously published by Arizona State University and later by Pacific Linguistics.

JSEALS welcomes articles that are topical, focused on linguistic (as opposed to cultural or anthropological) issues, and which further the lively debate that characterizes the annual SEALS conferences. Although we expect in practice that most JSEALS articles will have been presented and discussed at the SEALS conference, submission is open to all regardless of their participation in SEALS meetings. Papers are expected to be written in English.

Each paper is reviewed by at least two scholars, usually a member of the Advisory Board and one or more independent readers. Reviewers are volunteers, and we are grateful for their assistance in ensuring the quality of this publication. As an additional service we also admit data papers, reports and notes, subject to an internal review process.

JSEALS is published twice a year. Papers can be submitted to the Managing Editor, electronically (paul.sidwell@anu.edu.au or paulsidwell@yahoo.com) at any time.

Contents

Editorial	iv
Papers	
A Cognitive Linguistics Storyline Analysis of Iu-Mien Narrative Discourse <i>T. Daniel Arisawa</i>	1
Fixed Autosegmentism in Thai Emphatic Reduplication. <i>Leslie Lee</i>	41
Sorbung, an Undocumented Language of Manipur: its Phonology and Place in Tibeto-Burman <i>David Mortenson and Jennifer Keogh</i>	64
Floating Quantifiers in Burmese and Thai <i>Andrew Simpson</i>	115
The Evolution of Polyfunctionality of <i>dây</i> Construction in Thai: Split Patterns of Possibility-Related Modal Concepts <i>Kiyoko Takahashi</i>	147

Editorial

Welcome to JSEALS Volume 4.1. We are now into our fourth year and I can only report that our journal is humming along nicely; the quality of papers is excellent, and I am pleased to relate that there is quite a queue of papers shaping up for 4.2.

This volume also sees abstracts and keywords for all papers, reflecting the ongoing process of improving standards and, importantly, accessibility. And in this regard, I urge you please to check the updated *Guide for Authors* on the last page - things run so much smoother when authors strictly follow the submission guidelines.

As issue 4.1 is being released, the 2011 SEALS meeting is about to get underway. The host institution - Kasetsart University in Bangkok - is offering beautiful modern meeting rooms in a great campus setting. Acharn Kitima Indrambarya, Chair of the Organizing Committee, is doing a superb job, and we look forward to a very memorable meeting.

Finally, I would like to pay homage to Uri Tadmor, who recently stepped down from the SEALS International Committee and the JSEALS Editorial Advisory Board. Uri, now working as an Editorial Director for De Gruyter Mouton, has been a supporter of SEALS over many years, and we especially acknowledge the great job he did taking on and running the 16th SEALS meeting at Atma Jaya University, Jakarta, Indonesia, in 2006. Thank you Uri.

Paul Sidwell (Managing Editor)
May 2011

A COGNITIVE LINGUISTICS STORYLINE ANALYSIS OF IU-MIEN NARRATIVE DISCOURSE¹

T. Daniel Arisawa

Chiangrai Rajabhat University, Thailand
<danielta928@hotmail.com>

Abstract

A storyline is a foregrounded mainline of development in narrative discourse. Taking the narrative as a mental picture viewed by a hearer/reader (i.e. conceptualizer), the Cognitive Linguistics posits two groups of constructions which make storyline emerge to the foreground: progression and sequential structures.

The storyline progression in Iu-Mien narrative is encoded with the unmarked action/motion verbs, the development adverb *aengx* ‘furthermore’, Serial Verb Construction, topic chains, and the multiclausal constructions. The sequential structures include the prospective conjunction *ziouc* ‘and then (soon)’, the retrospective conjunction *cingx_daaih* ‘therefore’, the topic marker *aeqv* ‘as for’, and the ad-clause containing *V-liuz* ‘after V-ing’.

Keywords: discourse, cognitive linguistics, text analysis

0. Introduction

A storyline is a foregrounded mainline of development in narrative discourse. This study seeks to investigate a storyline in Iu-Mien narrative discourse. The purpose is to show what it means that the storyline is foregrounded. The approach taken here is Cognitive Linguistics (CL), particularly Cognitive Grammar (CG) theorized by Ronald W. Langacker (e.g. 1987, 1991a, 1991b, 1991c, 2000), rather than a traditional and popular textlinguistics developed by Robert E. Longacre (e.g. 1981, 1983, 2003a, 2003b). Though Longacre’s theory was much influenced by Hopper and Thompson’s correlation between high/low transitivity and the foreground/background in discourse, he has never elaborated the relation.

Longacre emphasizes (to simplify various the arguments) that a discovery the particular verb which has a morphosyntactic marking for the past tense or perfective aspect inevitably leads a researcher to identify the storyline. Iu-Mien, however, totally lacks such a marking in the verb; hence requires an alternative method. Following Somsonge’s non-verb-morphology dependent storyline identification method (i.e. temporal movement approach) in Thai (1990, 1992) and Hlai (2002), it turns out that the storyline in Iu-Mien can be analyzed in terms of two major factors: the chronological forward movement of action/event (e.g. topic chains and serial verb constructions, etc), and a bundle of the

¹ I would like to thank the anonymous reviewers for their remarks, and also to thank Paul Sidwell the managing editor for his invaluable comments for improvement of my paper, editorial advice and his patience and kindness. I take full responsibility for any flaws found in this paper.

linguistic expressions indicating sequentiality (e.g. a developmental adverb and an inchoative conjunction, etc.) in the narrative. In this study, the former is referred to as “progression”, and the latter “sequential structure”,² owing Talmy (2000).

The method in terms of progression and sequential structure explains cognitive linguistically why the storyline is foregrounded. From the CL perspective, a narrative is taken as a mental picture unfolded in front of the hearer/reader’s eyes. The narrative’s storyline is viewed as a thick, noticeable line that stands out (i.e. foregrounding) to the visual perception in the hearer/reader’s mental space as a result of combining two major factors: progression and sequential structure of the story.

Though seven story texts of Iu-Mien (over 700 sentences) were analyzed in search of storyline, among other discourse features, only one sample story is appended in this paper. Some examples are drawn from other sources.

1. Departure from Longacre’s theory

Longacre’s storyline theory within the general framework of textlinguistics is based on the discourse analyses of some eighty languages of the world. It is significantly influenced by his over fifty year study of Biblical Hebrew and Hopper and Thompson’s (1980) research on the correlation between high/low transitivity parameters and the foreground /background.

1.1. Longacre’s Premise

Longacre’s strong emphasis is on the virtual equation between the preterite verb which furthers the story and the foregrounded events or the storyline (e.g. 1981:337-47). For him, discovery of the preterite tense or completive aspect almost automatically eliminates the off-the-line materials as the background, i.e. non-storyline, then arrives at identification of the storyline.

1.2 Beyond Longacre: Temporal Movement

Gradually in the course of development of Longacre’s theory, some questions regarding the rigid connection between the morphosyntactic marking (e.g. preterite verb) and the foregrounded events or the storyline began to be raised even from among his colleagues and former students. One example is Somsonge³ who states:

² I appreciate the reviewers’ comments for improvement, particularly their recommendation that the term ‘sequentiality’ should be more general, non-particularized and straightforward. Upon consideration, however, I kept it in some places as in my original draft following Givon (1993, 1984), whose influence can be seen in Endo (1996, 2003). It is their notion of ‘topic-continuity’, ‘sequential-action’, ‘theme continuity’ (Givon 1984:296-97) and the materials which ‘control the flow of the story’ (Endo 1996:232, 244-97) that my ‘sequentiality’ means. Talmy’s ‘sequential structure’ is also comprehensive enough to signify the sense by Givon and Endo; it refers to ‘the patterns in which a number of the elements of some category combine in a sequence through time’ (Talmy 2000:463). Thus, in this paper both ‘sequentiality’ and ‘sequential structure’ can include verbs with the sequential form or adverb, and conjunctions. The use of ‘sequentiality’ in this study is restricted in the storyline. By contrast, the common textlinguistic term ‘coherence’ (i.e. a text’s property that enables a hearer/reader to ‘fit its different elements into a single overall mental representation’ (Dooley and Levinsohn 2001:23)) is broader and ‘cohesion’ (i.e. ‘the use of linguistic means to signal coherence’ (ibid: 27)) is narrower than the ‘sequentiality’.

³ Other examples are Somsonge (1992) and Dry (1992) in the anthology for Longacre.

The study of bipartite structure [i.e. storyline vs. non-storyline materials] of discourse information in Thai reveals that in language without verbal inflection as Thai, the bipartite structure is not expressed solely by the verb system as in English but by a conspiracy of non-systemic ways which include types of verbs, adverbs, time phrases, sequential signals, temporal clauses/phrases/words, auxiliaries, pre-serial verbs, and post-serial verbs. (Somsonge 1990:76) [Underline added]

She has developed, through the studies of Thai (1990, 1992) and Hlai (2002) and others, the way to identify the storyline by finding the temporal movement of narrative⁴ which may be indicated by the various constructions mentioned in the above quotation. It should be noted that such constructions as “adverbs, time phrases, sequential signals, temporal clauses/phrases/words, auxiliaries”, etc. were all (but punctiliar adverbs) categorized as the non-storyline materials by Longacre. However, they play an important role in identifying the storyline in Iu-Mien. Just as Thai is a language without verbal inflection, so is Iu-Mien. Inspired by Somsonge, the materials indicating temporal movement and sequentiality of narrative will be investigated.

1.3 Application to Iu-Mien: Progression and Sequential Structure

The materials contributing to the storyline in Iu-Mien can be categorized into two major groups. The first are those which indicate chronological forward movement, increment of new information, transition of actions/events in the flow of discourse: they include the development adverb *aengx* ‘furthermore’, Serial Verb Constructions (SVCs), and multiclausal constructions containing SVCs and topic chains. The property found in them can be summarized in the term “progression” (Talmy 2000:425-26).

The second involves those which indicate sequentiality within and across sentences: i.e. the prospective (or inchoative) conjunction *ziouc* ‘then’, the retrospective conjunction *cingx daaih* “therefore”, the ad-clause V + *liuz* “after V-ing”, and the topic marker *aeqv* “as for”. These constructions may be treated under the term “sequential structure” (Talmy 2000:463-76) or “sequentiality”.⁵

2. A Cognitive Grammar Outlook on Narrative: Current Discourse Space (CDS)

While textlinguistics⁶ tends to view narrative text outside a mind, e.g., as a written text on pages of a book or some sorts of recording, Cognitive Linguistics (CL) including CG considers it as a mental entity conceptualized in a mind. For example, Talmy⁷ expresses his fundamental assumption in approaching narrative as “the existence of a mind that has produced the narrative as well as of a mind that is cognizing the narrative” (2000:417-8). The schematization of cognizing or perceiving of the narrative is proposed in the framework of Current Discourse Space (CDS) by Langacker (2001:144-50; 2008:281, 464-67), an extension from his “stage model” (Langacker 1991a:211) at the clause level. The

⁴ Earlier than Somsonge, Dry (1981, 1983) studied this aspect.

⁵ In questioning Hopper and Thompson’s (1980) absence of mentioning “sequentiality”, Longacre stated, “Sequentiality must in fact be taken as prime characteristic of the storyline, i.e., of foregrounding in narrative” (1996:25) but has not developed this point further.

⁶ Earlier, in the 1960s and 70s, Longacre and his colleagues’ works were largely on the field-recorded written narrative discourses; hence their theory was usually called “discourse analysis”. However, when the term began to be applied to wider usage in such disciplines as conversation analysis or sociolinguistic study on people’s interaction as a discourse, Longacre adopted the term “textlinguistics”.

⁷ Other example is Gärdenfors (1999:21), who says, “Cognitive models are mainly perceptually determined (meaning is not independent of perception)”.

CDS as defined by Langacker is “the mental space comprising those elements and relations constructed as being shared by the speaker and hearer as a basis for communication at a given moment in the flow of discourse” (2001:144). The Figure 1 adapted from Langacker (2001, 2008) shows that the speaker (S) and the hearer (H) are interacting (indicated by the dashed two-way arrow) in mentally viewing (indicated by the upward dashed arrows) the focused part of a narrative which is profiled in the heavy line. “A discourse comprises a succession of frames each representing the scene being ‘viewed’ and acted on by the speaker and hearer at a given instant” (Langacker 2001:151).

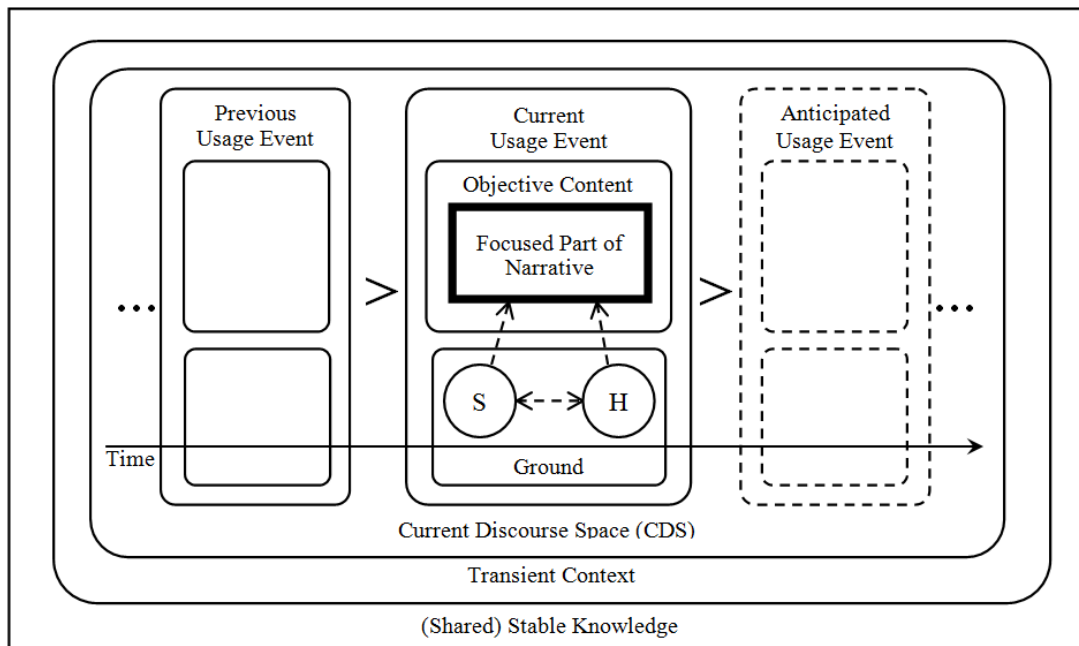


Figure 1. *Current Discourse Space in the Cognitive Grammar outlook on narrative*

Relevant parts of the Figure 1⁸ will be used to explain the conceptualization of storyline materials (e.g. briefly mentioned in 1.3) in the following pages. The important thing at this point is that a narrative discourse is viewed as a cognitive entity in a CL perspective.

3. Progression

The first component of the storyline is progression and it is analyzed based on a fundamental nature of human’s perceptive ability, called cognitive psychological principle I in this study. In analyzing some constructions contributing to progression, the principle I plays a role to explain how the storyline is foregrounded.

3.1 Cognitive Psychological Principle I

The analysis of progression in the storyline is based on a general nature of visual perception. Figure 2 shows that a moving object (b) is more conspicuous than a static one (a) to the perception of the viewer (V). This is a fundamental tendency of human cognitive

⁸ For the detail explanation of each part of the diagram, see Langacker (2001:145) and (2008:466).

ability. In the analysis of storyline, the viewer (V)⁹ can be represented by a speaker and writer or a hearer and reader of narrative, i.e. S and H in Figure 1.

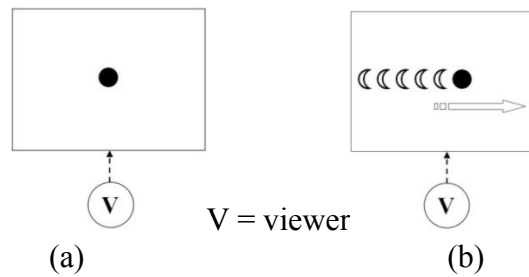


Figure 2. *Static and moving objects*

3.2 Unmarked action/motion verbs

The first linguistic construction on the storyline is the unmarked action/motion (as opposed to static) verb, which is the simplest and most common, as exemplified in (1).

(1) AS.070

<i>Ninh</i>	<i>mingh</i>	<i>wuov</i>	<i>ndiev</i>
3sg	go	there	down_there
pn.p	vi	pn.dem	adv.loc

‘He went down there.’

An intransitive verb is used in (1), whereas a transitive verb in (2).

(2) AS.015

<i>Nzopv</i>	<i>jiepv</i>
pierce	bear
vt	n

‘(She) pierced into the bear('s mouth).’

The verb *mingh* ‘go’ in (1) is depicted in Figure 3, where the mover (M) (c.f. Langacker 1991c:288) in the original location (L₁) moves to the new location (L₂) through an action of going. On the other hand, the conceptualization of the verb *nzopv* ‘pierce’ in (2) which has a patient (PAT) is described in Figure 4, following Langacker’s “stage model” (1991a:211), also called “Canonical Event Model” (1991c:285). The transmission of energy is depicted in these models and the energy transmission is a core element of events in a discourse. In this way the conceptualization of events is analyzed in a way analogous to a theater-goer viewing a play on the stage.

⁹ Elsewhere, Langacker sometimes uses “conceptualizer” (1991c:242-44).

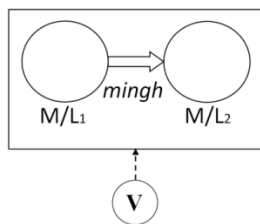


Figure 3. *mingh* ‘go’

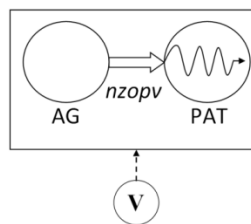


Figure 4. *nzopv* ‘pierce’

Besides the above quoted single unmarked verb in the storyline¹⁰, the main verb can be accompanied by an aspectual verb which always comes in a post verb position. Example (3) shows the construction ‘unmarked verb + post verb aspectual verb.’

(3) AS.097

97.1		97.2
<i>Mingh wuov ndiev!</i>		<u><i>gorngv ziangx</i></u>
go there down_there		speak finish
vi pn.dem adv.loc		v vi

‘^{97.1}(The governor’s people) went far down there ^{97.2}(and) finished talking (with the brave man).’

3.3 Development adverb *aengx* ‘further’

Secondly, another construction that has a force to push the story forward is ‘*aengx* + verb’, where *aengx* can be (i) most basically translated as ‘again’, ‘also’ meaning repetition; and (ii) ‘and then’, ‘further’, ‘furthermore’, or ‘moreover’ at the discourse level. This adverb often introduces a new development of a story, hence called the development adverb. In (4) the first *aengx* (65.1) introduces a chronological development; the second (65.2) presents one more happening on top of the previous ones in a story.

(4) AS.065

65.1		
<i>Aengx taux da'nyeic hnoi,</i>		
and_then reach second day		
adv v n n		

65.2			
<i>aengx maaih nda'maauh faanv.</i>			
further have tiger be_disturbing			
adv vt n v.st			

‘^{65.1}And then another day, ^{65.2}furthermore there was a disturbing tiger.’

Figure 4(a) represents the conceptualization of *aengx* ‘furthermore’, where the CDS holds three viewing frames: minus, zero and plus frames (c.f. Langacker 2001:151), each

¹⁰ Besides (a) the unmarked action/motion verb, (b) the unmarked verb + post verb aspectual verb, another expression that pushes the story forward that should come under section 3.2 is (c) the repetition of the same unmarked verb, e.g. *he ran, ran, ran*.

of which represents a usage event (corresponding to the previous, current, and anticipated usage events of Figure 1). The development adverb *aengx* ‘furthermore’ in the zero frame (i.e. the frame getting the present attention) introduces and profiles (described by the heavy line) a relationship in which an additional event (E_2) follows from another (E_1). Further, E_1 is identified (connected by the dashed line) with the preceding event (E) in the minus frame (i.e. the prior frame in discourse). The relationship of the extended E_2 and the present E_1 , treated as “anchor-increment”, profiles an evocation of or expectation toward a further event that is inherent in the construction, i.e. *aengx* ‘furthermore’. Inevitably, the incremental nature of *aengx* increases another frame, the plus frame, containing E_2 identical with (depicted by the dashed line) the one in the zero frame.¹¹ The basic notion of “anchor-increment” is sketched in Figure 4(b) (Langacker 2000:265).

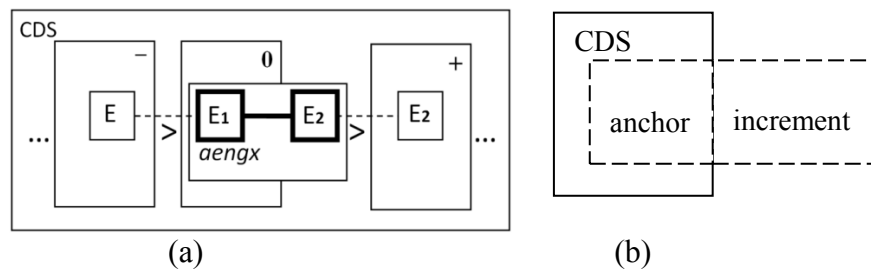


Figure 4. *aengx* ‘furthermore’ and incremental discourse space

The incremental characteristic of certain linguistic constructions discussed by Langacker under the notion “discourse expectations” (2001:151-53) is an important element that pushes the story forward. In the case of *aengx*, it has dual functions: prospective in that it evokes the subsequent discourse and retrospective in that it makes a relation to the prior discourse.

3.4 Serial Verb Constructions (SVCs) and Billiard-Ball Model

Thirdly, progression is most evidently expressed by a combination of Serial Verb Construction (SVC) and topic chain. Six defining factors of SVC are proposed by Aikhenvald (2005): 1) component verbs in SVC act as a single predicate, 2) monoclausality, 3) no intonation breaking between components, 4) shared tense, aspect, mood, modality and polarity value, 5) describes one event, 6) has one argument but may share arguments and obliques. An example of the simplest kind is (5).

(5) AS.008

<i>Nyornŋ</i>	<i>bieŋc</i>	<i>amz</i>	<i>mingh.</i>
creep	enter	granary	go
vi	vi	n	vi

‘They crept into the granary.’

¹¹ The translation value ‘again’ of *aengx* or the repetition of the same verb, rather than ‘furthermore’, is retrospect only. Hence, the retrospect use of *aengx* has a different conceptual structure than the dual functions; a possible diagram should neither have an increment arm of E_2 in the zero frame nor a dashed line toward the plus frame.

In (6), SVC is in the clause 23.2, in which the three verbs share one elliptical subject, i.e. the hero and his wife.

(6) AS.023

23.1

<i>Gengh</i>	<i>wuov_nzunc_hnoi</i>	<i>paaix</i>	<i>orv</i>	<i>daaih</i>
really	this_time	divide_up	meat	come
adv	adv.tmp	vt	n	v.asp

23.2

<i>dorh</i>	<i>mingh</i>	<i>bun</i>	<i>Am^pev.</i>
take	go	give	Ampe
vt	vi	vt	n.prp

^{23.1}Really this time (they) divided up the meat ^{23.2}(and) took it (to) give to the governor.'

Example (7) shows a case where an SVC has different subjects in it, which is referred to as “switch-subject SVC” (Aikhenvald 2005).

(7) AS.040

buonv daic ***mingh.***

shoot	die	go
vt	vi	v.asp

‘(Those who were inside the boat) shot (the crocodile and it) died.’

In order to account for conceptual structure of progression comprised of SVC and topic chains in the storyline, the notions of “action chain” and “billiard-ball model” are drawn upon from Langacker (1991c:282-329).¹² Figure 5 shows a transition of energy (indicated by the double arrows representing verbs) from one object (depicted by the circles representing nouns) to others in sequence. The energy transition starts from the first object named the Head and terminates in the last one termed the Tail.

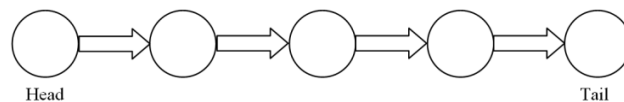


Figure 5. *Action chain in billiard-ball model*

As is easily observed, the action chain in the billiard-ball model is an extended chain of the Canonical Event Model in Figure 4. Whereas the term transitivity by Hopper and Thompson (1990) was restricted to lexical semantics and energy transmission within a clause, Langacker’s billiard-ball model can be applied to analyze SVCs, topic chains, and their combination in the form of multiclausal constructions (3.6).

¹² The notion of “action chain” can be said that it is the Langacker’s version of explaining transitivity in view of revealing its conceptual structure, which was discussed by Hopper and Thompson (1980) in relation to the foreground in discourse.

3.5 Topic Chains

Used in conjunction with SVCs, fourthly, topic chains are also substantially used in the storyline of Iu-Mien narratives. Topic chains are “chains of clauses which share a common topical participant” (Van Valin 2005:103). In (8), clauses 40.3 and 40.4 are the topic chain, followed by a switch-subject SVC with an aspectual verb.

(8) AS.040

40.1	40.2				40.3			
<i>Wuov_deix,</i>	<i>yiem</i>	<i>nzangv</i>	<i>gu'nyuoz</i>	<i>wuov_deix</i>	<i>zorqv</i>	<i>caang</i>	<i>nzopv,</i>	
those_people	be_in	boat	inside	those_people	take	spear	pierce	
pn.dem	vi	n	n	pn.dem	vt	n	vt	

40.4		40.5		
<i>zorqv congx</i>	<i>buonv,</i>	<i>buonv</i>	<i>daic</i>	<i>mingh.</i>
take gun	shoot	shoot	die	go
vt nvt	vt	vi	v.asp	

‘^{40.1}Those people, ^{40.2}those who were inside the boat, ^{40.3}took spears (and) pierced (crocodile), ^{40.4}took guns (and) shot, ^{40.5}shot (it and it) died.’

While SVCs give a story the sense of terseness and swiftness, the topic chains can elaborate the action and motions. The combined use of them forwards the story dynamically.

3.6 Multiclausal Constructions Containing SVCs and Topic Chains

The SVCs and topic chains can be combined to compose a long string of multiclausal construction as in (9a). An intonation break shows that it is composed of two sentences, but connected tightly describing one event comprising sub events. Clause 30.1 and Clause 30.2 are SVCs presenting a setting for Clause 30.3. There is no intonation break between Clauses 30.4-30.6, which comprise the three-member switch-subject SVCs. Clauses 30.3-30.6 is a topic chain.

(9a) AS.030

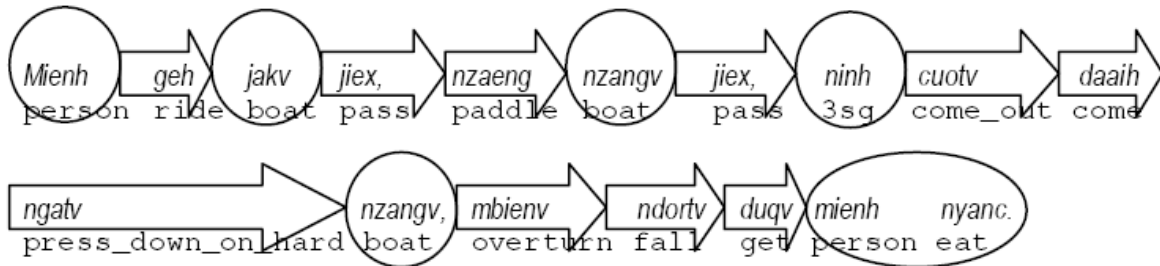
	30.1				30.2			
[St1]	<i>Mienh</i>	<i>geh</i>	<i>jakv</i>	<i>jiex,</i>	<i>nzaeng</i>	<i>nzangv</i>	<i>jiex,</i>	
	person	ride	boat	pass	paddle	boat	pass	
	n	vt	n	vi	v	n	vi	
	30.3							
	<i>ninh</i>	<i>cuotv</i>		<i>daaih</i>				
	pn.p	v		vi				
	3sg	come_out		come				
	30.4				30.5			
[St2]	<i>ngatv</i>			<i>nzangv,</i>	<i>mbienv</i>		<i>ndortv</i>	
	press_down_on_hard			boat	overturn		fall	
	vt			n	v		vi	

30.6
duqv mienh nyanc.
 get person eat
 vt n vt

^{30.1}People rode a boat to pass (the river), ^{30.2}paddled to pass, ^{30.3}it [the crocodile] came out ^{30.4}pressed down on the boat; ^{30.5}(the boat) turned over (and people) fell off, ^{30.6}(crocodile) got people (to) eat.'

(9b) is a schematic presentation of action chains mentioned in section 3.4. In (9b), "billiard-ball model" (3.4) is superimposed on the multicausal construction (9a), in which the circles represent NPs and the arrows Vs.

(9b) AS.030



This schematization is a simple technique to make the dynamism of action/event progression in the multicausal construction explicit. The longer the verbs are strung with interspersed NPs, the more vivid the actions become; the perception of the storyline thus becomes obvious.

An extended application of the billiard-ball model based on the cognitive psychological principle I (3.1) to both SVCs and topic chain is schematized in Figure 6. Circles represent nouns, double arrows verbs. Each time the verb is uttered by a narrator, an action occurs. As the action increases as indicated by a black arrow in each line along the axis Y, the mind of audience (i.e. hearer or reader) conceptualizes the motion of the black arrow as a progression of actions/events along the axis X.

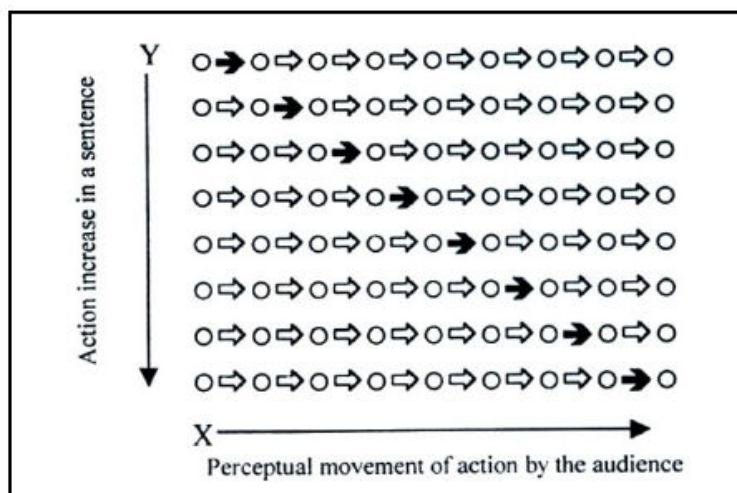


Figure 6. The conceptual schematization of action increase in SVCs and topic chains

3.7 Summary

Progression of a story is expressed in such constructions as unmarked action/motion verbs, the development adverb *aengx* ‘furthermore’, SVCs, topic chains, and the multiclausal construction containing SVCs and topic chains. The bundle of these constructions constitutes progression, i.e. advancement of action, energy, motion, and events, which forms the storyline. The progression in the storyline has been analyzed on the basis of the cognitive psychological principle I, i.e. a moving object is more conspicuous than a static one to the perception of the viewer. Taking a narrative as analogous to a theater play viewed by the audience, the storyline stands out as the foreground in the mental space of the hearer or reader through the progression energized by the above analyzed constructions. The clear and dynamic movement in the storyline is foregrounded in the visual perception in the mind of hearer/reader. In CL outlook on the narrative discourse, the storyline is an outcome of the energy, action, event progression.

4. Sequential Structures

Besides the progression, there is yet another factor by which storyline is foregrounded to the perception of the hearer or reader, that is, sequential structures.

4.1 Cognitive Psychological Principle II

A principle on which sequentiality is deemed as important to the foreground is schematized in Figure 7. That is, a condensed cluster of objects in a linear order (c) is more conspicuous to the perception of the viewer than scattered objects with no connection between them (a). This may be termed as cognitive psychological principle II.

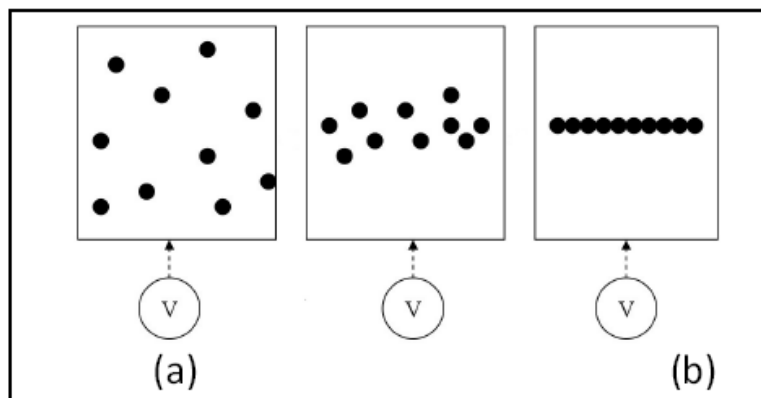


Figure 7. Scattered and condensed objects

Sequentiality is expressed by *ziouc* ‘then’, *cingx_daaih* ‘therefore’, *aeqv* ‘as for’, ad-clause containing *liuz* ‘after V-ing’, and independent clause marker *gau* ‘and then’. Four of them are selected to show the schematizations of their conceptual structures.

4.2 Sequential Marker (Prospective Conjunction) *ziouc* ‘then’

First, a very frequently used construction for connecting two clauses is S + *ziouc* + V, where *ziouc* can be translated as ‘and then, then soon, after which’. As in (10), *ziouc* always comes in the second position after S in the subsequent clause (126.2), which is connected to the preceding clause (126.1). With this grammatical behavior, *ziouc* can be

simply called a conjunction. At the same time, it gives a story a sense of quick transition of action/event across the clauses; hence its appropriate term may be the sequential marker. Example (10) shows a series of quick actions of a hero the main figure (126.2) after all soldiers fled (126.1).

(10) AS.126

126.1		126.2						
<i>Uv!</i>	<i>biaux</i>	<i>ninh</i>	<i>ziouc</i>	<i>zorqv</i>	<i>baeng</i>	<i>nyei</i>	<i>cong</i>	<i>daaih</i>
Oh!	flee	3sg	so_then	take	soldier	POSS	gun	come
intj	v	pn.p	seq.mk	vt	n	poss	n	v.dr

126.3

<i>buonv</i>	<i>nyei</i>	<i>buonv</i>
shoot	ADVBLZR	shoot
vt	part	vt

126.4

<i>nzuangv</i>	<i>nyei</i>	<i>nzuangv</i>	<i>zunc</i>	<i>jienv</i>	<i>mingh.</i>
aimless	ADVBLZR	aimless	chase	SML.ACT	go
v.st	part	v.st	vt	asp	vi

^{126.1}Oh! (they) ran away, ^{126.2}so he (immediately) took the soldier's guns ^{126.3}(and) shot repeatedly, ^{126.4}(and) chased them away furiously.'

In addition to expressing an immediate action following the first one, this construction also introduces an additional event in the course of happenings in succession as in (11).

(11) AS.071

<i>Wuov</i>	<i>hnoi</i>	<i>lungh_aanx</i>
that	day	after_noon
pn.dem	n	adv.tmp

<i>ziouc</i>	<i>haiz</i>	<i>wuov_jiex</i>	<i>laangz</i>	<i>mienh.</i>
so_then	hear	upper_side	village	people
seq.mk	vt	n	n	n

'In the afternoon of that day, then soon (they) heard (a sound of) the upper village people.'

Ziouc is more than a simple conjunction; it has a force to let a story advance. In parallel to *los* in the White Hmong, which means "(and) then/so, thus, therefore, yet, and it happens, and it turns out, with the result" (Clark 1988:93), *ziouc* in Iu-Mien seems to have the function of "an inchoative conjunction" (ibid). On this assumption, *ziouc* is considered to be an important element which belongs to the storyline. The "inchoative" semantic value is also testified by some native speakers of Iu-Mien, as saying that *ziouc* means "future".¹³

The conceptual structure of *ziouc* has much in common with that of *aengx* 'furthermore' (Figure 4). Figure 8 profiles the increment of E₂ upon the utterance of *ziouc*

¹³ Zanh Gueix-Fongc 2005: personal communication

in E₁ in the zero frame, which in turn evokes (indicated by the dashed line) the plus frame. What is different from *aengx* ‘furthermore’ in Figure 4 is that *ziouc* ‘and then’ lacks the anaphoric reference, hence Figure 8 does not have a dashed line which connects E and E₁.

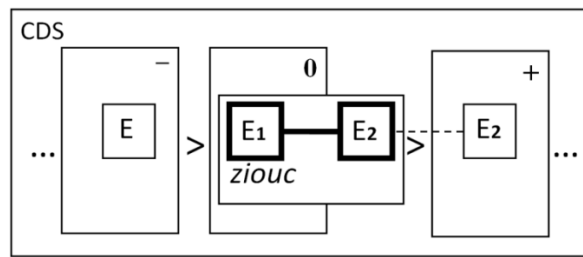


Figure 8. Prospective conjunction *ziouc* ‘and then’, ‘then soon’

While keeping connection with the previous clause, *ziouc* also pushes an event forward in CDS. Thus, the sequential marker *ziouc* can also be referred to as a prospective conjunction.

4.3 Retrospective Conjunction *cingx_daaih* “therefore”

The second material that ensures sequentiality is *cingx_daaih* ‘therefore’. Whereas *ziouc* is analyzed as the prospective conjunction, *cingx_daaih* ‘therefore’ is deemed as retrospective conjunction, thus, two are in the contrastive positions in their semantic value. (12a) and (12b) are connected discourse, where *cingx_daaih* bases (12b) on (12a). The clause after *cingx_daaih* presents a result of the previous sentence. (In fact, in this story *cingx_daaih* ‘therefore’, ‘as a result’ here is used humorously, since the hero is only accidentally honored by a success of others.)

(12a) AS.060

60.1	60.2		60.3				60.4	
<i>Wuov_deix</i>	<i>zorqv</i>	<i>caang</i>	<i>nzopv</i>	<i>nzopv</i>	<i>nzopv</i>	<i>nzopv</i>	<i>daic</i>	<i>mi'aqv.</i>
those_people	take	spear	pierce	pierce	pierce	pierce	die	RSLT
pn.dem	vt	n	vt	vt	vt	vt	vi	asp

^{60.1}Those people ^{60.2}took spears ^{60.3}pierced (and) pierced repeatedly (the snake) ^{60.4}(and it) died.

(12b) AS.061

<i>Cingx_daaih</i>	<i>aengx</i>	<i>orn</i>	<i>yietc</i>	<i>nzunc</i>
therefore	again	appoint	one	time
conj	adv	vt	numb	n

ninh zoux Aa^Han Nguv.
 3sg be Snake_the_Brave
 pn.p vi n.prp

Therefore, (the governor) again gave him another title (to) be Snake the Brave.

The conceptual structure of *cingx_daaih* ‘therefore’ is depicted in Figure 9.

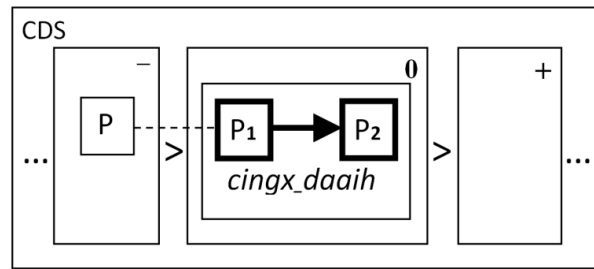


Figure 9. Retrospective conjunction *cingx_daaih* ‘therefore’

A parallelism with *therefore* in English analyzed by Langacker can be recognized to explain Figure 9. The following quotation is applicable to *cingx_daaih*: “*therefore* [and *cingx_daaih*] introduces and profiles a relationship in which one proposition, labeled P₂, follows from another, P₁. Moreover, P₁ is identified as the proposition expressed in the previous usage event, just earlier in the flow of discourse” (Langacker 2001:149).

It should be noted that *cingx_daaih* often emphasizes a logical consequence of the preceding preposition, in contrasted to the construction V + *liuz* ‘after V-ing’ which signifies temporal transition (4.4). As such, P₁ and P₂ are confined in the zero frame and there is no increment toward the plus frame.

4.4 Ad-Clause V + *liuz* “after V-ing”

The third construction to contribute to the storyline is an ad-clause containing the phrase V + *liuz* ‘after V-ing’, where *liuz* itself originally means ‘to finish’. The ad-clause always comes in the sentence initial position followed by the main clause. Compare (13a-d). They are not from the story in the appendix but all from actual situations except for (13a) which is unnatural.

(13a) *nyanc **liuz** hnaangx
 eat finish rice
 vt aux n
 (unnatural)

(13b) Nyanc **liuz** hnaangx aqv
 eat finish rice PFT
 vt aux n part.asp
 ‘I have finished having a meal.’

(13c) Nyanc **liuz** nhaangx, (yie) cingx_daaih daaih
 eat finish rice (1sg) therefore come
 vt aux n pn.p conj vi
 ‘I have finished having a meal, so I came.’

(13d) nyanc **liuz** nhaangx, (yie) ziouc daaih
 eat finish rice (1sg) consequently come
 vt aux n pn.p conj vi

‘Having finished a meal, I came right away.’

The native speakers’ response to (13a) would be “and then what happened?” This means that *nyanc liuz hnanx* is just a fragment until either it is completed by an aspectual particle *aqv* as in (13b) or followed by a main clause as in (13c-d). Both the narrator and hearer cannot stop after the *liuz*-clause. This ad-clause inevitably evokes an expectation in hearer’s mind.

Reutilizing the discourse space increment diagram from Figure 4(b), the conceptual structure of *liuz*-clause can be explained as in Figure 10. In the ad-clause *V + liuz* ‘after V-ing’, the *V* carries the old information and the whole construction projects a new discourse space for more information.

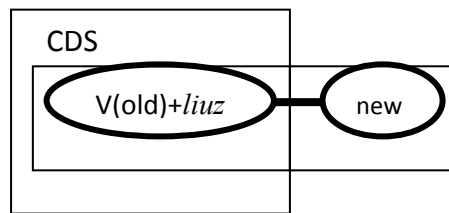


Figure 10. Discourse space increment through ad-clause *V + liuz* ‘after V-ing’

Thus, the construction is a prospective element in a similar vein with *ziouc* ‘and then (soon)’ in Figure 8. The effect on the audience in hearing and reading the ad-clause that contains *liuz* is to cause an expectation or a whetting of the appetite to listen for more. Naturally, it has an impact on the audience in creating a mental space even before hearing what comes next. As a result, the use of *liuz* pushes the story forward.

4.5 Topic marker *aeqv* “as for”

In the fourth place, a topic marker *aeqv* ‘as for’ is also an important element in the sequential structure. This is true when the above discussed retrospective and prospective elements (i.e. *cingx_daaih* ‘therefore’ and *ziouc* ‘and then’) are taken in consideration to understand sequentiality of storyline. The main function of topicalization, according to Levinsohn, is to indicate discontinuities while maintaining “the overall unity and continuity” (Levinsohn 1994:7). He says: “This is because topicalized constituents have a BIDIRECTIONAL function: (a) they serve as a POINT OF DEPARTURE for what follows and (b) they indicate the primary BASIS for linking what follows to its context”. In Examples (14)-(16), the underlined parts are topicalized by the topic mark *aeqv*; the clause after them (i.e. (113.2), (36.2), and (96.3)) are all main clauses.

The first example (14) has a personal pronoun topicalized; (15) exemplifies a verb phrase topicalized.

(14) AS.113

113.1	113.2			
<u><i>Ninh aeqv</i></u> ,	<i>sin</i>	<i>zinx</i>	<i>haic</i>	<i>aav lov!</i>
3sg TOP	body	tremble	very	SURPRISE
pn.p top	n	vi	adv	part.f

^{113.1}As for him, ^{113.2}his body trembled terribly.

(15) AS.036

36.1			36.2				
<i>Mingh</i>	<i>taux</i>	<i>aeqv</i> ,	<i>aengx</i>	<i>nzaeng</i>	<i>jienv</i>	<i>nzangv</i>	<i>jiex</i> .
go	reach	TOP	again	paddle	SML.ACT	boat	pass
vi	v	top	adv	v	asp	n	vi

^{36.1}As (they) reach there, ^{36.2}(they) again passed (the river) paddling a boat.

Not only topicalizing a noun phrase and a verb phrase, but also multiple phrases can be topicalized as in (16). Each time *aeqv* is uttered, the floor of narration is held until the whole sentence is concluded with the main clause.

(16) AS.096

96.1			96.2		
<i>Mv</i>	<i>zeiz</i>	<i>nor</i>	<i>aeqv</i> ,	<i>mv</i>	<i>hingh</i>
NEG	be	as	TOP	NEG	win
neg	vi	top	top	neg	vt

96.3

<i>mbuo</i>	<i>naaiv</i>	<i>mv</i>	<i>maaih</i>	<i>faatv</i> .
1pl	here	NEG	have	magic
pn.p	pn.dem	neg	vt	n

^{96.1}If we don't, ^{96.2}(we) can't win and then ^{96.3}we do not have magic here.

The held floor of narration forces the narrator to continue to tell story and the hearer to expect more to listen. Evidently, the use of the topic marker *aeqv* let the story go on.

The conceptual structure of the bi-directional function of topic marker *aeqv* 'as for', which evokes an upcoming proposition creating a new discourse space, is illustrated in Figure 11 following Langacker 2001.

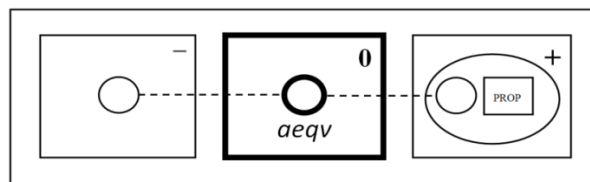


Figure 11. Topic marker *aeqv* 'as for'

A lengthy quotation from Langacker well explains Figure 11:

A topic marker refers schematically to the *thing* profiled by the noun phrase with which it combines; the nominal referent is represented as a circle in the focus frame. Such marker is prospective by virtue of signaling that the profiled entity will function as a conceptual *reference point* for purposes of interpreting a subsequent proposition, as shown in the plus frame. But it is also retrospective, in the sense that the topic needs to be an entity already accessible in the prior discourse. Thus the profile is shown as corresponding to a thing present in the minus frame. (Langacker 2001:152)

4.6 Summary

The sequential structure premises the analyses on the cognitive psychological principle II: a condensed cluster of objects in a linear order is more conspicuous to the perception of the viewer (i.e. hearer/reader of narrative) than scattered objects with no connection between them. On this premise, the linguistic constructions that have functions of combining and sequential nature have been analyzed. The prospective conjunction *ziouc* ‘and then (soon)’ and the retrospective conjunction *cingx_daaih* ‘therefore’ ensure the flow of discourse in sequence. Besides them, the topic marker *aeqv* ‘as for’ has the bi-directional function, which provides the sequentiality for the story as well as preventing the narration from stopping by its floor-holding function. The ad-clause containing V-*liuz* ‘after V-ing’ has the incremental nature of discourse space. This construction not only keeps sequentiality in the storyline but causes the hearer to expect to listen for more. Thus, the sequential structure makes the storyline obvious as the foreground to the visual perception in the mind of the hearer/reader.

5. Conclusion

The term “foreground” is itself pertains to an act of viewing: it is the locus of a view which is the nearest to an observer, hence the most salient or noticeable. Though the term has been used to describe the event-line/mainline/storyline of the narrative discourse in numerous literatures, the optical notion of the foreground has never been highlighted in the study of storyline. CL including CG and cognitive semantics, however, assumes that the narrative discourse is a cognitive entity viewed in the mind of the conceptualizer, i.e. viewer/hearer/reader. In the course of hearing/reading the narrative, by analogy with an act of viewing a play on the stage, the hearer/reader perceives the storyline as the foreground in the mental space. The storyline is optically foregrounded as an outcome of two main components of the narrative: progression and sequentiality.

The progression of the storyline comprises such constructions as the unmarked action/motion verbs, the development adverb *aengx* ‘furthermore’, SVCs, topic chains, and the multiclausal construction containing SVCs and topic chains. They advance the story. The sequential structures include such elements as the prospective conjunction *ziouc* ‘and then (soon)’, the retrospective conjunction *cingx_daaih* ‘therefore’, the topic marker *aeqv* ‘as for’ with the bi-directional function, and the ad-clause containing V-*liuz* ‘after V-ing’. The storyline encoded with these constructions of the forward-moving and the tightly sequenced becomes easily observable in front of the mental eyes of the hearer/reader, i.e. the viewer. In other words, progression and sequentiality in the narrative make the storyline emerge to the foreground in the perception of the viewer.

This approach works well for Iu-Mien which totally lacks morphological markers in the verb, as well as revealing semantics of storyline. It is also expected that a CL approach can assist in discourse analysis in various languages in Eastern and Southeast Asia, such as within Chinese, Tai, Hmong-Mien, and Vietnamese languages.¹⁴ As these languages are areally similar in their structure (e.g. little morphological inflection), adopting a similar cognitive-semantic approach can be illuminating.

¹⁴ Purnell 2009: personal communication

References

- Aikhenvald, Alexandra Y. 2005. *Serial verb constructions*. Oxford: Oxford University Press.
- Clark, Marybeth. 1988. An inchoative conjunction in Hmong: Extra-sentential topic marker? Paper presented at the International Symposium on Language and Linguistics, Faculty of Liberal Arts, Thammasat University, Bangkok, Thailand, 9-11 August 1988, ed. By Cholticha Bamroongraks et al., 93-106. Bangkok: Faculty of Liberal Arts, Thammasat University.
- Dooley, Robert A., and Stephen H. Levinsohn. 2001. *Analyzing discourse: A manual of basic concepts*. Dallas: SIL International.
- Dry, Helen Arister. 1981. Sentence aspect and the movement of narrative time. *Text* 1(3):233-40.
- _____. 1983. The movement of narrative time. *The journal of literary semantics* 12(2):19-54.
- _____. 1992. Foregrounding: an assessment. *Language in context: essays for Robert E. Longacre*, ed. by Shin Ja J. Hwang and William R.
- Endo, Yoshinobu. 1996. *The verbal system of Classical Hebrew in Joseph story: An approach from discourse analysis*. Assen: Van Gorcum & Comp.
- _____. 2003. Problems in the 'Grounding-theory' of R. E. Longacre and others in the study of Hebrew narrative discourse. *Exegetica: Studies in Biblical exegesis* 14: 77-90. Tokyo: Biblical Exegesis Study Group in Japan.
- Gärdenfors, Peter. 1999. Some tenets of cognitive semantics. *Cognitive semantics: Meaning and cognition*, ed. by Jens Allwood and Peter Gärdenfors, 19-36. Amsterdam: John Benjamins.
- Givón, Talmy. 1983. Topic continuity in discourse: An introduction. *Topic continuity in discourse*, ed. by Talmy Givón, 1-41. Amsterdam: John Benjamins.
- _____. 1984. *Syntax: A functional-typological introduction*, Vol. 1. Amsterdam: John Benjamins.
- Hopper, Paul J. and Sandra A. Thompson. 1980. Transitivity in grammar and discourse. *Language* 56:2. 251-99.
- Levinsohn, Stephen H. 1994. Discontinuities in coherent texts. *Discourse features of ten languages of West-Central Africa*, ed. by Stephen H. Levinsohn, 3-14. Dallas: The Summer Institute of Linguistics and The University of Texas at Arlington.
- Langacker, Ronald W. 1987. *Foundations of Cognitive Grammar. Vol. I: Theoretical prerequisites*. Stanford, CA: Stanford University Press.
- _____. 1991a. Cognitive Grammar. *Linguistic theory and grammatical description*, ed. by Flip G. Droste and John E. Joseph, 274-306. Amsterdam: John Benjamins.
- _____. 1991b. *Concept, image, and symbol: The cognitive basis of grammar*. Berlin: Mouton de Gruyter.
- _____. 1991c. *Foundations of Cognitive Grammar. Vol. II: Descriptive application*. Stanford, CA: Stanford University Press.
- _____. 2000. *Grammar and conceptualization*. Berlin: Mouton de Gruyter.
- _____. 2008. *Cognitive Grammar: A basic introduction*. Oxford: Oxford University Press.

- Longacre, Robert E. 1981. A spectrum and profile approach to discourse analysis. *Text* 1:4.337-59.
- _____. 1983. *The grammar of discourse*. 1st ed. New York: Plenum Press.
- _____. 2003a. Joseph: a story of divine providence—a text theoretical and textlinguistic analysis of Genesis 37 and 39-48, 2nd ed. Winona Lake: Eisenbrauns.
- _____. 2003b. Holistic textlinguistics. SIL Electronic working papers. n.p.: SIL International.
- Somsong Burusphat. 1990. The bipartite structure of information in Thai narrative discourse. [Language and Linguistics (*Phaasaa lae phaasaasaat*)] 9:1.65-78. (ภาษาและภาษาศาสตร์ ปีที่ 9 ฉบับที่ 1 กรกฎาคม-ธันวาคม 2533 สาขาวิชา ภาษาศาสตร์ คณะศิลปศาสตร์ มหาวิทยาลัยธรรมศาสตร์) Bangkok: Thammasat University.
- _____. 1992. The identification of storyline in Thai narrative discourse. *Language in context: Essays for Robert E. Longacre*, ed. by Shin Ja J. Hwang and William R. Merrifield, 419-33. Dallas: The Summer Institute of Linguistics and The University of Texas at Arlington.
- _____. 2002. The temporal movement of the Hlai (Li) origin myth. *Collected papers on Southeast Asian and Pacific Languages*, ed. by Robert S. Bauer, 141-51. Canberra: Pacific Linguistics.
- Talmy, Leonard. 2000. *Toward a cognitive semantics*, Vol. 2: Typology and process in concept structuring. Cambridge, Massachusetts: The MIT Press.
- Van Valin, Robert D., Jr. 2005. *Exploring the syntax-semantics interface*. Cambridge: Cambridge University Press.

Appendix: Text of “A Story of Aahan” (AS)

Aa^han Gouv

Yunh Zoih gorngv
 Yunzoi speak
 n.prp v

Mr. Yunzoi narrated.

AS.001

1.1				1.2		
<i>Maaih i_gox</i>		<i>mienh</i>		<i>za'gengh</i>	<i>jomc</i>	<i>nyei.</i>
have married_couple		person		really	be_poor	PRS.ST
v n		n		adv	v.st	asp

^{1.1}There was a couple ^{1.2}(who were) very poor.

AS.002

2.1		2.2			2.3		
<i>Jomc</i>	<i>nyei,</i>	<i>ninh</i>	<i>mingh</i>	<i>maaic</i>	<i>gong</i>	<i>lorz</i>	<i>hnaangx nyanc.</i>
be_poor	PRS.ST	3sg	go	sell	work	look_for	rice eat
v.st asp	pn.p vi	vt	n	vt	n	vt	

^{2.1}Poor they were, ^{2.2}they went (and) did labor for other people ^{2.3}looked for rice (to) eat [i.e. made their living.]

AS.003

3.1						
<i>l_gox</i>		<i>ndaam jienv</i>		<i>ndaamx</i>		
married_couple		carry SML.ACT		carrying_pole		
n		vt asp		n		

3.2			3.3				
<i>hnaangv Janx-Kor^lormx</i>		<i>nor</i>	<i>hnaeng,</i>	<i>hnaeng</i>	<i>hnaeng</i>	<i>hnaeng</i>	<i>mingh.</i>
as-if Northern Thai		as	dangle	dangle	dangle	dangle	go
adv n		top	v	v	v	v	vi

^{3.1}The couple would carry baskets on their shoulders with poles ^{3.2}just like Northern Thai would do, ^{3.3}(they) went dangling, dangling and dangling.

AS.004

4.1	4.2		
<i>Mingh gau</i>	<i>hmuangx</i>		<i>aqv.</i>
go DPCL	dark		CHG-O-ST
vi part	adj		part.asp

^{4.1}As they went, ^{4.2}(it) became dark.

AS.005

5.1		5.2			
<i>Hmuangx</i>	<i>aqv,</i>	<i>taux</i>	<i>wuov</i>	<i>ndaamv-jauv,</i>	
dark	CHG-O-ST	reach	there	half-way	
adj	part.asp	v	pn.dem	n	

5.3				5.4				
<i>hnaangv yie</i>	<i>nyei</i>	<i>lamz</i>	<i>nor</i>	<i>zoux</i>	<i>daaih</i>	<i>lauh</i>	<i>haic</i>	<i>aqv.</i>
like 1SG	POSS	granary	as	make	come	long-period	very	CHG-O-ST
adv pn.p	poss	n	top	vt	v.asp	v.st	adv	part.asp

^{5.1}It became dark ^{5.2}(when they) were on the way, ^{5.3}(then there was) a granary just like mine, ^{5.4}it has been built for long time.

AS.006

6.1		6.2	
<i>Taux</i>	<i>wuov</i>	<i>hmuangx</i>	<i>aqv.</i>
reach	there	dark	CHG-O-ST
v	pn.dem	v.st	part.asp

^{6.1}(When they) arrive there ^{6.2}(it) became dark.

AS.007

7.1		7.2							
<i>Aav,</i>	<i>mingh</i>	<i>haaix</i>	<i>yaac</i>	<i>mv</i>	<i>mingh</i>	<i>taux</i>	<i>bieqc</i>	<i>lamz</i>	<i>bueix.</i>
Ah	go	what	and_yet	NEG	go	reach	enter	granary	sleep
intj	vi	q	conj	neg	vi	v	vi	n	vi

^{7.1}Ah, wherever (they) went ^{7.2}(they) have not gone as far as to enter a granary (to) sleep.

AS.008

<i>Nyong</i>	<i>bieqc</i>	<i>lamz</i>	<i>mingh.</i>
creep	enter	granary	go
vi	vi	n	vi

They crept into the granary.

AS.009

<i>Jiepv</i>	<i>njiec</i>	<i>dorn.</i>
bear	give_birth_to	cub
n	vt	n

A bear had given birth to a cub (there).

AS.010

10.1				10.2		
<i>Jiepv</i>	<i>za'gengh</i>	<i>tiux</i>	<i>cuotv</i>	<i>oix</i>	<i>ngaac</i>	<i>ninh_mbuo.</i>
bear	really	run	go_out	want-to	bite	3_PL
n	adv	vi	v	v.mod	vt	pn.p

^{10.1}The bear really ran out ^{10.2}(it) wanted (to) bite the couple.

AS.011

<i>Ninh</i>	<i>nyei</i>	<i>nqox</i>	<i>biaux!</i>	<i>biaux.</i>
3sg	POSS	husband	flee	flee
pn.p	poss	n	vi	vi

Her husband ran away.

AS.012

12.1				12.2	
<i>Ninh</i>	<i>nyei</i>	<i>auv</i>	<i>maiv_haih_func_zoux,</i>	<i>maiv</i>	<i>henv.</i>
3sg	POSS	wife	not_be_able_to_anything	NEG	strong
pn.p	poss	n	idm	neg	v.st

^{12.1}There was nothing his wife could do, ^{12.2}(she) was not strong.

AS.013

<i>Ndaam</i>	<i>jienv</i>	<i>ga'naaiv.</i>
carry	DUR	thing
vt	asp	n

(She) was carrying things.

AS.014

14.1

<i>Hnungh</i>	<i>hnaeng!</i>	<i>hnungh</i>	<i>hnaeng</i>
ONOMA	dangle	ONOMA	dangle
onoma	v	onoma	v

14.2

<i>ziouc</i>	<i>zorqv</i>	<i>ninh</i>	<i>neyi</i>	<i>ndaamx</i>	<i>hngv</i>	<i>nv</i>	<i>nor,</i>
so_then	take	3sg	POSS	carrying_pole	like	this	as
seq.mk	vt	pn.p	poss	n	adv	pn.dem	top

14.3

<i>se</i>	<i>ninh</i>	<i>neyi</i>	<i>biaav-ndaav</i>	<i>neyi,</i>
that_is	3sg	POSS	stick-long	PRS.ST
copl	pn.p	poss	n-v.st	asp

14.4

<i>bungx</i>	<i>njiec</i>	<i>daaih</i>	<i>ndamc</i>	<i>mv</i>	<i>bung</i>	<i>ndutv</i>	<i>mingh,</i>
release	let_go_down	come	stamp	this	end	come_apart	go
vt	vt	v.dr	vt	pn.dem	n	vi	v.asp

14.7

<i>ndamc</i>	<i>mv</i>	<i>bung</i>	<i>ndutv</i>	<i>mingh</i>	<i>nzopv</i>	<i>jielv.</i>
stamp	this	end	come_apart	go	go	bear
vt	pn.de	n	vi	vi	vt	n

^{14.1}(She) swung (it) dangling, dangling, ^{14.2}then (she) took her carrying pole like this, ^{14.3}that is, her pole was long, ^{14.4}(she) released (it and) let (it) go down ^{14.5}(and) stamped on this end ^{14.6}(and it) came off; ^{14.7}(she) stamped on this end ^{14.8}let it drop (on the bear) ^{14.9}(and it) pierced the bear.

AS.015

15.1

<i>Nzopv</i>	<i>jielv.</i>	<i>nzopv</i>	<i>gau</i>
pierce	bear	pierce	DPCL
vt	n	vt	part

15.3

<i>nzopv!</i>	<i>nzopv</i>	<i>nzui</i>	<i>gau</i>	<i>jielv</i>	<i>neyi</i>	<i>nzuih_baengx,</i>	<i>daic</i>	<i>mi'aqv.</i>
pierce	pierce	shove	DPCL	bear	POSS	mouth	die	RSLT
vt	vt	v	part	n	poss	n	vi	asp

^{15.1}(She) pierced into the bear('s mouth) ^{15.2}as (she) pierced, ^{15.3}(she) shoved the pole really hard ^{15.4}(and the bear) died.

AS.016

<i>Wuov_nzunc_hnoi</i>	<i>i_gox</i>	<i>caangh_laangh</i>	<i>aqv.</i>
this-time	married_couple	discuss	PFT
adv.tmp	n	v.cmp	asp

Then the couple discussed.

AS.017

17.1

<i>Ninh</i>	<i>oix</i>	<i>paaix</i>	<i>nc</i>	<i>dauh</i>	<i>jielv</i>	<i>daaih</i>	<i>dorh</i>	<i>mingh</i>	<i>bun</i>	<i>Am^pev.</i>
3sg	want_to	divide	that	CLF	bear	come	take	go	give	Ampe
pn.p	v.mod	vt	pn.dem	clf	n	v.asp	vt	vi	vt	n.prp

^{17.1}He wanted to divide the bear (to) take it ^{17.2}(and) give to the district township official.

AS.023

23.1						23.2			
<i>Gengh</i>	<i>wuov_nzunc_hnoi</i>	<i>paaix</i>		<i>orv</i>	<i>daaih</i>	<i>dorh</i>	<i>mingh</i>	<i>bun</i>	<i>Am^pev.</i>
really	this_time	divide_up	meat	come	take	go	give	Ampe	
adv	adv.tmp	vt	n	v.asp	vt	vi	vt	n.prp	

^{23.1}Really this time (they) divided up the meat ^{23.2}(and) took it (to) give to the governor.

AS.024

24.1		24.2							
<i>Taux</i>	<i>wuov</i>	<i>Am^pev</i>	<i>naaic</i>	<i>gaax</i>		<i>ninh_mbuo</i>			
reach	there	Ampe	ask	try_and_see	3_PL				
v	pn.dem	n.prp	v	part		pn.p			

" <i>Mv</i>	<i>maaih</i>	<i>congx,</i>	<i>mv</i>	<i>maaih</i>	<i>nzuqc</i>	<i>hnangv_haaix_nor</i>	<i>duqv</i>	<i>jjepv?"</i>	
NEG	have	gun	NEG	have	knife	how	get	bear	
neg	vt	n	neg	vt	n	adv	vt	n	

^{24.1}(They) reached there, ^{24.2}the governor asked them, "You don't have a gun, don't have a knife. How could you get the bear?"

AS.025

" <i>Oh</i>	<i>yie_mbuo</i>	<i>mborqv</i>	<i>daaih.</i>	<i>Caa</i>	<i>jienv</i>	<i>mborqv</i>	<i>nyei</i>	<i>mborqv.</i>
Oh	1_PL	hit	come	catch	SML.ACT	hit	ADVBLZR	hit
intj	pn.p	vt	v.asp	vt	asp	vt	part	vt

<i>Dorh</i>	<i>biaav</i>	<i>nzuiz</i>	<i>nyei</i>	<i>nzuiz</i>	<i>nzuih_baengx</i>	<i>daic</i>	<i>mingh."</i>	
take	stick	shove	ADVBLZR	shove	mouth	die	go	
vt	n	v	part	v	n	vi	v.asp	

"Oh, we beat (it to death). (We) caught it (and) beat and beat. (We) took a stick and shoved (it) into its mouth repeatedly (and it) died."

AS.026

" <i>U...</i>	<i>Meih</i>	<i>za'gengh</i>	<i>henv</i>	<i>haic."</i>
Wow!	2sg	really	strong	very
intj	pn.p	adv	v.st	adv

"Wow! You are really so strong!"

AS.027

<i>Heuc</i>	<i>naaic</i>	<i>dauh</i>	<i>mienh</i>	<i>zoux</i>	<i>Aa^Han_Mix.</i>
name	that	CLF	person	be	Bear_the_Brave
vt	pn.dem	clf	n	vi	n.prp

(The governor) entitled that man (to) be Bear the Brave.

AS.028

<i>Wuov_nzunc_hnoi</i>	<i>duqv</i>	<i>yietc</i>	<i>norm</i>	<i>mbuox</i>	<i>hlo</i>	<i>deix</i>	<i>aqv.</i>
this-time	get	one	CLF	name	big	some	PFT
adv.tmp	vt	numb	clf	n	v.st	adv	part.asp

At this occasion (he) got one considerably big name.

AS.029

29.1		29.2				29.3			
<i>Da'nyeic</i>	<i>hnoi</i>	<i>hnangv</i>	<i>wuov</i>	<i>ndiev</i>	<i>ndaaih</i>	<i>naaic,</i>	<i>mbuo</i>	<i>geh</i>	<i>jakv</i>
second	day	like	that	under	river	there	lpl	ride	boat
n	n	adv	pn.dem	adv.loc	n	pn.dem	pn.p	vt	n

29.4
maaih douh_taanh yietc nyungc, hlo nyei.
 have big_lizard one kind big PRS.ST
 vt n numb n v.st asp

29.1 Then another day, ^{29.2}like the river down there ^{29.3}(where) we ride a boat, ^{29.4}there was a kind of crocodile, ^{29.5}(which) was big.

AS.030

30.1
Mienh geh jakv jiex, nzaeng nzangv jiex,
 person ride boat pass paddle boat pass
 n vt n vi v n vi

30.3
ninh cuotv daaih ngatv nzangv, mbienv ndortv
 3sg come_out come press_down_on_hard boat overturn fall
 pn.p v vi vt n v vi

30.6
duqv mienh nyanc.
 get person eat
 vt n vt

^{30.1}People rode a boat to pass (the river), ^{30.2}paddled to pass, ^{30.3}it [the crocodile] came out ^{30.4}pressed down on the boat; ^{30.5}(the boat) turned over (and people) fell off, ^{30.6}(crocodile) got people (to) eat.

AS.031

Janx aengx mv haih fungc zoux.
 non_Iu-Mienh again NEG be_able_to how-come do
 n adv neg aux adv vt

There was nothing the Northern Thai people could do about it again.

AS.032

Aav mv noic duqv aqv.
 Ah NEG manage can PFT
 intj neg vi v.aux part.asp

Ah, (we) can't manage to do anything.

AS.033

Aqv_zuqc mingh lorz naaic dauh Aa^Han_Mix daaih zorqv hnanv.
 must go look_for that CLF Bear_the_Brave come take no_other_way
 aux vi vt pn.dem clf n.prp v.dr vt adv

(We) must go find that Bear the Brave and bring him here to let him catch (the crocodile).

AS.034

Za'gengh mingh lorz.
 really go look_for
 adv vi vt

(They) really went to look for (him).

AS.035

Lorz buatc nc dauh Aa^Han_Mix.
 look_for see that CLF Bear_the_Brave
 vt v pn.dem clf n.prp

(They) found the Bear the Brave.

AS.036

36.1				36.2					
<i>Mingh</i>	<i>taux</i>	<i>aeqv.</i>	<i>aengx</i>	<i>nzaeng</i>	<i>jienv</i>	<i>nzangv</i>	<i>jiex.</i>		
go	reach	TOP	again	paddle	SML.ACT	boat	pass		
vi	v	top	adv	v	asp	n	vi		

^{36.1}As (they) reach there, ^{36.2}(they) again passed (the river) paddling a boat.

AS.037

37.1				
<i>Jiex</i>	<i>taux</i>	<i>ndaaih</i>	<i>mbu'ndongx</i>	<i>hnangv,</i>
pass	reach	river	middle	only
vi	v	n	n	adv

37.2						37.3		37.4	
<i>douh_taanh</i>	<i>za'gengh</i>	<i>hmatv</i>	<i>jieqv</i>	<i>nyei,</i>	<i>jangv</i>	<i>nyei</i>	<i>cuotv</i>	<i>daaih.</i>	
big_lizard	really	absolutely	black	PRS.ST	wide	PRS.ST	come_out	come	
n	adv	part	v.st	asp	v.st	asp	vi	vi.dr	

^{37.1}As soon as (the boat) passed (to) reach the middle of the river, ^{37.2}the crocodile, (which was) really absolutely black ^{37.3}(and) wide, ^{37.4}came out.

AS.038

38.1			38.2		38.3			
<i>Ninh</i>	<i>gamh_nziex</i>	<i>haic</i>	<i>sin</i>	<i>zinx</i>	<i>nyanh</i>	<i>nv</i>	<i>nyanh</i>	<i>nyei.</i>
3sg	fear	very	body	tremble	jerk	intensely	jerk	PRS.ST
pn.p	v.comp	adv	n	vi	vi	intj	vi	asp

^{38.1}He was very afraid, ^{38.2}(his) body trembled, ^{38.3}jerked intensely.

AS.039

<i>Nqaiiv</i>	<i>cuotv.</i>
feces	come_out
n	vi

Feces came out.

AS.040

40.1		40.2			40.3			
<i>Wuov_deix,</i>	<i>yiem</i>	<i>nzangv</i>	<i>gu'nyuoz</i>	<i>wuov_deix</i>	<i>zorqv</i>	<i>caang</i>	<i>nzopv,</i>	
those_people	be_in	boat	inside	those_people	take	spear	pierce	
pn.dem	vi	n	n	pn.dem	vt	n	vt	

40.4		40.5			
<i>zorqv</i>	<i>congx</i>	<i>buonv,</i>	<i>buonv</i>	<i>daic</i>	<i>mingh.</i>
take	gun	shoot	shoot	die	go
vt	n	vt	vt	vi	v.asp

^{40.1}Those people, ^{40.2}those who were inside the boat, ^{40.3}took spears (and) pierced (it), ^{40.4}took guns (and) shot, ^{40.5}shot (it and it) died.

AS.041

<i>"Meih_mbuo</i>	<i>mv_dungx</i>	<i>gaengh</i>	<i>baqv</i>	<i>saa.</i>
2_PL	NEG.CMD	really	pierce	SFT.CMD
pn.p	neg	adv	vt	part

"Don't you yet pierce it."

<i>Yie</i>	<i>jaang-jaang</i>	<i>bungx</i>	<i>nqaiiv</i>	<i>bun</i>	<i>ninh</i>	<i>nyanc.</i>
1sg	on_the_process	release	feces	let	3sg	eat
pn.p	adv	vt	n	vt	pn.p	vt

"I'm just releasing feces to let it eat."

Yie oix zorqv nangh.
 1sg want_to take be_alive
 pn.p v.mod vt v.st

“I want to catch it alive.”

Oix zorqv ndoh jienv dorh mingh fongc guen jien.
 want_to take bind DUR take go show_gratitude province officer
 v.mod vt vt asp vt vi vt n n

“I want to take it bound to show gratitude to the provincial officer.”

Meih_mbuo fungc zorqv baqv daic mingh?"
 2_PL how_come take pierce die go
 pn.p adv vt vt vi v.dr

“How come you guys pierced it to death?”

AS.042

Hnangv nc nor gomgv.
 like that as speak
 adv pn.dem adv v

He said like that.

AS.043

Hmz, wuov_nzunc_hnoi aengx mingh aqv.
 Yes! This_time again go CHG-O-ST
 intj adv.tmp adv vi part.asp

Yes, this time (he)'s got to go again.

AS.044

Dorh jienv nv dauh douh taanh mingh Am^pev wuov.
 take SML.ACT this CLF big_lizard go Ampe place
 vt asp pn.dem clf n vi n.prp pn.dem

(He) taking this crocodile (and) went to the provincial office.

AS.045

Aengx om ninh zoux Aa^Han_Suv^Tiqc.
 further_more appoint 3sg be Suthi the Brave
 adv vt pn.p vi n.prp

(The governor) appointed him again (to) be Suthi the Brave.

AS.046

46.1 Duqv i norm mbuox Aa^Han_Mix, Aa^Han_Suv^Tiqc 46.2 duqv daaih aqv.
 get two CLF name Bear_the_Brave Suthi_the_Brave get come PFT
 vt numb clf n n.prp n.prp vt v.asp part.asp

^{46.1}(He) has got two names, Bear the Brave (and) Suthi the Brave; ^{46.2}(he) got (them).

AS.047

47.1 Da'nyeic hnoi youc maaih diuh domh naang, hlo nyei,
 second day also have CLF big snake big PRS.ST
 n n seq.mk vt clf adj n v.st asp

47.3 ndortv wuov laangz mienh, yietc norm mungv,
 fall that village person one CLF town
 vi pn.dem n n numb clf n

<i>yietc</i>	<i>norm</i>	<i>zingh</i>	<i>nyei</i>	<i>mienh</i>	<i>nyei</i>	<i>wuom-kuotv.</i>
one	CLF	city	POSS	people	POSS	water_source
<i>nmb</i>	<i>clf</i>	<i>n</i>	<i>poss</i>	<i>n</i>	<i>poss</i>	<i>n</i>

^{47.1}Then another day, ^{47.2}there was a gigantic snake, (which) was big ^{47.3}fell into a city water source for the residents of the whole town, the whole city.

AS.048

<i>Za'gengh</i>	<i>mv</i>	<i>duqv</i>	<i>wuom</i>	<i>nyanc</i>	<i>aqv.</i>
really	NEG	get	water	eat	CHG-O-ST
<i>adv</i>	<i>neg</i>	<i>vt</i>	<i>n</i>	<i>vt</i>	<i>part.asp</i>

Indeed, (people) became unable to get the drinking water. (i.e. the water became undrinkable.)

AS.049

<i>Naang</i>	<i>za'gengh</i>	<i>hlo</i>	<i>haic</i>	<i>bieqc</i>	<i>gu'nyuoz</i>	<i>mingh.</i>
snake	really	big	very	enter	inside	go
<i>n</i>	<i>adv</i>	<i>v.st</i>	<i>adv</i>	<i>vi</i>	<i>n</i>	<i>v.asp</i>

The really very big snake fell into (the well).

AS.050

<i>Aengx</i>	<i>caangh_laangh</i>	<i>naaic</i>	<i>laangz</i>	<i>mienh,</i>
again	discuss	that	village	people
<i>adv</i>	<i>v</i>	<i>pn.dem</i>	<i>n</i>	<i>n</i>

<i>yietc</i>	<i>norm</i>	<i>fouv-zingh</i>	<i>nyei</i>	<i>mienh.</i>
one	CLF	capital_city	POSS	people
<i>numb</i>	<i>clf</i>	<i>n</i>	<i>poss</i>	<i>n</i>

And again the people of that village, of the whole capital discussed.

AS.051

<i>Hm,</i>	<i>maiv</i>	<i>haih</i>	<i>fungc</i>	<i>zoux.</i>
well...	NEG	be_able_to	how-come	do
<i>intj</i>	<i>neg</i>	<i>aux</i>	<i>adv</i>	<i>vt</i>

Well, there was nothing (they) could do about it.

AS.052

<i>Aqv_zuqc</i>	<i>mingh</i>	<i>lorz</i>	<i>Aa^Han_Mix,</i>	<i>Aa^Han_Suv^Tiqc</i>	<i>daaih</i>	<i>zorqv</i>	<i>hngangv.</i>
must	go	look_for	Bear_the_Brave	Suthi_the_Brave	come	take	no_other_way
<i>aux</i>	<i>vi</i>	<i>vt</i>	<i>n.prp</i>	<i>n.prp</i>	<i>v.dr</i>	<i>vt</i>	<i>adv</i>

There is no other way than (we) must go look for Bear the Brave, Suthi the Brave (and ask him to) catch (the snake).

AS.053

53.1			53.2				53.3			
<i>Mingh</i>	<i>lorz</i>	<i>ninh</i>	<i>mingh</i>	<i>taux</i>	<i>wuov</i>	<i>norm</i>	<i>laangz</i>	<i>buangh</i>	<i>zuqc</i>	<i>ninh</i>
go	look_for	3sg	go	reach	that	CLF	village	meet	GOAL	3sg
<i>vi</i>	<i>vt</i>	<i>pn.p</i>	<i>vi</i>	<i>v</i>	<i>pn.dem</i>	<i>clf</i>	<i>n</i>	<i>vt</i>	<i>part</i>	<i>pn.p</i>

53.4			53.5	
<i>hngangv</i>	<i>tengx</i>	<i>meih</i>	<i>weih</i>	<i>laatc.</i>
like	help	2sg	wall_in	fence
<i>adv</i>	<i>vt</i>	<i>pn.p</i>	<i>vt</i>	<i>n</i>

^{53.1}(They) went (to) look for him, ^{53.2}went (and) reached that village, ^{53.3}found him (there), ^{53.4}(he was) like help(ing) you ^{53.5}(to) make a fence.

AS.054

54.1			54.2		54.3					
<i>Yiem</i>	<i>wuov</i>		<i>weih</i>	<i>laatc</i>	<i>cipv</i>	<i>jienv</i>	<i>yie</i>	<i>wuov</i>	<i>nyungc</i>	<i>nzuqc_paiv</i>
be_there	there		wall_in	fence	stick	DUR	lsg	that	kind	scabbard
vi	pn.dem		vt	n	vt	asp	pn.p	pn.dem	n	n
54.4										
<i>heuc</i>	<i>ninh,</i>									
call	3sg									
v	pn.p									

"Aa *yie* *mingh* *mangc* *gaax* *wuov* *nyungc.*"
 Ah lsg go look try_and_see that kind
 intj pn.p vi v part pn.dem n

^{54.1}Being there, ^{54.2}(he was) build(ing) in a fence ^{54.3}sticking a scabbard like mine (in his back part of the belt); ^{54.4}(and they) called him. (He answered) "Ah, I will go and see that kind (of thing)."

AS.055

55.1									
<i>Ndau</i>	<i>juiz</i>		<i>deix</i>						
land	be_steep		some						
n	v.st		adv						
55.2	55.3								
<i>eix_leiz</i>	<i>wuom-kuotv</i>	<i>yiem</i>	<i>ndiev</i>	<i>maengx, ninh</i>	<i>yiem</i>	<i>jiex</i>	<i>maengx.</i>		
meaning	well	exist	down_there	side 3sg	be_in	up_there	side		
n	n	vi	adv.loc	n	pn.p	vi	n	n	

^{55.1}The land was steep; ^{55.2}which means the well was on the lower side, ^{55.3}he was on the upper side.

AS.056

56.1		56.2			56.3				
<i>Mingh</i>	<i>taux</i>	<i>yietv</i>	<i>ngamv</i>	<i>njiec,</i>	<i>nzuqc_paiv</i>	<i>yietv</i>	<i>donx</i>	<i>naaiv</i>	
go	reach	as_soon_as	squat	go_down	scabbard	as_soon_as	hit	this	
vi	v	adv	vi	vi	n	adv	vt	pn.dem	
56.4									
<i>ninh</i>	<i>ziang</i>	<i>wuov</i>	<i>biu</i>	<i>gu'nyuoz</i>	<i>mi'aqv.</i>				
3sg	just	that	jump	inside	RSLT				
pn.p	adv	pn.dem	vi	n	asp				

^{56.1}(They) went (and) reached (the well), ^{56.2}as soon as (he) squatted down ^{56.3}his scabbard hit this [i.e. his back], ^{56.4}he just fell down there inside (the well).

AS.057

<i>U</i>	<i>Janx</i>		<i>za'gengh</i>	<i>ceng</i>	<i>ninh</i>	<i>haic.</i>
Wow!	non_Iu-Mienh		really	praise	3sg	very
intj	n		adv	vt	pn.p	adv

Wow! The Northern Thai really praised him very much.

AS.058

58.1			58.2		58.3			
<i>Mangc</i>	<i>gaax</i>	<i>naaic</i>	<i>daaih</i>	<i>taux</i>	<i>ziang_naaic</i>	<i>biu</i>	<i>gu'nyuoz</i>	<i>mi'aqv.</i>
look	see	that	come	reach	just_then	jump	inside	RSLT
v	part	pn.dem	vi	v	adv	vi	n	asp

^{58.1}Look at him, ^{58.2}(he) arrived here ^{58.3}(he) immediately jumped into the well.

AS.059

Zorqv naang-jaang nanv jienv.
 take snake-neck grasp DUR
 vt n-n vt asp
 (He) took the snake by the neck, grabbing (it)

AS.060

60.1	60.2	60.3				60.4	
Wuov_deix	zorqv	caang	nzopv	nzopv	nzopv	nzopv	daic mi'aqv.
those_people	take	spear	pierce	pierce	pierce	pierce	die RSLT
pn.dem	vt	n	vt	vt	vt	vt	vi asp

^{60.1}Those people ^{60.2}took spears ^{60.3}pierced (and) pierced repeatedly (the snake) ^{60.4}(and it) died.

AS.061

Cingx_daaih	aengx om	yietc	nzunc	ninh	zoux	Aa^Han_Nguv.
therefore	again appoint	one	time	3sg	be	Snake_the_Brave
conj	adv vt	numb n	pn.p	vi	n.prp	

Therefore, (the governer) again gave him another title (to) be Snake the Brave.

AS.062

Za'gengh	longx	haic	aqv	wuov	nzunc.
really	be_good	very	CH-O-ST	that	time
adv	v.st	adv	asp	pn.dem	n

This time it was really good.

AS.063

Za'gengh	duqv	nyaanh	duqv	lui-houx.
really	get	money	get	clothes
adv	vt	n	vt	n

(He) actually got money and clothes.

AS.064

Gwye	(Jien)	za'gengh	a'hneiv	haic	ninh.
officer	official	really	be_happy	very	3sg
n	n	adv	v.st	adv	pn.p

The governer was really happy about him.

AS.065

65.1			65.2			
Aengx	taux	da'nyeic	hnoi,	aengx	maaih	nda'maauh
and_then	reach	second	day	further	have	tiger
adv	v	n	n	adv	vt	n
						faanv.
						be_disturbing
						v.st

^{65.1}And then another day, ^{65.2}furthermore there was a disturbing tiger.

AS.066

66.1	66.2	66.3	66.4	
Maaiah	laanh	mienh	nda'maauh	daaih
have	CLF	person	tiger	come
vt	clf	n	n	vi
				ngaatic
				gau,
				jai
				ngaatic
				gau,
				dungz
				ngaatic
				gau,
				bite
				DPCL
				chicken
				n
				vt
				part
				n
				vt
				part
				n
				vt
				top

^{66.1}There was a person ^{66.2}the tiger came (and) bit (him); ^{66.3}then (it) bit chickens, ^{66.4}then (it) bit pigs, ^{66.5}(it)

wanted all people!

AS.067

Caangh_laangh aa, maiv_haih_func _zoux.
 discuss Ah can' t_do_anything
 v intj idm

(They) discussed (saying) Ah, there's nothing (we) can do (about it).

AS.068

68.1
 Aqv_zuqc mingh lorz Aa^Han_Mix, Aa^Han_Nguv, Aa^Han_Suv^Tiqc
 must go look_for Bear_the_Brave Snake_the_Brave Suthi_the_Brave
 aux vi vt n.prp n.prp n.prp

68.2
 daaih zorqv hnavg.
 come take no_other_way
 vi vt adv

^{68.1}(We) must go look for Bear the Brave, Snake the Brave, Suthi the Brave ^{68.2}(and have him) catch (the tiger), there is no other way.

AS.069

Aengx mingh lorz buac ninh.
 and_then go look_for see 3sg
 adv vi vt v pn.p

And then (they) went (and) found him.

AS.070

70.1
 Ninh mingh wuov ndiev
 3sg go there down_there
 pn.p vi pn.dem adv.loc

70.2
 hnavg laangz yiem naaiv, 70.3
 like village exist here 3sg go there ndiev ndoqv wuov.
 adv n vi pn.dem pn.p vi pn.dem adv.loc n there
 pn.dem

^{70.1}He went down there, ^{70.2}like in this village, ^{70.3}he went down there at the stream.

AS.071

Wuov hnoi lung_haanx ziouc haiz wuov_jiex laangz mienh.
 that day after_noon so_then hear upper_side village people
 pn.dem n adv.tmp seq.mk vt n n n

In the afternoon of that day, then soon (they) heard (a sound of) the upper village people.

AS.072

72.1
 Mbiouh mv mbiouh nyei 72.2 72.3
 noisy intensely noisy ADVBLZR chase jienv nda'maauh daaih aqv.
 v.st intj v.st part vt asp n come PFT
 v.st intj v.st part vt asp n vi part.asp

^{72.1}(They) made a lot of noise ^{72.2}chasing the tiger ^{72.3}(and they) came.

AS.073

73.1	Zunc	jienv	nda'maauh	taux	wuov	73.2	ninh	fiaux	ndiangx	mingh	73.3	yiem	gu'nguaaic.
	chase	SML.ACT	tiger	reach	there		3sg	go_up	tree	go		be_there	above
	vt	asp	n	v	pn.dem		pn.p	vi	n	v.dr		vi	adv.loc

^{73.1}(When they) came chasing the tiger (and) reached there, ^{73.2}he climbed up a tree ^{73.3}(and) stayed above.

AS.074

Ga'ndiev	ndiangx	njongz	nyei.
underneath	tree	be_hollow	PRS.ST
adv.loc	n	v.st	asp

Underneath the tree was hollow.

AS.075

75.1	Nda'maauh	suoh	suoh	nyei	taux	wuov	75.2	youc	maiv	bieqc.
	tiger	MIM	MIM	ADVBLZR	reach	there		but	NEG	enter
	n	v.st	v.st	part	v	pn.dem		seq.mk	neg	vi

^{75.1}The tiger arrived there slinking along ^{75.2}but did not come in (the hole).

AS.076

76.1	Haaix	ndau	maiv	mingh	76.2	youc	bieqc	wuov	ndiangx-nqunx	mingh.
	what	ground	NEG	go	also	enter	there		tree-hole	go
	q	n	neg	vi	seq.mk	vi	pn.dem		n-n	v.dr

^{76.1}(It) did not go anywhere ^{76.2}but entered that hole of the tree.

AS.077

Dueiv	yiem	ga'nyiec.
tail	be_in	outside
n	vi	n

(Its) tail was outside.

AS.078

Dueiv	ndaauv	nyei	yiem	ga'nyiec.
tail	long	PRS.ST	exist	outside
n	adj	asp	vi	n

The tail was long (and) it stayed outside.

AS.079

79.1	Wuov_deix	huei,	huei	nyei	zunc	jienv	daaih
	those_people	MIM	MIM	ADVBLZR	chase	SML.ACT	come
	pn.dem	v.st	v.st	part	vt	asp	vi
79.2	bungx	jienv	congx	daaih.			
	Release	SML.ACT	gun	come			
	vt	asp	n	vi			

^{79.1}Those people came chasing (it) with a loud noise, ^{79.2}came shooting guns.

AS.080

80.1	Wuov_deix	daaih	fatv	deix	80.2	ninh	njiec	aiv	deix.
	those_people	come	near	some	3sg	descend		low	some
	pn.dem	vi	v.st	adv	pn.p	vi		v.st	adv

^{80.1}(As) they came a little closer, ^{80.2}he came down a little.

AS.081

81.1					81.2			
<i>Wuov_deix</i>	<i>daaih</i>	<i>fatv</i>	<i>deix</i>	<i>ninh</i>	<i>njec</i>	<i>aiv</i>	<i>deix.</i>	
those_people	come	near	some	3sg	descend	be_low	some	
pn.dem	vi	v.st	adv	pn.p	vi	v.st	adv	

^{81.1}(As) they came a little closer, ^{81.2}he came down a little.

AS.082

82.1					82.2			
<i>Hnangv</i>	<i>wuov_deix</i>	<i>taux</i>	<i>wuov</i>	<i>nda</i>	<i>ninh</i>	<i>tiux</i>	<i>njec</i>	<i>daaih</i>
as_if	those_people	reach	that	place	3sg	run	go_down	come
adv	pn.dem	v	pn.dem	n	pn.p	vi	vi	vi

82.3				
<i>baeng</i>	<i>jienv</i>	<i>nda'maauh</i>	<i>neyi</i>	<i>dueiv.</i>
pull	SML.ACT	tiger	POSS	tail
vt	asp	n	poss	n

^{82.1}It looked like (that) when those people arrived there ^{82.2}he ran down ^{82.3}pulling the tiger's tail.

AS.083

<i>Nauc</i>	<i>wuov_deix,</i>									
yell_at	those_people									
vt	pn.dem									
	<i>"Nqongh</i>	<i>daaih</i>	<i>nqongh</i>	<i>daaih</i>	<i>yie</i>	<i>zorqv</i>	<i>duqv</i>	<i>dauh</i>	<i>nda'maauh</i>	<i>aqv.</i>
	hurry	come	hurry	come	1sg	take	can	CLF	tiger	PFT
	v	vi	v	vi	pn.p	vt	v.aux	clf	n	part.asp

(He) shouted to them, "Come quickly! I've just managed to catch the tiger."

AS.084

<i>Wuov_deix</i>	<i>daaih</i>	<i>taux</i>	<i>buonv</i>	<i>neyi</i>	<i>buonv.</i>
those_people	come	reach	shoot	ADVBLZR	shoot
pn.dem	vi	v	vt	part	vt

Those people arrived, shooting repeatedly.

AS.085

85.1				85.2	85.3
<i>Baqv</i>	<i>neyi</i>	<i>baqv,</i>	<i>baqv</i>	<i>daic</i>	<i>mingh.</i>
pierce	ADVBLZR	pierce	pierce	die	go
vt	part	vt	vt	vi	v.asp

^{85.1}(They) pierced and pierced (the tiger) ^{85.2}pierced ^{85.3}(it) died.

AS.086

<i>Aengx</i>	<i>dorh</i>	<i>mingh</i>	<i>bun</i>	<i>gwyen</i>	<i>mangc.</i>
and_then	take	go	let	officer	look
adv	vt	vi	vt	n	vt

(He) took it back (to) let the governor look at (it).

AS.087

<i>Gwyen</i>	<i>aengx</i>	<i>heuc</i>	<i>ninh</i>	<i>zoux</i>	<i>Aa^Han_Siex,</i>	<i>Aa^Han_Siex.</i>
officer	further_more	name	3sg	be	Tiger_the_Brave	Tiger_the_Brave
n	adv	vt	pn.p	vi	n.prp	n.prp

The governer furthermore called him Tiger the Brave, Tiger the Brave.

AS.088

<i>Aa^Han_Siex</i>	<i>duqv</i>	<i>biei</i>	<i>nzunc</i>	<i>aqv.</i>
Tiger_the_Brave	get	four	time	PFT
n.prp	vt	numb	n	part.asp

Tiger the Brave got (awarded) four times.

AS.089

<i>Za'gengh</i>	<i>henv.</i>
really	strong
adv	adj

(He) was really strong.

AS.090

<i>Nqa'haav</i>	<i>aengx</i>	<i>deic-bung</i>	<i>aengx</i>	<i>faanv</i>	<i>hnavgv</i>	<i>ih_zanc</i>	<i>naaiv</i>	<i>mborqv</i>	<i>jaax.</i>
Later	again	country	again	unsettled	like	now	here	hit	REC.ACT
adv.tmp	adv	n	adv	v.st	adv	adv.tmp	pn.dem	vt	adv

Later, the country became unsettled again, like people here fight each other nowadays.

AS.091

91.1	<i>Mborqv</i>	<i>jaax</i>	<i>mienh</i>	<i>camv,</i>	<i>baeng</i>	<i>camv,</i>
hit	REC.ACT	people	many	soldier	many	
vt	adv	n	v.st	n	v.st	

91.2	<i>haaix_deix</i>	<i>yaac</i>	<i>mborqv</i>	<i>waaic</i>	<i>nzengc.</i>
wherever	also	hit	be_destroyed	consumed	
n	conj	vt	v.st	v	

^{91.1}(There) were many fighters, many soldiers; ^{91.2}(they) destroyed everywhere.

AS.092

<i>Maaih</i>	<i>naaiv</i>	<i>dauh</i>	<i>gwyen</i>	<i>yiem</i>	<i>naaic</i>	<i>yietc</i>	<i>norm</i>	<i>zingh.</i>
have	this	CLF	officer	live	there	one	CLF	city
vt	pn.dem	clf	n	vi	pn.dem	numb	clf	n

There was this officer living in that city.

AS.093

<i>Wuov_deix</i>	<i>baeng</i>	<i>aengx</i>	<i>oix</i>	<i>daaih</i>	<i>mborqv.</i>
those	soldier	again	want_to	come	hit
pn.dem	n	adv	v.mod	vi	vt

Further those soldiers wanted to come (to) fight.

AS.094

<i>Maiv_haih_func_zoux.</i>
can't_do_anything
idm

There was nothing (he) could do about it.

AS.095

95.1	<i>Aa</i>	<i>heuc</i>	<i>Aa^Han_Mix</i>	<i>Aa^Han_Suv^Tiqc</i>	<i>Aa^Han_Nguv</i>	<i>Aa^Han_Siex</i>	<i>mingh</i>
Ah	call	Bear_the_Brave	Suthi_the_Brave	Snake_the_Brave	Tiger_the_Brave	go	
intj	v	n.prp	n.prp	n.prp	n.prp	vi	

95.2
zorqv hñangv.
 take no_other_way
 vt adv

^{95.1}Ah, (we only) order Bear the Brave, Suthi the Brave, Snake the Brave, Tiger the Brave (to) go ^{95.2}(and) catch (the soldiers), there is no other way.

AS.096

96.1				96.2			96.3				
<i>Mv zeiz nor aeqv, mv hingh aeqv, mbuo naaiv mv maaih faatv.</i>											
NEG be as TOP NEG win TOP 1pl here NEG have magic											
neg vi top top neg vt top pn.p pn.dem neg vt n											

^{96.1}If we don't, ^{96.2}(we) can't win and then ^{96.3}we do not have magic here.

AS.097

97.1				97.2					
<i>Mingh wuov ndiev! gomgv ziangx</i>									
go there down_there speak finish									
vi pn.dem adv.loc v vi									

97.3									
<i>wuov ngaanc domh zomg yietc norm ndaamh.</i>									
that opposite_side big mountain one CLF slope									
pn.dem n adj n numb clf n									

^{97.1}(The governor's people) went far down there ^{97.2}(and) finished talking (with the brave man); ^{97.3}on the slope on the opposite side of the great mountain.

AS.098

98.1	98.2									
<i>Yiem wuov ninh faaux taux wuov aeqv, maaih yietc nyungc nyaaix-weih.</i>										
be_in there 3sg go_up reach there TOP have one kind fern										
vi pn.dem pn.p vi v pn.dem part vt numb n n										

^{98.1}From there as he went up there, ^{98.2}there was one kind of fern.

AS.099

<i>Ndiangx yiem naaiv mbu'ndongx yietc nyungc nyaaix-weih gormx mingh.</i>								
tree exist here middle one kind fern traverse_completely go								
n vi pn.dem n numb n n vt v.dr								

Among the trees here a kind of fern traversed completely.

AS.100

100.1			100.2		
<i>Ninh faaux mingh zueiz jienv wuov.</i>					
3sg go_up go sit DUR there					
pn.p vi vi vi asp pn.dem					

^{100.1}He climbed up ^{100.2}sitting there.

AS.101

<i>Baeng za'gengh daaih camv haic.</i>				
soldier really come many very				
n adv vi v.st adv				

The numerous soldiers really came.

AS.102

Ziex	cin	mbu'ziex	cin	mbu'ziex	waanc	yaac	maiv	hiuv.
many	thousand	how_many	thousand	how_many	ten_thousand	TOP	NEG	know
adj	numb	q	numb	q	numb	part	neg	v

Many thousand, (or we) don't know if it was how many thousand, how many ten-thousand.

AS.103

Daaih	taux	youc	haaix	ndaui	maiv	hitv.
come	reach	also	what	ground	NEG	rest
vi	v	seq.mk	q	n	neg	vi

Didn't know where they stayed either.

AS.104

104.1	104.2							
Mingh	hitv	jienv	ndiangx-gom-hlen,	camv!	nyei	hitv	jienv	wuov.
go	rest	DUR	tree-base-side	many	ADVBLZR	rest	DUR	there
vi	vi	asp	n-n-n	v.st	part	vi	asp	pn.dem

^{104.1}(They) went (and) were resting at the tree base, ^{104.2}numerously resting there.

AS.105

Yiem	wuov	caangh_laangh.
be_there	there	discuss
vi	pn.dem	v

They were there (and they) discussed (i.e. they were discussing there).

AS.106

"Uv!	Gamh_nziex	haic,
Oh!	fear	very
intj	v.cmp	adv

wuov	laangz	mienh	yie_mbuo	se	mv	gamh_nziex,
that	village	people	1_PL	that_is	NEG	fear
pn.dem	n	n	pn.p	copl	neg	v.cmp

wuov	norm	fouv-zingh	nyei	mienh	yie_mbuo	maiv	gamh_nziex.
that	CLF	capital_city	POSS	people	1_PL	NEG	fear
pn.dem	clf	n	poss	n	pn.p	neg	v.cmp

Gamh_nziex	Aa^Han_Nguv	Aa^Han_Mix	hnangv."
Fear	Snake_the_Brave	Bear_the_Brave	only
v.cmp	n.prp	n.prp	adv

"Oh, it's scary! We are not afraid of the village people, we are not afraid of the people of the city. We only fear Snake the Brave, Bear the Brave."

AS.107

Ninh	yiem	gu'nguaaic	ninh	haiz	nyei.
3sg	be_in	above	3sg	hear	PRS.ST
pn.p	vi	adv.loc	pn.p	vt	asp

He could hear (the soldiers' conversation) staying above.

AS.108

"Uv!	se_gomgv	duqv	Aa^Han_Mix	Aa^Han_Nguv	daic	mingh	nor	aeqv,
Oh!	if	get	Bear_the_Brave	Snake_the_Brave	die	go	as	TOP
intj	conj	vt	n.prp	n.prp	vi	vi	top	top
yietc	laangz	mienh	maiv	gaux	yie_mbuo	nyanc	lung	ndorm."
whole	village	people	NEG	enough	1_PL	eat	sky	morning
adj	n	n	neg	v.st	pn.p	vt	n	n

"Oh! If (we) get Bear the Brave, Snake the Brave killed, the people of the whole village are not enough for us to eat for breakfast."

AS.109

109.1			109.2				109.3			
<i>Ninh</i>	<i>gamh_nziex</i>	<i>haic</i> ,	<i>gamh_nziex</i>	<i>haic</i>	<i>sin</i>	<i>zinx</i>	<i>gam</i>	<i>jienv</i>	<i>ndiangx</i> .	
3sg	fear	very	fear	very	body	tremble	hold	DUR	tree	
pn.p	v.cmp	adv	v.cmp	adv	n	vi	vt	asp	n	

^{109.1}He was terrified very much, ^{109.2}terrified so much (that his) body trembled, ^{109.3}clinging to a tree.

AS.110

<i>Sin</i>	<i>zinx</i>	<i>za'gengh</i>	<i>nyanh</i>	<i>nyanh</i>	<i>nyanh</i>	<i>nyanh</i>	<i>nyanh</i>	<i>nyanh</i>	<i>nyanh</i>	<i>nyanh</i>	<i>nyanh</i>	<i>deix</i> .
body	tremble	really	jerk	jerk	jerk	jerk	jerk	jerk	jerk	jerk	jerk	some
n	vi	adv	vi	vi	vi	vi	vi	vi	vi	vi	vi	adv

(His) body trembled much continuously.

AS.111

111.1										
<i>Caang</i>	<i>wuov_deix</i>	<i>nyei</i>	<i>baeng</i>	<i>caang_congx</i>	<i>baeqc</i>	<i>nzengc</i>				
spear	those	RELT	soldier	gun	white	consumed				
n	pn.dem	relt	n	n	v.st	v				

111.2			111.3	
<i>yiem</i>	<i>wuov</i>	<i>zoux</i>	<i>nyanc</i> .	
be_there	there	make	eat	
vi	pn.dem	vt	vt	

^{111.1}(The area was) all white (with) the soldiers (holding) spears (and) guns, ^{111.2}(they) were there ^{111.3}cook (ing) food.

AS.112

112.1				112.2	
<i>Ciev</i>	<i>jienv</i>	<i>geh</i>	<i>daaih</i>	<i>siqv</i>	<i>nyei</i> .
direct	DUR	ride	come	red	PRS.ST
vt	asp	vt	vi	v.st	asp

^{112.1}(Others) came driving horses ^{112.2}(they) were red.

AS.113

113.1		113.2				
<i>Ninh</i>	<i>aeqv</i> ,	<i>sin</i>	<i>zinx</i>	<i>haic</i>	<i>aav</i>	<i>lov!</i>
3sg	TOP	body	tremble	very	SURPRISE	
pn.p	top	n	vi	adv	part.f	

^{113.1}As for him, ^{113.2}his body trembled terribly.

AS.114

114.1		114.2		114.3				
<i>Naaiv</i>	<i>aeqv</i> ,	<i>mbuo</i>	<i>yiem</i>	<i>naaiv</i> ,	<i>gorv_laaic</i>	<i>ninh_mbuo</i>	<i>maiv</i>	<i>daaih</i> .
here	TOP	1pl	live	here	guess	3_PL	NEG	come
pn.dem	top	pn.p	vi	pn.dem	v.cmp	pn.p	neg	vi

^{114.1}As for this place, ^{114.2}we are here, ^{114.3}(we) suspect (that) they won't come here.

AS.115

115.1				115.2		115.3		
<i>Za'gengh</i>	<i>daaih</i>	<i>naaiv</i>	<i>aeqv</i> ,	<i>cong</i>	<i>mv</i>	<i>maaih</i> ,	<i>nzuqc</i>	<i>mv</i>
really	come	here	TOP	gun	NEG	have	sword	NEG
adv	vi	pn.dem	top	n	neg	vt	n	neg
								vt

^{115.1}If (they) really come here, ^{115.2}(we) don't have guns, ^{115.3}don't have swords.

AS.116

Fungc zoux.
adv v

Can't do anything.

AS.117

117.1				117.2	
<i>Aav_dangh</i>	<i>ninh_mbuo</i>	<i>ciepv</i>	<i>buatc</i>	<i>ndortv</i>	<i>aqv,</i>
in_a_short_while	3_PL	look_at	see	fall	PFT
adv.tm	pn.p	vt	v	vi	part.asp

117.3			
<i>za'gengh</i>	<i>gamh_nziex</i>	<i>haic.</i>	
really	fear	very	
adv	vt	adv	

^{117.1}Suddenly they saw ^{117.2}(the brave man) fall down, ^{117.3}(they) were really terrified very much.

AS.118

<i>Ninh</i>	<i>oix</i>	<i>daic</i>	<i>haic.</i>
3sg	want_to	die	very
pn.p	v.mod	vi	adv

He was scared very much.

AS.119

<i>Sin</i>	<i>zinx</i>	<i>nyanh</i>	<i>gau.</i>
body	tremble	jerk	very
n	vi	vi	adv

His body trembled very much.

AS.120

120.1			
<i>Nyaaix-weih</i>	<i>nquatv</i>		
fern	crack_down		
n	vi		

120.2			
<i>ziouc</i>	<i>suz!</i>	<i>suz</i>	<i>nyei</i>
so_then	sound_of_spurting_air	sound_of_spurting_air	ADVBLZR
seq.mk	onoma	onoma	part

<i>fortc</i>	<i>wuov</i>	<i>ndau</i>	<i>daaih.</i>
cry	that	ground	come
vi	pn.dem	n	vi

^{120.1}The fern cracked down, ^{120.2}so (he) zoomed down to the ground with a sound of spurting air.

AS.121

<i>Maiv_haih_fungc_zoux</i>	<i>aqv.</i>
can't_do_anything	CHG-O-ST
idm	part.asp

He couldn't help.

AS.122

122.1			122.2		122.3	
<i>Nquaqv</i>	<i>sin</i>	<i>daaih</i>	<i>mbaix</i>	<i>buoz</i>	<i>mbaix</i>	<i>zoux.</i>
get_up_quickly	body	come	clap	hand	slap	foot
vt	n	v.dr	vt	n	vt	n

^{122.1}(He) got up quickly ^{122.2}clapped (his) hands ^{122.3}beat (his) thighs.

AS.123

"Aa^Han_Mix	Aa^Han_Nguv	Aa^Han_Suv^Tiqc	dongh	ye	aqv.
Bear_the_Brave	Snake_the_Brave	Suthi_the_Brave	that_is	1sg	I_am_sure
n.prp	n.prp	n.prp	idntf	pn.p	part.mod

"Bear the Brave, Snake the Brave, Suthi the Brave, that's me myself!"

AS.124

124.1		124.2			124.3			
lh_hnoi	ziangh,	ih_hnoi	daic	yaac	ih_hnoi	haeqv!	wuov_deix	aeqv.
today	alive	today	die	TOP	today	frighten	those_people	TOP
adv.tmp	v.st	adv.tmp	vi	top	adv.tmp	vt	pn.dem	top

124.4							
mangc	gaax	mbuo	yiem	naaiv	caangh_laangh	hnanv,	
look	try_and_see	1pl	be_in	here	discuss	only	
v	part	pn.p	vi	pn.dem	v	adv	

124.5								
ninh	mv	beiv	duqv	yiem	haaix_ndau	ndaix	daaih	aqv.
3sg	NEG	compare	can	be_there	what_ground	fly	come	PFT
pn.p	neg	v	v.aux	vi	q	vi	vi	part.asp

^{124.1}(Even though) today (he) is alive, ^{124.2}(or) dies today, ^{124.3}(since he) frightened them today, ^{124.4}let's see, we only discuss here, ^{124.5}there is no way to describe where he flew from.

AS.125

Maiv	jangx	mangc	nyaaix-weih.
NEG	remember	look	fern
neg	v	v	n

(He) didn't remember to look at the fern.

AS.126

126.1			126.2					
Uv!	biaux	ninh	ziouc	zorqv	baeng	neyi	congx	daaih
Oh!	flee	3sg	so_then	take	soldier	POSS	gun	come
intj	v	pn.p	seq.mk	vt	n	poss	n	v.dr

126.3						126.4	
buonv	neyi	buonv	nzuangv	neyi	nzuangv	zunc	jienv
shoot	ADVBLZR	shoot	aimless	ADVBLZR	aimless	chase	SML.ACT
vt	part	vt	v.st	part	v.st	vt	asp
							mingh.
							go
							vi

^{126.1}Oh! (they) ran away, ^{126.2}so he took the soldier's guns ^{126.3}(and) shot repeatedly, ^{126.4}(and) chased them away furiously.

AS.127

Wuov_nzunc_hnoi	ninh	aengx	ndaam	wuov_deix	neyi	congx.
this-time	3sg	again	carry	those_people	POSS	gun
adv.tmp	pn.p	adv	vt	pn.dem	poss	n

Then he further carried their guns.

AS.128

Wuov_deix	biaux	nzengc	mi'aqv.
those_people	flee	consumed	RSLT
pn.dem	vi	vi	asp

They all ran away.

AS.129

129.1				129.2			
<i>Ndaam</i>	<i>jienv</i>	<i>congx</i>	<i>nzuonx</i>	<i>mingh</i>	<i>bun</i>	<i>gwyen</i>	<i>mangc.</i>
carry	SML.ACT	gun	return	go	let	officer	look
vt	asp	n	v	vi	vt	n	vt

^{129.1}(He) went back carrying the guns on his shoulders, ^{129.2}(and) let the governor look (at them).

AS.130

130.1			130.2		
<i>Buonv</i>	<i>baaic</i>	<i>mi'aqv,</i>	<i>ninh_mbuo</i>	<i>biaux</i>	<i>nzengc</i>
Shoot	defeat	RSLT	3_PL	flee	consumed
vt	vt	asp	pn.p	v	vi
					<i>mi'aqv.</i>
					RSLT
					asp

^{130.1}(He) shot (and) defeated (them) all; ^{130.2}they all completely ran away.

AS.131

<i>Gwyen</i>	<i>aengx</i>	<i>bun</i>	<i>mienh</i>	<i>mingh</i>
Officer	further	let	person	go
n	adv	vt	n	vi
<i>zorqv</i>	<i>congx</i>	<i>lo_haaix</i>	<i>nzuonx</i>	<i>daaih.</i>
take	gun	and_so_forth	return	come
vt	n	n	vi	vi

The officer further let people go (to) get guns and so forth back.

AS.132

132.1					132.2		
<i>Wuov_nzunc_hnoi</i>	<i>paaiv</i>	<i>yietc</i>	<i>gorqv</i>	<i>deic-bung</i>	<i>yietc</i>	<i>norm</i>	<i>fouv-zingh</i>
this_time	appoint	one	corner	region	one	CLF	capital_city
adv.tmp	vt	numb	n	n	numb	clf	n
132.3							
<i>bun</i>	<i>ninh</i>	<i>nyanc.</i>					
let	3sg	eat					
vt	pn.p	vt					

^{132.1}Then this time, (the governor) assigned one corner of a region ^{132.2}(with) one capital city ^{132.3}(to) let him make a living.

AS.133

<i>Za'gengh!</i>	<i>ninh</i>	<i>yaac</i>	<i>duqv</i>	<i>zoux</i>	<i>hlo.</i>
really	3sg	also	get	be	big
adv	pn.p	conj	vt	vi	v.st

Really he also became big.

AS.134

<i>l_gox</i>	<i>loz-hnoi</i>	<i>za'gengh</i>	<i>jomc</i>	<i>haic.</i>
married_couple	old-days	really	poor	very
n	n	adv	v.st	adv

Before that the couple was really very poor.

AS.135

Nae!
See
intj

I have told you.

FIXED AUTOSEGMENTISM IN THAI EMPHATIC REDUPLICATION¹⁵

Leslie Lee

University of California at San Diego
<leslie@ling.ucsd.edu>

Abstract

The classical approach to dealing with fixed material in reduplication is to assume that the reduplicant is prespecified with the fixed content (Marantz 1982, Yip 1982b). However, researchers working in OT have argued that fixed segmentism in reduplication can be explained without resorting to prespecification, and that the emergence of fixed segmentism in the reduplicant is either phonologically or morphologically determined (Alderete *et al.* 1999, Walker 2000). This paper extends this argument to Thai emphatic reduplication, where, instead of segmental content, it is the autosegmental content of tone that is fixed. I argue that the analysis of Thai emphatic reduplication does not require prespecification in the reduplicant, while accounting for the occurrence of high tone on CVVO syllables in emphatic reduplication - a phenomenon that is generally prohibited in the language.

Keywords: phonology, tones, reduplication

1. Introduction

It has been argued that fixed segmentism in reduplication can be explained without having to resort to prespecification in the reduplicant, and that the emergence of fixed segmentism in the reduplicant is either phonologically determined or morphologically determined (Alderete *et al.* 1999, Walker 2000). In this paper, I explore a related phenomenon in Thai emphatic reduplication, where the reduplicant appears with an invariant high tone regardless of the tone on the base, i.e. it is not segmental content that is fixed, but autosegmental content. In addition to the issue of fixed autosegmentism, the occurrence of a high tone on the reduplicant also highlights a corresponding puzzle in the tonal phonology of the language: high tones are generally prohibited from occurring on CVVO syllables in Thai, yet, this combination does occur on the CVVO reduplicant in emphatic reduplication. Therefore, the aim of this paper is twofold: (i) to provide an account of fixed autosegmentism in reduplication without prespecification, and (ii) to account for the presence of high tone on CVVO reduplicants.

¹⁵ I would like to thank Sharon Rose and an anonymous reviewer for detailed comments that have led to the improvement of this paper. Any remaining errors are my own.

Leslie Lee. 2011. Fixed Autosegmentism in Thai Emphatic Reduplication.

Journal of the Southeast Asian Linguistics Society 4.1:41-63

Copyright vested in the author

Received 5/2/2010, revised text accepted 16/3/2011

1.1 Thai tones

Thai has five lexical tones. The five tones are traditionally (since Abramson 1962) described and transcribed as mid, low, high, falling, and rising. These are illustrated in (1):

- (1) Mid naa ‘rice field’
 Low nàa ‘custard apple’
 High náa ‘aunt’
 Falling nâa ‘face’
 Rising nǎa ‘thick’
- (Morén and Zsiga 2006: 114)

Not all syllables can bear all five tones. Only open syllables with long vowels (CVV) or syllables closed by a sonorant coda (CVS, CVVS) can bear all five tones:

- (2) CVV CVS CVVS
 naa ‘rice field’ laŋ ‘crate’ laaŋ ‘omen’
 nàa ‘custard apple’ làŋ ‘to flow’ làaŋ ‘chime of a bell’
 náa ‘aunt’ lám ‘to go beyond’ láaŋ ‘to wash’
 nâa ‘face’ lâm ‘sturdy’ låaŋ ‘below’
 nǎa ‘thick’ lǎŋ ‘back’ lǎan¹⁶ ‘grandchild’
- (Morén and Zsiga 2006: 115)

In syllables closed by obstruents, tonal distribution is restricted. In CVO syllables, only low and high tones may appear, while in CVVO syllables, only low and falling tones are permitted:

- (3) CVO CVVO
 mid --- ---
 low làk ‘stake’ làak ‘various’
 high lák ‘to steal’ ---
 falling --- làak ‘to tow’
 rising --- ---
- (Morén and Zsiga 2006: 116)

The table in (4) summarises the general distribution of tones in Thai with respect to syllable type:

(4) General distribution of tones in Thai

σ Type/Tone	Mid	Low	High	Falling	Rising
CVV, CVS, CVVS	✓	✓	✓	✓	✓
CVVO	✗	✓	✗	✓	✗
CVO	✗	✓	✓	✗	✗

¹⁶ The original transcription is laaŋ, which is surely a typographical error.

1.2 Reduplication in Thai

There are two types of reduplication in Thai that copy the base segments faithfully: simple reduplication and emphatic reduplication. In simple reduplication, the reduplicant is an exact copy of the base, including its tone, as illustrated in (5):

(5)	dii	‘good’	dii-dii	‘really good, better, (do) well’
	cháa	‘slow’	cháa-cháa	‘really slow, slower, (do) slowly’
	rew	‘fast’	rew-rew	‘really fast, faster, (do) fast’
	săaw	‘young’	săaw-săaw	‘really young, younger (female)’
	nùm	‘young’	nùm-nùm	‘really young, younger (male)’
	kèɛ	‘old’	kèɛ-kèɛ	‘really old (person)’

(Iwasaki and Ingkaphirom 2005: 34)

Since the reduplicant is simply an exact copy of the base, simple reduplication obeys the general distribution pattern of tones in the language.

On the other hand, emphatic reduplication copies only the base segments, and assigns an emphatic high tone to the reduplicant irrespective of the tone on the base.¹⁷ (6) illustrates:

(6)	dii	‘good’	díi-dii	‘very good’
	ʔim	‘full’	ʔím-ʔim	‘very full’ ¹⁸
	yâak	‘difficult’	yáak-yâak	‘very difficult’
	róɔn	‘hot’	róɔn-róɔn	‘very hot’
	wăan	‘sweet’	wáan-wăan	‘very sweet’

(Iwasaki and Ingkaphirom 2005: 35)

Unlike the simple reduplicant which fills its autosegmental content through copy from the base, the emphatic reduplicant displays fixed autosegmental material. Recall that high tones do not occur on CVVO syllables in the language (see (3-4)). The emphatic reduplication of a CVVO syllable, as in the case of *yáak-yâak*, then presents a puzzling situation where a high tone is permitted on a CVVO syllable in spite of the fact that this is generally prohibited in the tonal system of the language. This suggests that the requirement on the reduplicant to bear a high tone overrides the general phonotactic constraints in the language.

In light of the data concerning emphatic reduplication, two questions beg to be answered:

- (i) Why does the reduplicant surface with a high tone in emphatic reduplication? In other words, what is the basis for the fixed autosegmentism?
- (ii) What is the explanation for the high tone occurring on CVVO syllables?

¹⁷ According to Haas, phonemically, the emphatic high tone can be treated as a modification of the high tone. Phonetically, however, it is not identical with the high tone, for the emphatic high is always higher in pitch and usually more protracted in length than the normal high (Haas 1946). I assume that the emphatic high is underlyingly the same as the regular high tone.

¹⁸ Iwasaki and Ingkaphirom omitted the onset glottal stops in order to represent casual speech. In this paper, I assume the underlying phonological representations.

2. Analyses of Thai tonal distribution

Before proceeding to answer these questions, the gaps in the Thai tonal system must first be accounted for. More specifically for present purposes, the occurrence of high tone on CVVO syllables must be ruled out in the general phonology of the language.

Prior to Morén and Zsiga (2006), no complete explanation has been proposed for the distributional gaps. As they note, two different partial explanations have been proposed among previous analyses, each with their own shortcomings. Gandour (1974, 1977) and Zhang (2002) account for the lack of rising tone on obstruent final syllables by appealing to duration. Citing phonetic studies showing that rising contours tend to be durationally longer than falling contours (Ohala and Ewan 1972; Sundberg 1973, 1979; Xu 1999a, 1999b), these authors argue that the relatively shorter duration of the sonorant portion of the rhyme in obstruent-final syllables does not allow sufficient time for the phonetic realisation of a rising tone. However, an account based on duration cannot account for the prohibition against mid tones, nor can it explain why high tones are prohibited on CVVO but allowed on CVO, the shortest syllable of all. Instead of duration, Yip (1982a, 2002) grounds the prohibition against rising contours in glottalisation. Noting that syllable-final obstruents in Thai are both voiceless and glottalised, she formulates a negative constraint “*LH[+glottal]” which prevents rising contours from associating with obstruent-final syllables. She does not, however, extend the analysis to the other gaps in the system (Morén and Zsiga 2006: 117).

Since Morén and Zsiga (2006) is the only account that successfully accounts for all the gaps in the Thai tonal system, I assume that this is the correct analysis. On this analysis, the lack of mid, high, and rising tones on CVVO syllables is due to a positive constraint that requires obstruent coda segments to be associated with a low tone. This analysis is outlined in the following section. As the distributional gaps in CVO syllables are not crucial to our discussion of emphatic reduplication, I shall ignore this issue here and refer the interested reader to the original article.

2.1 *Bimoraicity in Thai*

Morén and Zsiga argue that all stressed words in Thai are bimoraic (2006: 127-129). First, long and short vowels contrast, but there are no extra long vowels in the language. That is, vowels may be monomoraic or bimoraic, but not trimoraic. Second, short open syllables cannot occur in stressed positions, and underlyingly CV syllables that occur in stressed positions are made heavy by the addition of a glottal stop. Thus, all stressed syllables are (minimally) bimoraic. Third, since CVC syllables can bear stress, the final coda in a CVC syllable is moraic. Finally, based on acoustic measurements, they note that phonetic shortening of the long vowel and coda occurs in CVVC syllables, suggesting that Thai words are maximally bimoraic, and that the second mora is shared between the vowel and the coda in a CVVC syllable. Thus, they propose the moraic structure in (7):

$$(7) \quad \begin{array}{ccc} \mu \mu & \mu \mu & \mu \mu \\ | | & | | & | \backslash \\ \text{CVV} & \text{CVC} & \text{CVVC} \end{array}$$

2.2 *Phonetic realisation of Thai tones in citation form*

Based on data recorded from two native female speakers, Morén and Zsiga provide a detailed description of the shapes of the five tones in citation form by dividing the

frequency scale into three subranges: high (240-280 Hz), mid (200-240 Hz), low (160-200 Hz):

- The mid tone stays level or falls gradually through the mid range.
- The low tone begins in the mid range and falls, reaching its lowest point at the bottom of the range at the end of the syllable.
- The high tone stays level or falls slightly in the mid range through the first mora, then rises to reach the high range at the end of the syllable.
- The falling tone begins in the high range, starting with a rising high during the first mora, and then falls to the low range during the second mora.
- The rising tone begins in the mid range, falling and staying low during the first mora before rising during the second mora, ending in the mid range.¹⁹

(Morén and Zsiga 2006: 130)

The tonal shapes described here are generally consistent with the findings of previous acoustic analyses of the tonal shapes in citation form (Abramson 1962, Gandour et al. 1991).

2.3 Representation of Thai tones

Morén and Zsiga noticed that, across syllable types and between their two speakers, the inflection points in tonal contours correspond closely to the right edges of morae: mid syllable inflections occurred nearly exactly at the end of the first mora. This led them to argue that the tone-bearing unit (TBU) in Thai is the mora, and that the specified tones are aligned to the right of morae. On the assumption that phonologically toneless morae are phonetically mapped onto the neutral pitch range, mid tone syllables have no phonological tonal specification and remain at neutral pitch throughout; high tones, which remain in the mid range through the first half of the syllable and reach a high point only at the end of the rhyme, begin with no tonal specification on the first mora and end with a high tone on the second; low tones, which begin in the mid range and fall to a low point at the end, begin with a tonally unspecified first mora and end with a low tone on the second mora; falling tones, which first exhibit a rise in the high range followed by a fall to the low range, begin with a high tone on the first mora and end with a low tone on the second; rising tones, which fall and stay low before rising, begin with a low tone on the first mora and end with a high tone on the second. Their proposed phonological representations of Thai tones are shown below:

(8)	<u>Mid</u>	<u>High</u>	<u>Low</u>	<u>Falling</u>	<u>Rising</u>
		H	L	H L	L H
	μ μ	μ μ	μ μ	μ μ	μ μ

¹⁹ Morén and Zsiga (p. 130) note that the rising tone ends in what they classify as the mid range, and does not reach the target range for high. Nonetheless, the rising tone consistently ends at a higher frequency than the mid tone.

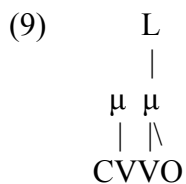
(Morén and Zsiga 2006: 134)

2.4 Analysis of CVVO syllables

Given the representation of tones in (8), the restriction against mid, high, and rising tones on CVVO syllables can be generalised as a requirement for CVVO syllables to end low. In order to rule out CVVO syllables that do not end low, Morén and Zsiga (p. 143) propose C.G.CODA→L, a positive constraint that requires obstruent coda segments to be associated with a low tone:

C.G.CODA→L Constricted glottis coda segments must be associated with L tone.

Since the second mora of a CVVO syllable is assumed to be shared between the vowel and the coda, as shown in (7), C.G.CODA→L is only satisfied when the second mora of a CVVO syllable is associated with a L tone:



On this analysis, the lack of mid tone on CVVO syllables is due to the ranking of C.G.CODA→L above *L, the markedness constraint against L tones:

(10) Neutralisation of mid tone to low tone in CVVO

	/laak/ 'various'	C.G.CODA→L	*L
☞ a.	$ \begin{array}{c} L \\ \\ \mu \mu \\ \ \backslash \\ laak \end{array} $		*
b.	$ \begin{array}{c} \mu \mu \\ \ \backslash \\ laak \end{array} $	*!	

Given a CVVO input with no tonal specification, C.G.CODA→L must outrank *L in order for candidate (10a), which surfaces with a L tone, to be preferred over the faithful candidate in (10b). In other words, it is less costly to have an extra L tone than to risk having an obstruent coda segment that is not associated with a L tone.²⁰

The avoidance of high tones on CVVO syllables is achieved by ranking C.G.CODA→L above both *[TT]_σ and ALIGN-R(T,σ), as shown in tableau (11).

²⁰ One might ask why the markedness constraint *L is used instead of the faithfulness constraint DEP-IO(L), since candidate (10a) involves an unfaithful correspondence between the input and the output. This is certainly an option, but crucially, both *L and DEP-IO(L) must be ranked below C.G.CODA→L: ranking *L above C.G.CODA→L incorrectly predicts (10b) to be the optimal candidate.

*[TT]_σ Two tones within the same syllable domain are prohibited.

ALIGN-R(T,σ) Align the tone at the right edge of the syllable.

Given an input specified with a high tone, it is more harmonic to insert a L tone to satisfy C.G.CODA→L, as in the optimal candidate (11a), than to violate C.G.CODA→L, as in (11b), even though (11a) incurs violations of both *[TT]_σ and ALIGN-R(T, σ). Finally, rising tones on CVVO syllables are ruled out by ranking both C.G.CODA→L and MAX-IO(H) above LINEARITY, shown in tableau (12).

MAX-IO(H) Every H tone in the input has a correspondent in the output.

LINEARITY Preserve the linear order of features and segments.

(11) Neutralisation of high tone to falling tone in CVVO

	/laak-H/ 'to tow'	C.G.CODA→L	*[TT] _σ	ALIGN-R(T, σ)
☞ a.	H L μμ \n laak		*	*
b.	H μμ \n laak	*!		

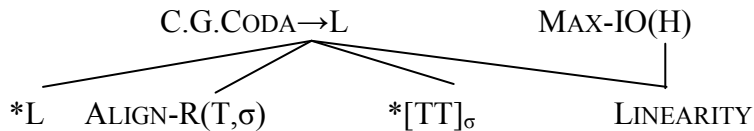
(12) Neutralisation of rising tone to falling tone in CVVO

	/laak-LH/ 'to tow'	C.G.CODA→L	MAX-IO(H)	LINEARITY
☞ a.	H L μμ \n laak			*
b.	LH μμ \n laak	*!		
c.	L μμ \n laak		*!	

Given an input specified with a rising tone, it is preferable to inverse the linearity of the tones in order to satisfy C.G.CODA→L, as in the optimal candidate (12a), than to have the faithful candidate in (12b), which violates C.G.CODA→L, or to delete the H tone as in (12c), which violates MAX-IO(H).

To summarise, the ranking of constraints shown in (13) ensures that mid, high, and rising tones do not occur on CVVO syllables, so that these syllables always end low and surface with either low or falling tone:

(13) Ranking of constraints required for CVVO syllables to end low



3. Reduplication

Having shown how the distributional gaps in CVVO syllables are accounted for in Morén and Zsiga's analysis, I now turn my focus to reduplication. I begin by outlining key theoretical assumptions.

3.1. Theoretical assumptions

The analysis is formalised in the constraint-based framework of OT (Prince and Smolensky 1993). The standard OT model of reduplication is shown in (14):

(14) Standard model of reduplication (McCarthy and Prince 1995: 273)

Input:	/Af _{RED} + Stem/
	↑↓ <i>Stem-IO faithfulness</i>
Output	[Red ↔ Base]
	<i>B-R identity</i>

In this model, input-output (IO) correspondence is evaluated with respect to the stem and the base, while base-reduplicant (B-R) correspondence is evaluated with respect to the base and the reduplicant. This model assumes the definition of the reduplicant, RED, in (15):

(15) Definition of RED (McCarthy and Prince 1994 (Part 1): 2)

RED_k is a morpheme lexically unspecified for segmentism, but requiring a correspondence relation with its base, the phonological structure to which it attaches.

However, Walker (2000) observes that this definition is at odds with the principle of Richness of the Base (Prince and Smolensky 1993: 191), which hypothesises that all inputs are possible. Since OT constraints do not apply to inputs and merely evaluate output candidates, the null hypothesis is that prespecification *could* actually occur in the inputs of reduplicative affixes: otherwise, we would be stipulating restrictions on the input (Walker 2000: 90). Accordingly, she revises the definition of RED:

(16) Revised definition of RED (Walker 2000: 90)

RED_k is a morpheme requiring a correspondence relation with its base, the phonological structure to which it attaches.

Having eliminated the assumption that the reduplicative affix is lexically unspecified, there is now an input form of the reduplicative affix to which the output can correspond. Walker then proposes an elaborated model of reduplication, with correspondence between the input and output forms of the reduplicative affix:

(17) Elaborated model of reduplication (Walker 2000: 90)

Input	/Af _{RED} + Stem/	
<i>Affix_{RED}-IO faithfulness</i>	↑↓	↑↓
Output	[Af _{RED} ↔ Base]	<i>Stem-IO faithfulness</i>
	<i>B-R identity</i>	

This is the model that I will assume in my analysis. Unlike the model in (14), this model evaluates IO correspondences not merely with respect to the stem and base, but also the input and output forms of the reduplicative affix.

The three core families of faithfulness constraints on correspondence relations are given in (18), where S₁ and S₂ refer to strings in a correspondence relation (McCarthy and Prince 1995: 264).

- (18) a. MAX
 Every segment of S₁ has a correspondent in S₂.
 b. DEP
 Every segment of S₂ has a correspondent in S₁.
 c. IDENT(F)
 Let α be a segment in S₁ and β be any correspondent of in S₂. If α is [γF], then β is [γF]. (Correspondent segments are identical in feature F).

McCarthy and Prince cast these constraints in terms of correspondence relations on segments. (19) interprets these in terms of tone:

- (19) a. MAX(T)
 Every tone of S₁ has a correspondent in S₂.
 b. DEP(T)
 Every tone of S₂ has a correspondent in S₁.
 c. IDENT(T)
 Let α be a mora in S₁ and β be any correspondent of in S₂. If α is [γT], then β is [γT]. (Correspondent morae are identical in tonal affiliations).²¹

Since the concern of the present paper is with the form of the reduplicant, it will be unnecessary to focus on stem-IO correspondence, which is covered in the scope of Morén

²¹ Since the TBU is the mora on Morén and Zsiga’s analysis, IDENT(T) has to be assessed with respect to morae and their tonal affiliation.

and Zsiga (2006). I shall thus ignore candidates that involve bases ruled out by Morén and Zsiga's analysis.

Given the model in (17), if a reduplicative affix input came with prespecified input material, it has to be subject to both $\text{Affix}_{\text{RED}}\text{-IO}$ faithfulness constraints and B-R identity constraints. The two general configurations that may arise from the competition between these sets of constraints are shown in (20):

- (20) a. $\text{Faith-BR} \gg \text{Affix}_{\text{RED}}\text{-Faith-IO}$
 b. $\text{Affix}_{\text{RED}}\text{-Faith-IO} \gg \text{Faith-BR}$ (Walker 2000: 90)

The ranking in (20a) yields a pattern in which the reduplicant is a maximally faithful copy of the base, i.e. prespecified material in the reduplicant does not overwrite the copied material from the base. This outcome corresponds to one in which there is no apparent prespecification. On the other hand, the ranking in (20b), which places some or all of $\text{Affix}_{\text{RED}}\text{-Faith-IO}$ above Faith-BR , allows prespecified material to appear in the output of the reduplicative affix at the cost of maximising copied material from the base. The outcome yields prespecified material as the source of fixed segmentism in reduplication. Walker observes that if rankings such as that in (20b) could be eliminated, this would rule out the possibility of the emergence of prespecified material in the reduplicant. Thus, she proposes the following metaconstraint on reduplication correspondence:

- (21) Reduplication correspondence metaconstraint (Walker 2000: 92)
 $\text{Faith-BR} \gg \text{Affix}_{\text{RED}}\text{-Faith-IO}$

Given this metaconstraint, Faith-BR must be ranked above $\text{Affix}_{\text{RED}}\text{-Faith-IO}$, and prespecification is not an option for the source of fixed segmentism.

3.2. *Simple reduplication*

Having laid out the theoretical assumptions, I will now first develop an analysis of simple reduplication in Thai. As seen in section 1.2, tones from the base appear faithfully on the reduplicant in simple reduplication. In an OT analysis, this entails that the Faith-BR constraints outrank the $\text{Affix}_{\text{RED}}\text{-Faith-IO}$ constraints as well as the general markedness constraints in the language. I illustrate with the CVV syllable *kɛ̀ɛ̀* 'old' but the same basic analysis applies for all other syllable types as well, since surfacing as a faithful copy of the base entails that the reduplicant obeys the general distribution of tones in the language. I assume that the tonal alignment in reduplication is no different from that in other parts of the language. On this assumption, tonal alignment on the reduplicant falls out from Morén and Zsiga's analysis, in which $\text{ALIGN-R}(T, \sigma)$ outranks $\text{ALIGN-L}(T, \sigma)$.

The tableau in (22) illustrates the faithful copying of the L tone from the base onto the reduplicant when a tonally unspecified reduplicative affix is attached to *kɛ̀ɛ̀* 'old':

(22) Simple reduplication

	/RED+kεε-L/	MAX-BR(T)	DEP-AF _{RED} (T)	*L
a.	<p style="text-align: center;">L L</p> <p style="text-align: center;"> </p> <p style="text-align: center;">μμ μμ</p> <p style="text-align: center;"> </p> <p style="text-align: center;">kεε – kεε</p>		*	**
b.	<p style="text-align: center;"> L</p> <p style="text-align: center;"> </p> <p style="text-align: center;">μμ μμ</p> <p style="text-align: center;"> </p> <p style="text-align: center;">kεε – kεε</p>	*!		*

The optimal candidate in (22a), which copies the L tone from the base onto the reduplicant, violates DEP-AF_{RED}(T) since the L tone on reduplicant does not have a correspondent in its input. In faithfully reduplicating the base, the candidate in (22a) also incurs an additional violation of *L. In contrast, the candidate in (22b) only violates *L once and does not incur any Affix_{RED}-Faith-IO violations. However, it violates MAX-BR(T) as the L tone on the base does not have a correspondent in the reduplicant. Therefore, in order for the reduplicant to surface as a faithful copy of the base as in (22a), MAX-BR(T) must be ranked above DEP-AF_{RED}(T) and *L. The ranking order MAX-BR(T) >> {DEP-AF_{RED}(T), *L} ensures that a L tone from the base is copied faithfully onto the reduplicant.

The definition of RED in (16) allows prespecification in the reduplicative affix, so it is possible for the reduplicative affix to be tonally specified. This creates three possible scenarios: (i) overwriting, where the prespecified tone on the reduplicative affix replaces the tone that we would otherwise expect to be copied from the base, (ii) the creation of a contour tone, where the prespecified tone on reduplicative affix and the copied tone from the base each occupy one mora on the reduplicant, and (iii) backcopy, where the prespecified tone on the reduplicative affix is copied onto the base. Since simple reduplication simply involves the faithful copying of tones from the base to the reduplicant, these possibilities must be ruled out. For this to happen, Faith-BR must outrank Affix_{RED}-Faith-IO, as predicted by Walker's metaconstraint in (21).

First consider overwriting, illustrated in tableau (23) with a reduplicative affix specified with a H tone. The faithful candidate in (23b) violates MAX-BR(T), since the L tone on the base does not have a correspondent in the reduplicant. However, it does not violate any Affix_{RED}-Faith-IO constraints. On the other hand, the optimal candidate in (23a), which surfaces without the prespecified H tone and instead copies the L tone from the base, violates both DEP-AF_{RED}(T) and MAX-AF_{RED}(T), but not MAX-BR(T). It has been shown in (22) that MAX-BR(T) must be ranked above DEP-AF_{RED}(T). (23) now shows that for (23a) to be the optimal candidate, MAX-BR(T) must also outrank MAX-AF_{RED}(T). Thus overwriting in simple reduplication is ruled out by the ranking order MAX-BR(T) >> {DEP-AF_{RED}(T), MAX-AF_{RED}(T)}.

(23) No overwriting in simple reduplication

/RED-H+kεε-L/	MAX-BR(T)	DEP-AF _{RED} (T)	MAX-AF _{RED} (T)
a. L L μμ μμ kεε – kεε		*	*
b. H L μμ μμ kεε – kεε	*!		

In order to avoid violating MAX-BR(T) like the candidate in (23b), the L tone from the base must be copied onto the reduplicant. However, it is also possible for this to happen without violating MAX-AF_{RED}(T): if the prespecified H tone on the reduplicative affix and the copied L tone from the base each occupy one mora on the reduplicant, forming a derived contour tone, the H tone on the reduplicant input would have a correspondent in its output and MAX-AF_{RED}(T) would be satisfied. This is illustrated in tableau (24). Since both candidates in (24) copy the L tone from the base onto the reduplicant, both violate DEP-AF_{RED}(T) equally. However, in allowing the prespecified H tone to surface on the RED, the candidate in (24b) violates DEP-BR(T), as the prespecified H tone does not have a correspondent in the base. Therefore, in order for (24a) to be the optimal candidate, DEP-BR(T) must be ranked above MAX-AF_{RED}(T). The ranking order DEP-BR(T) >> MAX-AF_{RED}(T) ensures that a derived contour tone does not result on the reduplicant as a result of the concatenation of a prespecified tone and a copied tone.

(24) No derived contour tone in simple reduplication

/RED-H+kεε-L/	DEP-BR(T)	DEP-AF _{RED} (T)	MAX-AF _{RED} (T)
a. L L μμ μμ kεε – kεε		*	*
b. HL L μμ μμ kεε – kεε	*!	*	

Backcopy can be avoided by ranking Stem-Faith-IO above Affix_{RED}-Faith-IO. This is illustrated in (25). Unlike the candidate in (24b), the candidate in (25b) does not violate DEP-BR(T) as each tone on the reduplicant has a correspondent in the base. Therefore, in order to rule out this candidate and prevent a tone that was prespecified on the reduplicative affix from appearing on the base, DEP-IO(T) must be ranked above MAX-

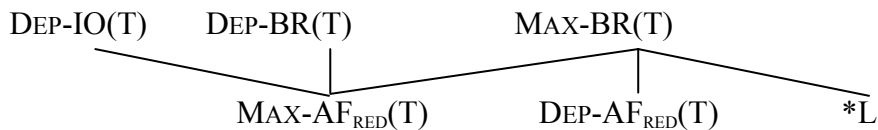
AF_{RED}(T). The ranking order DEP-IO(T) >> MAX-AF_{RED}(T) thus prevents backcopy in simple reduplication.

(25) No backcopy in simple reduplication

	/RED-H+kεε-L/	DEP-IO(T)	DEP-AF _{RED} (T)	MAX-AF _{RED} (T)
a.	L L μμ μμ kεε – kεε		*	*
b.	HL HL μμ μμ kεε – kεε	*!	*	

The ranking of constraints required for simple reduplication is summarised in (26). Generally speaking, in order for the reduplicant to be an exact copy of the base, the Faith-BR constraints must be ranked above the Affix_{RED}-Faith-IO constraints, in accordance with the general configuration in the metaconstraint in (21).

(26) Ranking of constraints for simple reduplication



3.3. Emphatic reduplication

Having presented an analysis of simple reduplication, I turn now to emphatic reduplication. The first question with respect to emphatic reduplication was why the reduplicant always surfaces with a high tone; in other words, why does the reduplicant appear with invariant autosegmental material? The classical approach to dealing with fixed material in the reduplicant is to assume that the reduplicant is prespecified with the fixed content (Marantz 1982, Yip 1982b). On this analysis, the reduplicative affix would simply be prespecified with a H tone when used in emphatic reduplication.

However, recent work within OT has explored the notion that fixed segmentism in reduplication can be explained without prespecification, and that the emergence of any fixed content in the reduplicant is either phonologically determined or morphologically determined (Alderete *et al.* 1999, Walker 2000). Moreover, as shown in (23-25), a prespecified tone on the reduplicative affix cannot be allowed to surface on the reduplicant for simple reduplication to take place. Now, in order to have a unified analysis of both simple and emphatic reduplication, the grammar in (26) presented for simple reduplication must also hold for emphatic reduplication. Therefore, not only is the prespecification of a H tone on the reduplicative affix in emphatic reduplication unnecessary, assuming it would contradict the rest of the grammar of the language. Thus, in line with the spirit of recent work, I will eschew the prespecification approach and explore instead the option of

deriving the fixed autosegmental content in the reduplicant from the premises of OT; reduplicative affixes prespecified with autosegmental content will not be considered from this point on.

Fixed segmentism that has a phonological basis falls under the OT rubric of the emergence of the unmarked (TETU), whereby only unmarked structure is allowed on the reduplicant while the corresponding marked structure is allowed elsewhere in the language (Alderete *et al.* 1999: 328). On Morén and Zsiga's analysis of Thai tones, high tones are marked. If emphatic reduplication were indeed an instance of TETU, we would expect the emphatic reduplicant to surface instead with a mid tone, which is treated as the neutral pitch. This means that if Morén and Zsiga's analysis is correct, the invariant H tone on the reduplicant in emphatic reduplication cannot have a phonological basis.

Morphological fixed segmentism, on the other hand, is a kind of affixation, whereby the fixed content is treated as an affix that is realised simultaneously with the reduplicant, overwriting part of it (Alderete *et al.* 1999: 328). Thus, the identity of the fixed content is morphologically determined. In the case of Thai emphatic reduplication, it is not segmental content that is fixed, but tone, i.e. autosegmental content. In order to pursue the line of phonological reasoning that the invariant high tone that occurs in Thai emphatic reduplication has a morphological basis, I assume that the emphatic high tone is an affix that is realised simultaneously with the reduplicant. As Yip notes, "Although we usually begin by thinking of a morpheme as something with segments, purely tonal morphemes abound" (Yip 2002: 106). On this conception, the emphatic high tone can be treated as a tonal morpheme sans segments that affixes to the reduplicative compound. The concatenation of the emphatic tonal morpheme with the reduplicative compound results in the following model:

- (27) Elaborated model of Thai emphatic reduplication
- | | | |
|--------|--|--|
| Input | $/Af_H + Af_{RED} + Stem/$ | |
| | <i>Affix_{RED}-IO faithfulness</i> $\uparrow\downarrow$ $\uparrow\downarrow$ <i>Stem-IO faithfulness</i> | |
| Output | $[Af_{RED} \leftrightarrow Base]$ | |
| | <i>B-R identity</i> | |

IO faithfulness constraints govern the correspondence relations between the stem and base as well as the input and output forms of the reduplicative affix, while B-R faithfulness constraints regulate the correspondence between the reduplicant and the base. However, these constraints do not apply to the tonal morpheme, since it is not part of the stem, base, or reduplicative affix. Hence, in addition to the correspondence relations between related phonological structures, it is necessary to assume a mapping between morphology and phonology, in order to guarantee the appearance of the tonal morpheme in the phonological output. I adopt a constraint requiring a morpheme to be realised in the phonological content of the output (Samek-Lodovici 1993; Akinlabi 1996; Gnanadesikan 1997; Rose 1997a,b; Walker 1998, 2000):

REALISE- μ A morpheme must have some phonological exponent in the output.

In essence, when applied to emphatic reduplication in Thai, REALISE- μ ensures that the tonal morpheme will be realised in the output.

3.3.1. Fixed Autosegmentism

I now proceed to account for the fixed autosegmentism in Thai emphatic reduplication. Since CVV/CVS/CVVS syllables are the least marked in terms of tonal distribution, it would be useful to begin the analysis with these syllables before moving on to the CVVO syllables.

First consider why the emphatic high tone appears, and why it appears on the reduplicant. Given an input stem that is not specified for tone, e.g. *dii* ‘good’, the reduplicative compound consists of a reduplicant that appears with a high tone and a base that remains phonologically toneless, as in *dii-dii* ‘very good’. In order for the emphatic high tone to appear only on the first half of the reduplicative compound, both REALISE- μ and Stem-Faith-IO must outrank Faith-BR.

(28) compares the optimal candidate with one in which the emphatic tonal morpheme does not surface. The optimal candidate in (28a) violates DEP-BR(T) as the H tone on the reduplicant does not have a correspondent in the base. It also violates DEP-AF_{RED}(T), since the H tone does not have a correspondent in the input of the reduplicant. Candidate (28b) does not incur any faithfulness violations, since the stem, base, and reduplicant stand in perfect correspondence. However, it incurs a violation of REALISE- μ , as the tonal morpheme does not have a phonological exponent in the output. Thus, in order for candidate (28a) to be more harmonic than candidate (28b), REALISE- μ must be ranked above DEP-BR(T) and DEP-AF_{RED}(T). The ranking order REALISE- μ >> {DEP-BR(T), DEP-AF_{RED}(T)} ensures that the emphatic tonal morpheme is realised in the phonological output.

(28) Emphatic high tone appears on the RED

	/H+RED+dii/	REALISE- μ	DEP-BR(T) : DEP-AF _{RED} (T)
a.	<p style="text-align: center;">H $\mu\mu$ $\mu\mu$ d i i – d i i</p>		* *
b.	dii – dii	*!	:

Now compare the optimal candidate with one in which the tonal morpheme appears not on the reduplicant, but on the base, as shown in (29). Both candidates satisfy REALISE- μ . However, the output candidate in (29b) contains a tone in the base that does not have a correspondent in the stem and thus violates DEP-IO(T). (29b) also violates MAX-BR(T), since the H tone on the base does not have a correspondent in the reduplicant. Thus, in order for (29a) to be more harmonic than (29b), either DEP-IO(H) or MAX-BR(T) must outrank DEP-BR(T), so that it is preferable to have the tonal morpheme realised on the reduplicant than on the base.

(29) Emphatic high tone appears on the red instead of the Base

/H+RED+dii/	DEP-IO(T)	MAX-BR(T)	DEP-BR(T)
a. H μμ μμ d i i - d i i			*
b. H μμ μμ d i i - d i i	*!	*(!)	

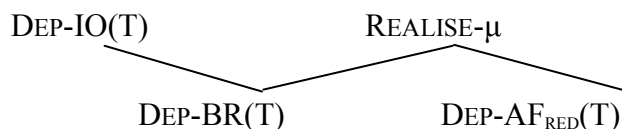
The ranking order DEP-IO(H) >> DEP-BR(T) is also relevant in ensuring that there is no backcopy, and that the emphatic high tone does not appear on the base to match the tone on the reduplicant, as shown in (30). Even though candidate (30b) satisfies Faith-BR, it violates Stem-Faith-IO and is less harmonic than the optimal candidate in (30a) which violates DEP-BR(T). Thus, DEP-IO(H) >> DEP-BR(T) ensures that backcopy does not occur in emphatic reduplication. Since this ranking is independently required, I assume, for reasons of parsimony, that the violation of DEP-IO(H) is the fatal violation in (29), and leave open the possibility of MAX-BR(T) and DEP-BR(T) being unranked with respect to each other.

(30) No backcopy

/H+RED+dii/	DEP-IO(T)	DEP-BR(T)
a. H μμ μμ d i i - d i i		*
b. H H μμ μμ d i i - d i i	*!	

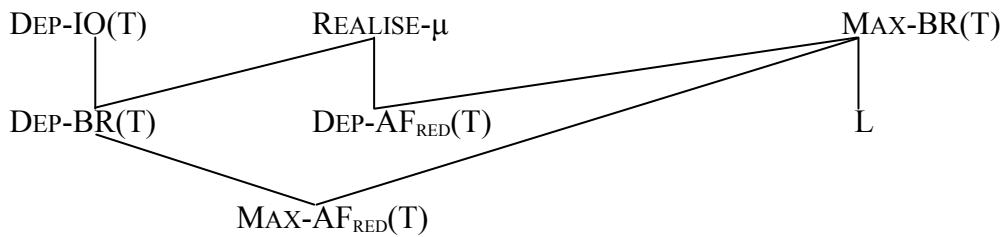
To summarise thus far, fixed autosegmentism in the emphatic reduplicant is accounted for by the ranking of constraints in (31):

(31) Ranking of constraints for fixed autosegmentism in emphatic reduplication



Combining this with the grammar in (26), the interim ranking of constraints required for reduplication in Thai is shown in (32):

(32) Interim ranking of constraints for Thai reduplication



(28-30) illustrated fixed autosegmentism with a stem that is unspecified for tone. Now consider emphatic reduplication that involves a stem that is specified for tone. Given an input stem specified for low tone, such as *ʔim* ‘full’, the reduplicative compound consists of a reduplicant that appears with a high tone and a base that retains the low tone, as in *ʔim-ʔim* ‘very full’. Unlike the case of the toneless base in (28-30), where the emphatic high tone is merely added onto the reduplicant, the emphatic high tone now overwrites the tone on the reduplicant that we would have otherwise expected to be copied from the base when the base is specified for tone. Since high and low tones are, on Morén and Zsiga’s analysis, specified on only one mora, one wonders why the concatenation of the high tonal morpheme and the low tone from the base does not result in a contour tone on the reduplicant, e.g. **ʔim-ʔim*. Such a candidate, which copies the base tone onto the reduplicant, would better satisfy MAX-BR(T) than one for which overwriting occurs, while satisfying REALISE-μ at the same time. (33) illustrates.

(33) Derived contour in emphatic reduplication?

/H+RED+ʔim- L/	IDENT- BR(T)	DEP- BR(T)	MAX- BR(T)	DEP- AF _{RED} (T)	*L
a. H L μ μ μ μ ʔ i m - ʔ i m	*	*	*!	*	*
☞ b. HL L μ μ μ μ ʔ i m - ʔ i m	*	*		**	**

Both candidates in (33) satisfy REALISE-μ, since the emphatic tonal morpheme surfaces on the reduplicant. Both candidates also violate IDENT-BR(T) and DEP-BR(T) equally: in each case, (i) there is one set of corresponding morae that do not have identical tonal affiliations, and (ii) the emphatic tonal morpheme that surfaces on the reduplicant does not have a correspondent in the base. The candidate in (33a) violates MAX-BR(T) as the L tone on the base has no correspondent in the reduplicant. In contrast, the candidate in

(33b) satisfies MAX-BR(T), at the expense of incurring additional penalty for violating DEP-AF_{RED}(T) and *L. However, as the latter two constraints have been established in (22) to be ranked below MAX-BR(T), these extra violations do not count against the candidate in (33b). Therefore, while the markedness constraint *L penalises the candidate in (33b), it is ranked too low to affect the choice of the optimal candidate. On the other hand, the faithfulness constraint MAX-BR(T) penalises candidate (33a) and not (33b). This leads to the incorrect predictions that the candidate in (33b) is the optimal candidate, and that overwriting does not occur in emphatic reduplication.

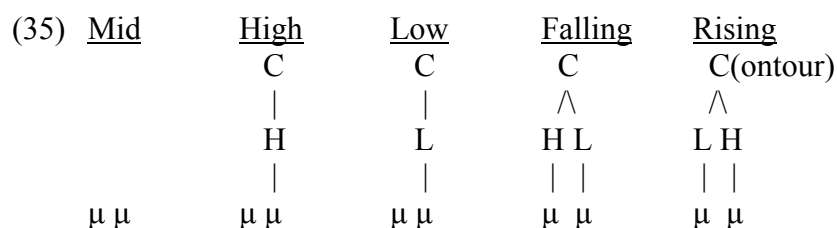
3.3.2. Feature geometry

It appears that neither markedness nor faithfulness constraints can account for the overwriting phenomenon in emphatic reduplication. What then is the explanation for the overwriting? I propose that we need to refer to the internal feature geometry of the tonal features. Bao (1990) proposes the model of internal feature geometry of tonal features in (34):



In this model, Tone features are dominated by a node of their own, called a Contour, which is a sister of the Register feature. The Contour node and Register feature are in turn dominated by a Tonal node. For present purposes, it is not necessary to consider the Register feature or the Tonal node; what is relevant is the idea that individual tones can form a constituent, as represented in the part of the model highlighted in bold.

Assuming that the individual tones within a syllable form a constituent, the constituent as a whole can be assessed for constraint violation. This entails a modification of Morén and Zsiga's proposed representation of Thai tones, illustrated in (35):



On this analysis, both the emphatic tonal morpheme and the tone on the base project their own Contour nodes. Since only the emphatic high tone can occur on the reduplicant, overwriting can be understood as the result of a prohibition against multiple Contour nodes within a syllable:

*[CC]_σ Two Contour nodes within the same syllable domain are prohibited.

By assuming the elaborated representation of tones in (35) and the markedness constraint $*[CC]_{\sigma}$, overwriting in emphatic reduplication can now be accounted for, as shown in (36):

(36) Overwriting in emphatic reduplication

	/H+RED+?im-L/	$*[CC]_{\sigma}$	MAX-BR(T)
a.	C C H L μ μ μ μ ? i m - ? i m		*
b.	C C C H L L μ μ μ μ ? i m - ? i m	*!	

As we saw earlier, the optimal candidate in (36a) violates MAX-BR(T). This constraint is not violated in candidate (36b), which copies the L tone from the Base. However, candidate (36b) incurs a violation of $*[CC]_{\sigma}$. This is because the entire Contour node projected by the L tone is copied from the base, while the emphatic high tone projects its own Contour node, resulting in two Contour nodes within the single syllable domain of the reduplicant. By ranking $*[CC]_{\sigma}$ above MAX-BR(T), the analysis correctly chooses (36a) as the optimal candidate.

Thus, feature geometry seems to provide a plausible explanation for the overwriting phenomenon. Notice that postulating the elaborated representation in (35) does not necessarily affect Morén and Zsiga’s analysis in any way, since one could still make reference to the individual tone features in the assessment of constraint violation, i.e. we could simply retain Morén and Zsiga’s analysis while ranking the faithfulness constraints on Contour nodes low enough so that their impact is not felt in the grammar.

3.3.3. H tone on CVVO syllables

The second question that this paper seeks to answer is why the emphatic high tone is permitted on CVVO syllables when such a combination is otherwise prohibited. That the emphatic reduplicant can surface with a high tone must mean that the overwriting process does not only override MAX-BR(T), but also the requirement for these syllables to end low. On Morén and Zsiga’s analysis, the positive constraint C.G.CODA→L rules out CVVO syllables that do not end low. Therefore, in order for CVVO syllables to surface with a H tone in emphatic reduplication, the two constraints REALISE-μ and $*[CC]_{\sigma}$ must be ranked higher than C.G.CODA→L. Consider (37), which illustrates overwriting in CVVO syllables in emphatic reduplication:

(37) Overwriting in CVVO in emphatic reduplication

/H+RED+CVVO/	REALISE- μ	C.G.CODA \rightarrow L	DEP-BR(T)	MAX-BR(T)
a. H L μ μ μ μ \ \ CVVO - CVVO		*	*	*
b. L L μ μ μ μ \ \ CVVO - CVVO	*!			

By not copying the L tone from the base onto the reduplicant, the candidate in (37a) violates both MAX-BR(T) and C.G.CODA \rightarrow L. Additionally, it also incurs a violation of DEP-BR(T) since the emphatic tonal morpheme has no correspondent in the base. The candidate in (37b) on the other hand, does not violate any of these constraints: it ends on a L tone, and the tones on the base and the reduplicant are in correspondence with each other. However, it fails to realise the emphatic tonal morpheme and thus violates REALISE- μ . Therefore, in order for (37a) to be chosen as the optimal candidate, REALISE- μ must be ranked above C.G.CODA \rightarrow L, DEP-BR(T), and MAX-BR(T). The ranking order REALISE- μ >> {C.G.CODA \rightarrow L, DEP-BR(T), MAX-BR(T)} guarantees that overwriting takes place in CVVO syllables during emphatic reduplication so that the emphatic high tone surfaces on the CVVO reduplicant.

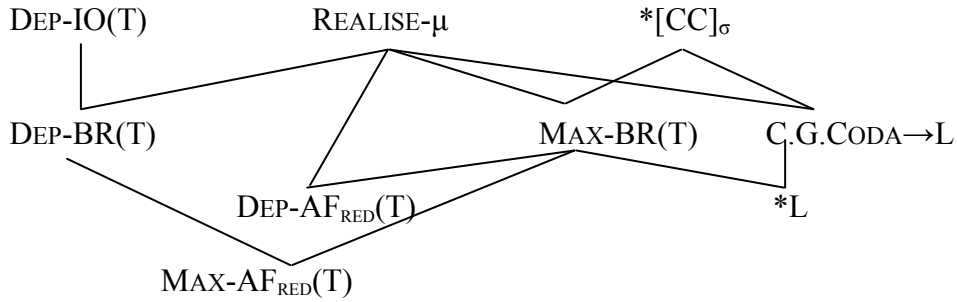
Besides having to ensure that the emphatic tonal morpheme appears on the reduplicant, we also need to prevent a contour tone from occurring on the reduplicant due to the concatenation of the emphatic tonal morpheme and the tone copied from the base. For this to happen, *[CC] $_{\sigma}$ must be ranked above C.G.CODA \rightarrow L, so as to resist the tendency of CVVO syllables to end low. This is illustrated in tableau (38):

(38) No derived contour tone in emphatic reduplication

/H+RED+CVVO/	*[CC] $_{\sigma}$	C.G.CODA \rightarrow L	MAX-BR(T)
a. C C H L μ μ μ μ \ \ CVVO - CVVO		*	*
b. C C C H L L μ μ μ μ \ \ CVVO - CVVO	*!		

The candidate in (38a) violates both MAX-BR(T) and C.G.CODA→L by not copying the L tone from the base onto the reduplicant. On the other hand, neither of these constraints is violated by the candidate in (38b), which copies the L tone. However, by copying the L tone along with its Contour node, while having the emphatic tonal morpheme and its associated Contour node appear within the same syllable domain, (38b) violates *[CC]_σ. Thus, to rule out (38b) in favour of (38a), *[CC]_σ must be ranked above MAX-BR(T) and C.G.CODA→L. The ranking order *[CC]_σ >> {MAX-BR(T), C.G.CODA→L} prevents a derived contour tone from occurring on a CVVO syllable during emphatic reduplication.

(39) Ranking of constraints for Thai reduplication



Therefore, the occurrence of H tone on CVVO syllables in emphatic reduplication falls out naturally from the existing tools of the analysis, providing additional evidence for the utility of the modified representation in (35). This completes the rankings that sanction the reduplication patterns in Thai, summarised in (39).

4. Conclusion

This paper presented an analysis of reduplication in Thai, with special emphasis on two issues related to emphatic reduplication: fixed autosegmentism and the occurrence of high tone on CVVO syllables. In the spirit of work in OT showing that fixed segmentism can be accounted for without appealing to prespecification, the first goal of this paper was to provide an account of fixed autosegmentism in Thai emphatic reduplication without prespecification. Indeed, in order to have a coherent account of both simple and emphatic reduplication, one *cannot* appeal to tonal prespecification in the reduplicative affix. This is because in order for simple reduplication to take place successfully, the faithfulness constraints on base-reduplicant identity must be ranked higher than the faithfulness constraints on the reduplicative affix. In turn, this ranking configuration rules out prespecification as the source of fixed autosegmentism in emphatic reduplication. I argued that the emergence of the invariant high tone is morphologically determined and proposed that the emphatic high tone that occurs on the reduplicant has its origins in a tonal morpheme that attaches to the reduplicative compound. As long as REALISE-μ, which ensures that a morpheme has a phonological exponent in the output, remains undominated, the emphatic tonal morpheme is guaranteed to be realised in the phonological output. In order to ensure that overwriting takes place in emphatic reduplication so that the reduplicant surfaces with an invariant high tone, I proposed a modification to Morén and Zsiga’s representations of Thai tones by allowing individual tones within a syllable domain to form a constituent, or Contour node. On this analysis, both the emphatic tonal morpheme and the tone on the base project their own Contour nodes, and overwriting in emphatic reduplication is analysed as the result of a prohibition against multiple Contour

nodes occurring within the same syllable domain, *[CC]_σ. With regard to the puzzling phenomenon of high tone bearing CVVO syllables that results from emphatic reduplication, I showed that the occurrence of the emphatic high tone on CVVO syllables falls out naturally from the interaction of *[CC]_σ and C.G.CODA→L, thereby providing additional motivation for and utility to the modified representation of Thai tones.

References

- Abramson, Arthur. 1962. The Vowels and Tones of Standard Thai: Acoustical Measurements and Experiments. *International Journal of American Linguistics* 28, 2, part II.
- Akinlabi, Akinbiyi. 1996. Featural affixation. *Journal of Linguistics* 32: 239-289.
- Alderete, John, Jill Beckman, Laura Benua, Amalia Gnanadesikan, John McCarthy and Suzanne
- Urbanczyk. 1999. Reduplication with fixed segmentism. *Linguistic Inquiry* 30: 327-364.
- Bao, Zhiming. 1990. *On the nature of Tone*. PhD dissertation, MIT.
- Gandour, Jackson. 1974. On the Representation of Tone in Siamese. In J.G. Harris and J.R. Chamberlain (eds.), *Studies in Tai Linguistics in Honor of William J. Gedney*. Central Institute of English Language, Bangkok: 170-195.
- Gandour, Jackson. 1977. On the Interaction Between Tone and Vowel Length: Evidence from Thai Dialects. *Phonetica* 34: 54-67.
- Gandour, Jackson, Siripong Potisuk, Suvit Ponglorpisit and Sumalee Dechongkit. 1991. Inter-and Intraspeaker Variability in Fundamental Frequency of Thai tones. *Speech Communication* 10: 355-372.
- Gnanadesikan, Amalia. 1997. *Phonology with ternary scales*. PhD dissertation, University of Massachusetts, Amherst.
- Haas, Mary. 1946. Techniques of Intensifying in Thai. *Word* 2: 127-130. Reprinted in Haas 1978.
- Haas, Mary. 1978. *Language, Culture, and History*. Stanford University Press: Stanford, California.
- Iwasaki, Shoichi and Preeya Ingkaphirom. 2005. *A Reference Grammar of Thai*. Cambridge University Press: Cambridge, UK.
- Marantz, Alec. 1982. Re reduplication. *Linguistic Inquiry* 13: 435-482.
- McCarthy, John and Alan Prince. 1994. An overview of Prosodic Morphology. Parts 1 and 2. Papers presented at the OTS/HIL Workshop on Prosodic Morphology, Utrecht.
- McCarthy, John and Alan Prince. 1995. Faithfulness and Reduplicative Identity. In *University of Massachusetts Occasional Papers in Linguistics 18: Papers in Optimality Theory*, 249-384. GLSA, University of Massachusetts, Amherst.
- Morén, Bruce and Elizabeth Zsiga. 2006. The Lexical and Post-Lexical Phonology of Thai Tones. *Natural Language and Linguistic Theory* 24: 113-178.
- Ohala, John J. and William G. Ewan. 1972. Speed of Pitch Change. *Journal of the Acoustical Society of America* 53: 345.
- Prince, Alan and Paul Smolensky. 1993. *Optimality Theory: constraint interaction in generative grammar*. Ms, Rutgers University and University of Colorado, Boulder.

- Rose, Sharon. 1997a. Theoretical issues in comparative Ethio-Semitic phonology and morphology. PhD dissertation, McGill University.
- Rose, Sharon. 1997b. Multiple correspondence in reduplication. In Bailey, Ashlee C, Kevin E. Moore, and Jeri L. Moxley (eds.), *Proceedings of the Annual Meeting of the Berkeley Linguistics Society* 23 :315-326.
- Samek-Lodovici, Vieri. 1993. Morphological gemination crosslinguistically. Paper presented at the Rutgers Optimality Workshop.
- Sundberg, Johan. 1973. Data on Maximum Speed of Pitch Changes. *STL Quarterly Progress and Status Report* 4: 39-47.
- Sundberg, Johan. 1979. Maximum Speed of Pitch Changes in Singers and Untrained subjects. *Journal of Phonetics* 7: 71-79.
- Walker, Rachel. 1998. *Nasalization, neutral segments, and opacity effects*. PhD dissertation, University of California, Santa Cruz.
- Walker, Rachel. 2000. Nasal reduplication in Mbe affixation. *Phonology* 17: 65-115.
- Xu, Yi. 1999a. F0 Peak Delay: When, Where and Why it Occurs. In J. Ohala (ed.), *International Congress of Phonetic Sciences 1999*: 1881-1884.
- Xu, Yi. 1999b. Effects of Tone and Focus on the Formation and Alignment of F0 Contours. *Journal of Phonetics* 27: 55-105.
- Yip, Moira. 1982a. Against a Segmental Analysis of Zhaoh and Thai: A Laryngeal Tier Proposal. *Linguistic Analysis* 9: 79-94.
- Yip, Moira. 1982b. Reduplication and C-V skeleta in Chinese secret languages. *Linguistic Inquiry* 13: 637-661.
- Yip Moira. 2002. *Tone*. Cambridge University Press: Cambridge.
- Zhang, Jie. 2002. *The Effects of Duration and Sonority on Contour Tone Distribution: A Typological Survey and Formal Analysis*. Routledge: New York.

SORBUNG, AN UNDOCUMENTED LANGUAGE OF MANIPUR: ITS PHONOLOGY AND PLACE IN TIBETO-BURMAN

David Mortensen²²
University of Pittsburgh
<drm31@pitt.edu>

Jennifer Keogh
Independent Scholar
<jennifer.r.keogh@gmail.com>

Abstract

Sorbung is a Tibeto-Burman language of Manipur, India. This study outlines the phonology of the language based on data elicited from a speaker of residing temporarily in the United States, where he was studying theology. Sorbung shares characteristics with at least two branches of Tibeto-Burman—Tangkhulic and Kuki-Chin—without showing unambiguous evidence of belonging to either group. A word list based upon this collection of data is provided as an appendix.

Keywords: phonology, classification, Tibeto-Burman

1. Introduction

Sorbung is a previously undocumented Tibeto-Burman language of Manipur, India. The current study outlines the major features of the phonology of the language, describing the segment inventories, prosodic structures, phonotactic constraints, and phonological alternations that characterize Sorbung. This sketch is based on lexical and other data elicited from a single native speaker of Sorbung (a male in his thirties) residing temporarily in the United States, where he was studying theology²³. Elicitation was conducted

²² The original field research on Sorbung was supported by a UC Berkeley Summer Research grant (2004) to the first author.

²³ Although we have no reason to doubt that our consultant's speech was typical of the speech of the broader language community, caution should be used in interpreting these results. Like most other speakers of Sorbung in his age group, our consultant was proficient to some degree in Standard Tangkhul, Meithei, Hindi, and English. Some instances of language interference are therefore likely. Since his spouse was not from Sorbung, he spoke to her and his children primarily in Standard Tangkhul. He had resided in the United States for approximately three years at the time the first author worked with him and his use of Sorbung was primarily confined to telephone conversations with family members during that time. However, he expressed considerable confidence in his Sorbung language competence, and was both prompt and consistent in his responses to elicitation prompts. Unfortunately, since the current political situation in Manipur makes fieldwork difficult and dangerous, it is unlikely that work on Sorbung based

David Mortenson and Jennifer Keogh. 2011. Sorbung, an undocumented language of Manipur: its phonology and place in Tibeto-Burman.

Journal of the Southeast Asian Linguistics Society 4.1:64-114

Copyright vested in the author

Received 29/9/10, revised text accepted 26/3/11

primarily in English but Standard Tangkhul and visual illustrations were used for difficult forms. Audio recordings were made of most of the lexical material elicited from the consultant, and these were used to verify and correct the transcriptions made while the first author was working with him. A word list based upon this collection of data is provided as an appendix.

Sorbung is phonologically interesting for a number of reasons. From a comparative standpoint, it shares characteristics with at least two branches of Tibeto-Burman—Tangkhulic (Mortensen & Miller 2009a; Mortensen & Miller 2009b) and Kuki-Chin (VanBik 2009)—without showing unambiguous evidence of belonging to either group. From a synchronic point of view, it displays a variety of productive alternations (both segmental and tonal) that may be of theoretical interest. It also shows interesting patterns of synchronic variation that may provide evidence regarding diachronic developments that have occurred elsewhere in Tibeto-Burman, specifically the development of manner distinctions from the “collapse” of lexical prefixes into the onsets of roots.

1.1. Geographic and Demographic Information

Sorbung is spoken in Sorbung village, a community of about 300 located in the southeast corner of Ukhrul District, Manipur State, India (about 40 km east of the border with Burma). Despite some significant linguistic and cultural differences, the inhabitants of Sorbung village consider themselves to be ethnic Tangkhuls, the majority ethnic group in Ukhrul District. The immediately neighboring villages are also ethnic Tangkhuls, though there is significant Kuki (Thadou) settlement that separates Sorbung from the larger group of Tangkhul villages to the north. According to our consultant, many of these villages are homes to dialects closely related to Sorbung. Others, like Tusom village, are inhabited by speakers of languages belonging unambiguously to the Tangkhulic group. To the south and southwest is a network of Maring and Old Kuki villages. The linguistic and ethnic affiliation of the villages to the east is not clear.

It is not well established whether Tangkhulic (or Kuki-Chin) are separate top-level branches of Tibeto-Burman or whether they form a subgroup with one another or other Tibeto-Burman languages. Traditionally, they have been grouped together as part of a Kuki-Chin-Naga group. However, as Burling (2003) notes, this subgroup has been based more on intuition than argument. Mortensen and Miller (2009b) present some evidence for innovations shared among Kuki-Chin, Tangkhulic, and an additional subset of “Naga” languages and suggest that they share a common ancestor more recent than Proto-Tibeto-Burman. Since there is no consensus on this subject, we will treat Tangkhulic and Kuki-Chin only as distinct taxonomic units, without specifying the relationship between them.

Like other ethnic Tangkhul, the inhabitants of Sorbung village are multilingual. Most aspects of family and village life are negotiated in the local language (here called Sorbung). Church sermons and some schooling are carried out in the ethnic lingua franca we call Standard Tangkhul (a Tangkhulic language originally based on the speech of Ukhrul village, now the district headquarters). Other schooling is done in Meithei—the dominant regional language, Hindi, and English. English is important not only because of its international stature, but because the overwhelming majority are Baptist Christians with a

on the speech of a larger number of individuals will be available for some time unless it is carried out by local linguists.

persisting historical relationship to the American Baptist denomination. Loanwords from all of these languages appear in the Sorbung lexicon.

1.2. Previous Work

There is no previous work on Sorbung directly, but at least one published data set from a different dialect of the same language exists. Brown (1837) published word lists collected by Capt. George Gordon, a British political agent in Manipur, for a variety of languages in Manipur. These included “Luhuppa,” which is largely identical to the Tangkhul language of Ukhrul town, “Champung,” the language of another Tangkhul village, and three varieties explicitly labeled as Tangkhul: “Northern Tangkhul,” “Central Tangkhul,” and “Southern Tangkhul.” Of these, Southern Tangkhul is the least like the others but very similar to Sorbung in its lexical, morphological, and phonological properties. Unfortunately, Brown’s word lists were only 60 words long, and provide the only published information on this language up until the present.

Mortensen (2003) and Mortensen and Miller (2009a; 2009b) show that most of the languages spoken by members of the Tangkhul tribe belong to a single subgroup within Tibeto-Burman. They establish criteria for membership in this group, “Tangkhulic,” on the basis of phonological, morphological, and lexical innovations. All of the Tangkhul languages enumerated by Brown (1837) appear to belong to this group, with the exception of “Southern Tangkhul.” Our more complete corpus of data from Sorbung makes it clear that it does not descend from the Proto-Tangkhulic reconstructed by Mortensen and Miller. While Sorbung speakers belong to the same ethnic group as the speakers of Tangkhulic language, their language appears to come from a somewhat different branch of the Tibeto-Burman family.

1.3. The Place of Sorbung within Tibeto-Burman

It is well known that ethnic and linguistic boundaries do not align neatly. This is no less true in South and Southeast Asia than elsewhere. It appears likely that Tangkhuls, as an ethnicity, may be the result of the convergence of two or more populations already related in language and culture. One of these comprises, roughly, the forebears of Tangkhulic speakers; another comprises the ancestors of Sorbung and “Southern Tangkhul” speakers. We contend that this second group spoke a language closely related to the so-called “Old Kuki” languages that are spoken to the south and west of Sorbung. While the population around Sorbung came to identify ethnically with population of Tangkhulic speakers, and while their language was subject to some lexical and structural influence from Tangkhulic languages, they retain a linguistic tradition distinct from their northern neighbors. This was facilitated, no doubt, by the high degree of linguistic heterogeneity present even among Tangkhulic languages. It is apparent that Sorbung and “Southern Tangkhul” fall outside of Tangkhul as a group only when they are compared with Old Kuki languages like Kom (Kom Rem).

This is not to say that Sorbung does not resemble Tangkhul languages in some respects (only that its resemblance to Old Kuki languages is stronger). Like Tangkhul but unlike most Kuki-Chin languages, Sorbung shows no evidence of subject agreement in verbs. It also shows little or no evidence of verbal stem alternation, an important Kuki-

Chin feature²⁴. Phonologically, too, Sorbung resembles Tangkhul in some respects. For example PTB *kr-, *kl- become c- in Sorbung, just as they did in Proto-Tangkhulic. Compare the following Sorbung forms with related reconstructions from Proto-Tangkhulic (PTk) and Proto-Tibeto-Burman (PTB).²⁵

Table 1: *Palatal reflexes of velar clusters in Proto-Tangkhul and Sorbung.*

PTB	PTK	SORBUNG	
*krap	*cap	caap	‘cry; weep’
*kri(y)	*ci	-cii	‘fear’
*klaw	*cow	ciu	‘dig’

However, in Tangkhulic, this appears to be a general process of palatalization before palatal vowels and glides, medial liquids having previously become palatal glides (Mortensen & Miller 2009b). In Sorbung, it is limited to PTB *kr- and *kl- specifically; *k- does not become c- before high front vowels, when has undergone secondary aspiration, does not become c- even before medial *-l-:²⁶

Table 2: *Forms displaying palatization in Tangkhulic but not Proto-Kuki-Chin or Sorbung.*

PTB	PTK	PKC	SORBUNG	
*klum	*ʃim	*khlum	k ^h uum	‘sweet’
*g-ləy	*-ʃi	*khlīi	k ^h ii	‘wind (n.)’
*g-la	—	*khlaa	k ^h aa	‘moon/month’
—	—	*khluaq	k ^h uaq	‘brain’
—	*-ci	*kii	kii	‘horn’
*kəy		*sa-kii	səkii	‘deer’
*g(y)əy	*-ci	*kii (PNC)	bəkii	‘parrot’

Furthermore, when lexical items with palatal stops are compared with their cognates in the Old Kuki languages Kom and Moyon, and the lexically-similar language Puiron²⁷, it is apparent that these languages share the same pattern of palatalization as Sorbung:

²⁴ Many Kuki-Chin languages have a morphological process that derives an inflected stem form (Form II) from the basic form (Form I). For a useful comparative study of this phenomenon, see Button (2009:182ff). Note, however, that there are Old Kuki languages which display evidence of neither subject agreement nor stem alternation (Grierson 1903:288-290).

²⁵ Except where otherwise indicated, Proto-Tibeto-Burman reconstructions are from Matisoff (2003) and Proto-Tangkhulic reconstructions are from Mortensen & Miller (2009b).

²⁶ Except where otherwise indicated Proto-Kuki-Chin (PKC) reconstructions are from VanBik (2009). Reconstructions labeled “DRM” are due to the first author.

²⁷ Ethnically, the speakers of Puiron are a subset of the Nruangmei tribe, most of whom speak languages belonging to the Zeme group (Burling 2003). Burling suggests that Puiron is a Zeme language with many Kuki loans. Below we suggest that Puiron is, like Sorbung, a “Nagafied” Old Kuki language.

Table 3: *Parallels in palatalization between Sorbung, Old Kuki (Kom and Moyon), and Puiron.*

PTB	PKC	SORBUNG	KOM	MOYON	PUIRON	
*krap	*krap	<i>càap</i>	<i>kə-čəp</i>	—	<i>cap</i>	‘cry/weep’
*kri(y)	*kriʔ	<i>ʔənci</i>	—	—	<i>kaci</i>	‘fear’
*klaŋ	*kroŋ (DRM)	<i>cóŋ</i>	<i>toŋ rih</i>	<i>lcoŋ</i>	—	‘speech’
*tsywap	*tsuap	<i>ʔəciap</i>	‘spleen’	<i>əčip</i>	—	‘lung’
*m-t(s)i	*tsii	<i>məcii</i>	—	—	<i>maci</i>	‘salt’
*dzyup	—	<i>cúup</i>	<i>čuču</i>	<i>cucu?</i>	—	‘breast’

In other words, palatal reflexes of PTB *kr-, *kl-, etc. cannot be used to associate Sorbung with Tangkhulic in preference to Kuki-Chin.

Another area of similarity between Tangkhulic and Sorbung is their personal pronouns. In this respect, Sorbung actually is quite different from most Kuki-Chin languages but similar to Tangkhulic languages (when the family is taken as an aggregate). **Table 4** compares the Sorbung person pronouns (both the bound-root forms and the free forms) with their cognates in Proto-Tangkhul:

Table 4: *Comparison of Sorbung and Proto-Tangkhulic pronominal forms.*

	FREE PRONOUNS		BOUND PRONOMINAL PROCLITICS	
	SORBUNG	PTK	SORBUNG	PTK
FIRST	ʔoo (<*ej)	*ʔej, *ʔi	ʔi-, ʔəN- (<*ŋ- ?)	*ʔi-
SECOND	naŋ	*naŋ	na-	*na-
THIRD	maa	*mə, *ʔa	ma-	*ʔa-

Apart from the third person proclitic, each of the Sorbung pronouns has a clear cognate in Tangkhulic. This is unsurprising for the second person pronouns, which are widespread in Tibeto-Burman (Matisoff 2003) and are also found in Kuki-Chin languages. For first and third person pronouns, however, Kuki-Chin languages usually have both free forms and proclitics from *kai ‘1’ and *ʔa ‘3’ (La Polla 2006). However, this match does not tie Sorbung to Tangkhulic exclusively. Meithei pronouns match those in Sorbung even more exactly:

Table 5: *Comparison of Sorbung and Meithei personal pronouns.*

	FREE PRONOUNS		BOUND PRONOMINAL PROCLITICS	
	SORBUNG	MEITHEI	SORBUNG	MEITHEI
FIRST	ʔoo (<*ej)	əy	ʔi-, ʔəN-	i-
SECOND	naŋ	nəŋ	na-	nə-
THIRD	maa	má	ma-	mə-

The idea that the Sorbung pronouns (and possibly, those of Tangkhulic) have been borrowed from the politically and economically powerful Meithei-speaking community cannot be dismissed out of hand. It is also true that all of the Sorbung pronouns seem to have cognates among Kuki-Chin languages. Consider, for example, the pronouns of Kuki Thadou (Hyman 2007b):

Table 6: Comparison of Sorbung and Kuki Thadou personal pronouns.

	FREE PRONOUNS		BOUND PRONOMINAL PROCLITICS	
	SORBUNG	KUKI THADOU	SORBUNG	KUKI THADOU
FIRST EXCLUSIVE	ʔoo (<*ej)	kei	ʔi-, ʔəN-	ka-
FIRST INCLUSIVE		ei		i-
SECOND	naŋ	naŋ	na-	na-
THIRD	maa	ama	ma-	a-

The proclitic *i-* marks ‘first person inclusive’ in some Kuki-Chin languages. The free first person pronoun in Puiron is *ai*, probably cognate with Sorbung *ʔoo*. In many Old Kuki languages, the third-person pronoun consists of *a- ‘3’ plus a formative like *ma*: Hallam *a-ma*, Aimol *a-ma*, Kolren *a-ma*. In Purum, this formative acts by itself as the marker of third person, as in *mo-ju* ‘3-NOM’. The same is true of Lamgang, where the 3sg prefix is *ma-* (Grierson 1903:282-284). In summary, the evidence supporting a special relationship between Sorbung and Tangkhulic breaks down under examination. It is unusual for Sorbung to share a characteristic with Tangkhulic unless the same characteristic is shared with Old Kuki languages.

The term “Old Kuki” is not without its complexities. As an ethnic term, it has been used to describe the Kuki-identified peoples who settled in Manipur at a relatively early time, prior to the arrival of the larger and more powerful Kuki groups like the Mizo (Lushai) and Thadou. These include the Anal, Aimol, Chothe, Purum, Koiren, Kom (Kom Rem), Lamkang, Moyon, Chiru, Tarao, and Vaiphei. Some of these groups, for political reasons, now identify as Naga (e.g. the Lamkang and the Tarao). We propose that certain other Naga groups, such as the Puiron (a ethnic subgroup of the Nruangmei) are the descendents of outlying Old Kuki groups that affiliated themselves ethnically with their Naga neighbors. It is not completely clear what groups should be characterized as Old Kuki and whether the languages of these people form a genetic subgroup (p.c., Kenneth VanBik). The internal coherence of Old Kuki was assumed in LSI, and this assumption has been carried on in part because of the lack of reliable data with which to test it.

VanBik, in his dissertation on Proto-Kuki-Chin, tentatively classified Old Kuki languages as part of his Northern (Zo) group, which is, in turn, part of his Peripheral group. This grouping is based on lexical resemblances between Northern Chin languages like Thadou and Paite and the Old Kuki language Purum. However, this speculation was removed from the published version of this work (VanBik 2009). We argue that the old Kuki languages cannot be part of VanBik’s Northern subgroup. The criterial innovation defining the Peripheral group is the fortition of PKC *r- > *g-. However, there is no evidence of this change in Old Kuki languages (or in Sorbung). Etyma with PKC *r- retain *r-* in Anal, Kom, Moyon, Puiron, and Sorbung, as they do in Mizo (a Central Kuki-Chin language) but develop *g-* in Tedim, a Northern Kuki-Chin language, as shown in **Table 7**.

Table 7: Reflexes of Proto-Kuki-Chin *r- and *hr- in Tedim, Old Kuki (Anal, Kom, and Moyon), Puiron, and Sorbung.

PKC	MIZO	TEDIM	ANAL	KOM	MOYON	PUIRON	SORBUNG	
*ruʔ	<i>ruʔ</i>	<i>guʔ</i>	<i>rú</i>	<i>ru</i>	<i>row</i>	<i>ru</i>	<i>ʔə-rúu</i>	‘bone’
*p-ruul	<i>rúul</i>	<i>gu:lʔ</i>	<i>pù-rùul</i>	—	—	<i>ma-run</i>	<i>bə-ríi</i>	‘snake’
*rik	<i>rit</i>	<i>gikʔ</i>	—	—	—	<i>rit-nu</i>	<i>Riit</i>	‘heavy’
*ril	<i>ril</i>	<i>gilʔ</i>	—	<i>ə-ri</i>	<i>l-rə́r</i>	—	<i>ʔə-ríi</i>	‘intestines’
*raʔ	<i>ràʔ</i>	<i>gàʔ</i>	—	—	—	<i>tak-ra</i>	<i>Raa</i>	‘fruit’
*ruy ~ *hruy	<i>hrüy</i>	<i>gui</i>	—	<i>rui</i>	<i>rí</i>	<i>rui</i>	<i>ruu</i>	‘vine; tendon’
		(Paite)						
*hriŋ	<i>hríŋ</i>	—	—		<i>l-ríŋ</i>	<i>riŋ</i>	<i>Riíŋ</i>	‘alive’

If all Northern Kuki-Chin languages are in the Peripheral group, and PKC *r- > g- in the common ancestor of all peripheral languages, the Old Kuki languages cannot be part of the Northern group. Given the morphological and lexical conservatism of Old Kuki languages, this introduces the possibility that they constitute a separate top-level branch of Kuki-Chin. However, they have been little-exploited in comparative reconstruction because relatively little data on them has been available. If it can be established that Sorbung belongs to this group, the body of data that we have assembled here could be of considerable value in reconstructing the history of Kuki-Chin as a family.

Aside from the evidence that has been presented thus far, there are three types of evidence that suggest Sorbung is a member of the Kuki-Chin family with specific affinities to the Old Kuki languages: shared lexical innovations, shared morphological innovations, and shared phonological innovations.

One striking aspect of the Sorbung lexicon is the presence of many lexical items that are widespread within Kuki-Chin but which are not widely found elsewhere in Tibeto-Burman:

Table 8: *Lexical resemblances between Sorbung and Kuki-Chin.*

SORBUNG	PKC		COMMENTS
<i>som</i>	* <i>som</i>	‘ten’	
<i>pan</i>	* <i>puan</i>	‘blanket; cloth’	
<i>bùu</i>	* <i>buʔ</i>	‘rice (cooked)’	
<i>kèe</i>	* <i>kee</i>	‘foot’	Cp. perhaps Karbi <i>kej</i> .
<i>lu-kaan</i>	* <i>luu</i>	‘head’	Cp. also Maring <i>lu</i> , Mru <i>lu</i> .
<i>lu-siip</i>	* <i>tship</i>	‘crown (of head)’	
<i>ʔèk</i>	* <i>ʔeek</i>	‘feces; shit’	
<i>mə-rái</i>	* <i>raay</i>	‘be pregnant’	
<i>cèm</i>	* <i>tsem</i>	‘knife’	
<i>ʔəták</i>	* <i>tak</i>	‘flesh’	
<i>wàm</i>	* <i>wom</i>	‘black’	
<i>dài</i>	* <i>daay</i>	‘cool; quiet’	
<i>paasàa</i>	* <i>pa-sal</i>	‘husband; man’	

However, Old Kuki languages (and Puiro) have even more specific lexical resemblances to Sorbung:

Table 9: *Innovative lexical items shared between Sorbung, Old Kuki (Kom, Aimol, Moyon, Chiru) and Puiro.*

SORBUNG	KOM	AIMOL	MOYON	CHIRU	UIRON	
<i>ʔóo</i>	<i>ei</i>	—	<i>ə</i>	—	<i>ai</i>	‘bite; chew’
<i>hóoŋ</i>	<i>ka-hoŋ</i>	<i>ə-joŋ</i>	—	<i>a-hoŋ</i>	<i>hoŋ</i>	‘come’
<i>ʔiin</i>	<i>in</i>	—	<i>lí</i>	—	<i>in</i>	‘drink’
<i>kʰòop</i>	—	<i>Khop</i>	—	—	—	‘be full’
<i>ʔən-cáaŋ-lám</i>	<i>čəŋ</i>	—	<i>ti-cáŋ</i>	—	<i>caŋ-lam</i>	‘right (side)’
<i>ʔən-fúŋ</i>	<i>in-suŋ</i>	<i>ə-suŋ</i>	—	<i>a-su</i>	—	‘sit’
<i>ʔiin</i>	—	—	—	—	<i>in</i>	‘sleep’
<i>lée</i>	—	—	—	—	<i>lo</i>	‘take’
<i>Kaa</i>	<i>kha-wa</i>	<i>Kha</i>	—	—	—	‘that’
<i>Wa</i>	<i>hi-wa</i>	—	—	—	—	‘this’
<i>ham-búu</i>	<i>hum-pui</i>	—	—	—	—	‘tiger’
<i>kòoŋ</i>	<i>koŋ</i>	—	<i>kuŋ</i>	—	<i>koŋ</i>	‘waist; back’

The lexical similarities between Sorbung and Old Kuki languages are very strong and indicate, at the very least, a history of contact between southern Tangkhuls and Kuki-Chin speaking peoples. The morphological similarities (in both inflection and derivation) help to reinforce the idea that the relationship is a genetic one. Hartmann (2001) shows that there are a pair of valence-related prefixes in Daai Chin (a Southern Kuki-Chin language): *ng-*, occurring in low-valence verbs, and *m-*, occurring in transitive (especially causative) verbs:

- (1) a. *thei* ‘be clever’ *ng-thei* ‘learn’ *m-thei* ‘teach’
 b. *püi* ‘be together’ *ng-püi* ‘be included’
 c. *kyüh* ‘be afraid’ *m-kyüh* ‘make afraid’

Cognates of these two prefixes appear in Sorbung as *ʔəN-* and *mə-*, respectively:

- (2) a. *ʔən-tʰèe* ‘awake; be awake’ *mə-tʰée* ‘get somebody up’
 b. *ʔən-cii* ‘fear; be afraid’ *mə-cíp* ‘frighten’

Various innovative inflectional markers found in Sorbung are also found in Old Kuki languages. The Sorbung progressive suffix *-om* (< *ʔóm* ‘exist; be a place’) has a cognate at least in Aimol. Compare Aimol *sekor a-chongaa-om* ‘horse 3-ride-PROG (he is riding a horse)’ with Sorbung *tʰiŋkooŋ-ua kiák-óm-òo* ‘tree-DEM fall-PROG-DECL (the tree is falling)’. The Sorbung genitive suffix *-ta* also appears to have cognates in Langrong, Purum, and Kolren (all of which have a formally similar suffix in genitive pronouns).

Since comparative reconstruction of languages in the India-Burma borderlands area remains in its infancy, relatively speaking, it is difficult to speak of phonological innovations in terms as explicit as would be desirable. As noted above, Sorbung shares with Old Kuki languages a common development of PTB **kr-*, **kl-*, **ts(y)-*, and **dzy-*. It was also noted that a similar change has occurred in Tangkhulic and that it is not entirely

possible to establish the complete independence of these changes. The same may be said for another cluster of changes that are shared by Kuki-Chin, Sorbung, and Tangkhulic:

Table 10: *Reflexes of Proto-Tibeto-Burman coronal fricatives and affricates in Proto-Kuki-Chin, Proto-Tangkhulic, and Sorbung.*

PTB	PTK	PKC	SORBUNG
*s-, *sy ₁ -	*t ^h -	*t ^h -	t ^h -
*sy ₂ -	*s-	*s ^h -	s-
*ts-	*s-	*s-, *θ-	s-
*tsy-	*ts-, *c-	*ts-, s-	c-, s-
*dz-	*ts-	*ts-	c-
*dzy-	*c-, *ts-	*ts-	c-

One piece of evidence that at least the Kuki-Chin and Sorbung innovations are shared is the fact that irregular and variable etyma tend to pattern the same way in PKC and Sorbung (given that PKC *ts- regularly corresponds to Sorbung c-):

Table 11: *Comparison of regular and variable developments from PTB *ts-, *ts(y)-, and *dzy-. Note that Sorbung patterns like PKC rather than PTK.*

PTB	PTK	PKC	SORBUNG	
*tsam	*sam	*sam	sàam	‘hair of head’
*ts(y)a-t	*tsa	*saa	saa	‘hot; ill’
*ts(y)a	*tsa	*θaa	saa	‘child’
*dzya-k/t/n	*tsa	*θa?	saa ‘eat’; sak ‘feed’	‘eat; food; feed’
*tsywap	*tsup	*tsuap	cúap ‘spleen’	‘lung’

It is also the case that Sorbung and Proto-Kuki-Chin developed initial glottal stop from the same—reasonably diverse—set of onsets:

Table 12: *Sources of Proto-Kuki-Chin and Sorbung initial glottal stop.*

PTB	PTk	PKC	Sorbung	
*ʔa:r	*ar	*ʔaar	ʔaa	chicken; fowl
*k ^w əy	*hwi	*ʔuy	ʔúu	dog (<i>Canis familiaris</i>)
*hyen	—	*ʔen	ʔen	look
*k-yim	*ʃim	*ʔim	ʔin	house

Perhaps the most striking evidence of a close relationship between Old Kuki and Sorbung is that fact that at least one Old Kuki language, Kom, shares with Sorbung a productive voicing alternation that seems to be active in the same environments. This alternation is discussed at length in Section 0 below.

Based on the evidence that we have presented here, we argue that Sorbung is best seen as a Kuki-Chin language (closely related to Old Kuki languages like Kom, Aimol, and Moyon) that has come under influence from Tangkhulic languages. We further suggest that it is not the only “Nagafied” Kuki language and that Puiron might be another case of this type, a matter that cannot be discussed at length here.

1.4. Important Characteristics

The significance of Sorbung partly lies in that it demonstrates of changes in progress that are important for accounting for historical developments elsewhere in Tibeto-Burman. These changes orbit around minor syllables (“presyllables” or “prefixes”), both their development from etymological major syllables and their collapse into following major syllables, yielding a new stop series.

Sorbung shows a pervasive pattern of reduction in pretonic syllables. Compounds consisting of two noun roots that would consist of heavy syllables in isolation are realized as a sequence of a light syllable and heavy syllable. This stress-conditioned pattern of reduction creates a kind of structure that is intermediate between full disyllables and sesquisyllables and provides a likely pathway for the development of new minor syllable prefixes from compounded roots.

In true minor syllables, though, another process is in progress. The vocalism in such syllables is now completely predictable; thus, the /ə/ vowels in minor syllables are analyzed here as epenthetic. However, in certain environments, epenthesis variably fails to occur. Thus, instead of [ʔəkái] ‘hips’, our consultant often produced [ʔkái]. Such preglottalized productions could easily develop into fortis [kʰái] or geminate stops [k:ái], which could be reanalyzed by learners as aspirates [kʰái]. Since the prefix [ʔə- ~ ʔ-] has an unpredictable distribution, such a change would lead to what would appear to be, retrospectively, an unconditioned change in manner. Matisoff (2003:87ff) has suggested that many of the apparently sporadic manner correspondences among Tibeto-Burman onsets should be attributed to the effects of lexical prefixes that have collapsed into the following major syllable onset. Because the distribution of Tibeto-Burman lexical prefixes seems to be heavily determined by analogical processes, but seldom shows deterministic grammatical conditioning, the presence of a prefix often has to be inferred from the effect its existence is meant to explain. While this presents a methodological problem (invoking a hidden cause), cases like Sorbung give additional empirical support to the idea that semi-productive lexical prefixes can give rise to new laryngeal distinctions in onsets.

2. Syllable Structure and Prosody

Sorbung, like many other Tibeto-Burman languages (indeed, many Southeast Asian languages generally), tends to have sesquisyllabic word stems. That is, word stems often consist of a minor-syllable “prefix” (which is sometimes a morphological prefix) and a major syllable “root”. This pattern is consistent with a general tendency towards right-headed (iambic) prosodic constituents. It is also essential to understanding the distribution to tone (which is only contrastive on major syllables) and the general syllable-structure constraints in the language.

2.1. Stress

Stress in Sorbung has the following correlates:

1. **Duration:** Unstressed vowels are never long; underlying long vowels are shorted in unstressed syllables.
2. **Amplitude:** Stressed syllables are louder than unstressed syllables.
3. **Pitch:** Pitch excursions are more pronounced in stressed syllables.

All of these correlates point to a general iambic pattern of stress. This is not to say, however, that Sorbung stress is predictable without reference to morphological structure. The following general principles seem to hold:

1. Minor syllables are never stressed. This may be due to the fact that the vowels in these syllables are epenthetic.
2. The last root syllable in a compound stem, not counting derivational suffixes, is stressed.
3. Two consecutive unstressed syllables are not tolerated.
4. Stress assignment is cyclic. ‘Underarm hair,’ from the compound *cub'laa* (*cuup* ‘breast’ plus *laa* ‘crotch’) plus *míi* ‘hair’ is *cub,laa'míi* rather than **,cuubla'míi*.
5. Suffixes do not affect the stress of the base to which they are attached: ‘look’ ‘big’, *'loog-,oo* ‘big-DECL’, *ʔ'uʉ* ‘dog’, *ʔ'uʉ-e* ‘dog-NOM’.
6. The members of one set of suffixes are always stressed (including the obligatory final suffixes marking mood). Another set of suffixes are always unstressed (including case suffixes).
7. Some non-lexicalized compounds, particularly verbs compounded with nominal objects, may not display stress subordination. Thus, while *móo* ‘fire’ is unstressed and reduced in *mò-sèm* ‘fire-blow; blow on fire’, *cuup* ‘breast’ is fully stressed and not reduced in *cuup-ʔiin-èe* ‘breast-drink-IMP; suckle!’

The following data illustrate these principles more fully. It can be seen in (3) that /*ŋaa*/ ‘fish’ is stressed when final or when preceded by a minor syllable, but is unstressed when it is initial:

- | | | | | | |
|-----|----|--------------------|-------------------------------|--|---------------|
| (3) | a. | <i>'ŋaa</i> | ‘fish’ | | ‘fish’ |
| | b. | <i>ʔə-'ŋaa</i> | ‘0’ + ‘fish’ | | ‘fish’ |
| | c. | <i>ʔə-ŋaa-'míi</i> | ‘0’ + ‘fish’ + ‘hair/feather’ | | ‘scale; fin’ |
| | d. | <i>ŋa-'míi</i> | ‘fish’ + ‘hair/feather’ | | ‘flying fish’ |

The examples in (4-6) show that /*kee*/ ‘foot’ is unstressed when in it occurs penultimately (5) but is stressed before a minor syllable (6):

- | | | | | |
|-----|-------------|--------------------------|----------------------------|--------------|
| (4) | <i>'kèe</i> | ‘foot’ | | ‘foot’ |
| (5) | a. | <i>kè-'lúuŋ</i> | ‘foot’ + ‘print’ | ‘footprint’ |
| | b. | <i>kè-'kíuk</i> | ‘foot’ + ‘crippled’ | ‘lame’ |
| | c. | <i>kè-'búu</i> | ‘foot’ + ‘great one’ | ‘big toe’ |
| | d. | <i>kè-'sáa</i> | ‘foot’ + ‘child’ | ‘little toe’ |
| (6) | a. | <i>,kèe-mə'tóo</i> | ‘foot’ + ‘thigh’ | ‘crotch’ |
| | b. | <i>,kèe-mə'rái</i> | ‘foot’ + ‘calf’ | ‘calf’ |
| | c. | <i>,kèe-mə'jáa</i> | ‘foot’ + ‘palm/sole’ | ‘sole’ |
| | d. | <i>,keè-mə,júum-'ràa</i> | ‘foot’ + ‘digit’ + ‘fruit’ | ‘toe’ |

In each word, one stress is more prominent than the others (specifically, it has a greater amplitude and longer duration). This is indicated here as primary stress [']. Other stressed syllables are marked as bearing secondary stress [,]. By way of illustration, duration of ‘foot’ in (5a-d) is much shorter than the duration of ‘foot’ in (6a-d). This, in turn, is somewhat shorter than the duration of ‘foot’ in (4). We have not identified any

cases where the distinction between primary and secondary stress is contrastive. However, its patterning does suggest that both Sorbung feet and prosodic words are right-headed.

2.2. Word Structure

Barring loanwords, non-compound stems in Sorbung are always monosyllabic or sesquisyllabic. Monosyllabic stems consist of a single major syllable. Sesquisyllabic stems consist of a minor syllable followed by a major syllable. Underlyingly, we believe that all Sorbung roots have a single vowel and that the predictable schwa-vocalism in minor syllables is the result of epenthesis. However, the system is easier to describe if it is approached from the surface.

2.3. Minor syllables

With one exception, the minor syllables in Sorbung have the shape Cə, where C can be any of the onsets allowed in major syllables except for /p/, /t/, /w/, /j/, and /ʃ/. The exception to this pattern is /ʔəN/, where N is a nasal that shares the same place of articulation as the following consonant. In all instances of minor syllables, [ə] is the only vowel to appear (though it is possible that the quality of these vowel is affected by coarticulation with following vowels). The complete inventory is given in the following table:

Table 13: *Minor syllables in Sorbung.*

bə				
<i>bərii</i> ‘snake’				
<i>bəkii</i> ‘parrot’				
pə	tə	cə	kə	ʔə
<i>pəláii</i> ‘navel’	<i>tətír</i> ‘crossbar’	<i>cəháa</i>	<i>kəʃuá</i> ‘rain’	<i>ʔəwám</i> ‘bear’
	<i>təràt</i> ‘tie’	‘langur’	<i>kələéɛŋ</i> ‘loose’	<i>ʔətʰár</i> ‘new’
		<i>cəkòo</i> ‘river’		
pʰə	tʰə		kʰə	
<i>pʰənaat</i> ‘grind’	<i>tʰəwáaj</i> ‘fly (n.)’		<i>kʰəràan</i>	
<i>pʰəláa</i> ‘wing’	<i>tʰərii</i> ‘boundary’		‘shoulder’	
mə	nə		ŋə	
<i>mətír</i> ‘shrew’	<i>nəmooməkʰúu</i>		<i>ŋənáar</i> ‘snore’	
<i>məmìt</i>	‘widow’			
‘extinguish’				
	sə			
	<i>səloo</i> ‘buffalo’			
	<i>səkúu</i> ‘porcupine’			
	rə			
	<i>rəŋáa</i> ‘five’			
	lə			
	<i>ləkʰàa</i> ‘shoulder’			
	<i>ləʃuá</i> ‘put’			
	təN			ʔəN
	<i>təŋguáp</i> ‘headdress’			<i>ʔəntʰée</i> ‘awake’
	<i>təmbuelàm</i> ‘river			<i>ʔəŋkʰaaw</i> ‘fish
	vally’			sp.’

In actual production, each of the Cə minor syllables has a variant without a schwa. Thus, /s-mùk/ ‘cattle’ can be realized either as [səmùk] or [smùk] and /ʔ-k^haaw/ ‘grasshopper’ can be realized either as [ʔək^haaw] or [ʔk^haaw]. The variants with initial consonant clusters occur more frequently when there is a sharp increase in sonority between the two consonants.

Problematically, our database contains minor syllables with /b/ but not with any other voiced stop (in spite of the fact that /d/ appears as the onset of major syllables). We believe this to be an accidental gap. Also, aspirated-stop minor syllables have a restricted distribution, occurring only before sonorant onsets. However, in this environment they contrast with their unaspirated counterparts, at least on the surface:

- (7) a. *k^həràan* ‘spider’ *kəraak* ‘voice’
 b. *p^həláa* ‘wash’ *pəlái* ‘snake’
 b. *t^hərii* ‘boundary’ *təràt* ‘tie’

We propose that this pattern actually results from a morphological cause: the aspirated stop minor syllables occur only in monomorphemic stems and are the result of epenthesis into underlying stop-sonorant clusters. The minor syllables with unaspirated stops occur in both monomorphemic stems, as in (7) and bimorphemic stems where the minor syllable is a lexical prefix. Underlying, we posit that monomorphemic stems like those in (7) have forms as in (8) while bimorphic stems have forms like those in (9):

- (8) a. /k^hràan/ ‘spider’ /kraak/ ‘voice’
 b. /p^hláa/ ‘wash’ /plái/ ‘snake’
 b. /t^hrii/ ‘boundary’ /tràt/ ‘tie’
 (9) a. /m-mìt/ ‘extinguish’ /t-tír/ ‘bar’
 b. /r-ŋáa/ ‘five’ /l-fùa/ ‘put’
 b. /s-loo/ ‘buffalo’ /ŋ-náar/ ‘tie’

This accounts for the freer distribution of unaspirated-stop minor syllables versus aspirated-stop minor syllables and for the greater frequencies of schwa-less variants stems in the aspirated-stop minor syllable group.

In some cases, there is compelling evidence that minor syllables are distinct morphological units; in other cases, the data suggest that minor syllables are simply part of a root. Evidence for morphological independence includes the following:

1. A root may occur with or without a minor syllable, depending on morphological or syntactic context.
2. The same root may occur with different minor syllables in related lexical items.
3. A cluster of stems containing the same minor syllable share a component or meaning or belong to the same semantic field.
4. The minor syllable is transparently related to an independent lexical item.

According to these criteria, at least the following minor syllables should be considered to be prefixes, rather than parts of the root:

- (10) **bə-** **fossilized animal prefix**
 a. *bərii* 'snake'
 b. *bakii* 'parrot'
- (11) **cə-** **'water' prefix** (*cii* 'water')
 a. *cəkáaŋ* 'drought'
 b. *cəlòk* 'flood'
 c. *cəluáŋ* 'flow'
- (12) **kə-** **lexical prefix**
 a. *kəniin* 'two' (*semnii* 'twenty')
 b. *kəriúuk* 'six'
- (13) **ʔə-** **lexical prefix**
 a. *ʔəriúu* 'bone'
 b. *ʔəcii* 'egg' (*ʔəcii-məcii* 'lay egg')
 c. *ʔəŋaa* 'fish' (*ŋámii* 'flying fish')
 d. *ʔəmii* 'hair; fur' (*mòrmii* 'facial hair')
- (14) **mə-** **body-part prefix** (< PTB **mi* 'human?')
 a. *məsàa* 'body'
 b. *mətóo* 'lap; thigh'
 c. *mətín* 'nail; claw'
- (15) **mə-** **'mouth' prefix** (< *mor* 'mouth?')
 a. *məlit* 'vomit'
 b. *məjúup* 'kiss'
 c. *məcámci* 'saliva'
 d. *mək'áa* 'jaw; chin'
- (16) **mə-** **fossilized stative prefix**
 a. *mənée* 'be soft'
 b. *mək'ái* 'be crooked'
 c. *məsáaŋ* 'be high'
 d. *mənèem* 'be low'
- (17) **mə-** **productive causative prefix** (< **mii* 'give?')
 a. *məthée* 'wake up (v.t.)' (*ʔənthée* 'wake up (v.i.)')
 b. *məcàap* 'make cry' (*càap* 'hurt')
 c. *məʔóo* 'make bite' (*ʔóo* 'bite')
- (18) **ʔəN-** **valence-decreasing prefix**
 a. *ʔənt'hée* 'wake up (v.i.)'
 b. *ʔəncii* 'fear (v.i.)'
 c. *ʔəndía* 'fall (v.i.)'

- (19) **lə-** ‘**hand/arm**’ prefix (< PTB *l(y)ak ‘hand/arm’?)
- a. *lək^hàa* ‘shoulder’
 - b. *ləfùa* ‘put’
 - c. *ləkium* ‘bar’
- (20) **sə-** ‘**animal**’ prefix (< *saa* ‘animal’)
- a. *səloo* ‘buffalo’
 - b. *səriám* ‘mithun’
 - c. *səki* ‘deer’
 - d. *səkúu* ‘porcupine’

None of the aspirated stop-minor syllables belong in this set, consistent with the hypothesis that the aspirated stops are morphologically part of the root.

Minor syllables are subject to an alternation that does not appear in major syllables. In compounds, nasal codas do not assimilate in place to following obstruents:

- (21) a. [inpù] ‘host (house owner)’
- b. [puank^hóoŋnaa] ‘loom’
 - c. [somt^húum] ‘thirty’
 - d. [ciŋtòŋ] ‘hill’

The nasal of the minor syllables [ʔəN] and [təN] assimilates in place to the following segment. The following examples show this for [ʔəN], which appears as a possessive prefix on kinship terms and as a stative/valency-reducing prefix on verbs:

- (22) a. [ʔəmbíi] ‘paternal grandfather (of ego)’
- b. [ʔəmbúu] ‘maternal grandfather (of ego)’
 - c. [ʔənt^hée] ‘awake’
 - d. [ʔəntiæk] ‘green’
 - e. [ʔəndía] ‘fall’
 - f. [ʔənnuu] ‘mother (of ego)’
 - g. [ʔəŋguu] ‘child’
 - h. [ʔəŋk^hau] ‘fish species’

Similarly, the nasal in [təN], for which no general function is discernable, shares the place of articulation with the subsequent onset.

- (23) a. [təmbuelàm] ‘left hand’
- b. [təmbàak] ‘valley’²⁸
 - c. [təŋʃuk] ‘pestle’
 - d. [təŋguáp] ‘headdress’
 - e. [təŋgoon] ‘ant-eater’
 - f. [təŋk^hai] ‘half’

²⁸ This form is probably borrowed from Meithei (see Chelliah 1997).

Before palatals, as in (23c) ‘pestle’, the nasal becomes coronal. This may be due to a general restriction against palatal nasals in Sorbung. Another interpretation of the evidence is that nasal is underlyingly /n/. However, we suggest that the nasal is underlyingly unspecified for place, for the following reasons:

1. Throughout the corpus, nasal codas generally do not assimilate in place to following obstruents; the different, assimilating behavior of [ʔəN] and [təN] can be explained by the “need” for the final nasals to acquire place features.
2. No where in the data does one of these formatives not occur before an obstruent from which it could acquire place features.

This position is subject to verification, pending the availability of further data.

2.4. Major syllable types

Major syllables in Sorbung may be either open or closed and may have either long or short nuclei. Length is only contrastive in closed syllables and only when syllables are stressed. Unstressed syllables always surface with short nuclei and underlying long vowels are shortened when they occur in weak positions. Stressed syllables, on the other hand, are always closed, have long nuclei, or both.

Table 14: *Sorbung major syllable types.*

	Stressed Syllables	Unstressed Syllables
Open Syllables	CVV	CV
Closed Syllables	CVC, CVVC	CVC

Long nuclei may be either long monophthongs or diphthongs. There appears to be no quantitative (length) contrast among diphthongs. As discussed in section 0 below, monophthongs and rising-sonority diphthongs occur in closed syllables relatively freely. However, (phonological) falling sonority diphthongs do not occur in closed syllables. Interactions between constraints on stress assignment and quantity produce length alternations, which will be discussed at greater length in section 0 below.

2.5. Tone

Tone is contrastive in Sorbung, but the lexical load born by tonal contrasts is relatively low and the phonetic distance between tonal categories is relatively small. Tone interacts with intonation to such an extent that it was sometimes impossible, working from recordings, to determine the tones of a lexical item with complete certainty. Furthermore, all statements about the tonal inventory of Sorbung remain tentative until the tone sandhi system of Sorbung is worked out in greater detail.

2.5.1. Inventory

Sorbung has three contrasting tones, which we will refer to as H, L, and M. These are indicated in our transcriptions by diacritics on the first vowel of a rhyme: acute (acute), grave (grave), or no diacritic (x), for H, L, and M, respectively. Minimal and near-minimal sets illustrating these contrasts are given in **Table 15** below:

Table 15: *Example of the three contrastive tones of Sorbung.*

	H		L		M
<i>sáa</i>	‘hot’	<i>sàa</i>	‘animal’	<i>saa</i>	‘eat’, ‘child’
<i>móo</i>	‘fire’	<i>mòo</i>	‘QUEST’	<i>moo</i>	‘brother’
<i>t^hóo</i>	‘hear’	<i>k^hòo</i>	‘split, break’	<i>k^hoo</i>	‘bee’
<i>fī</i>	‘rain cover’	<i>fī</i>	‘sand’	<i>fī</i>	‘crab’, ‘star’
<i>wáj</i>	‘fly (v.)’	<i>wám</i>	‘black’	<i>wan</i>	‘belly’
<i>rée</i>	‘war’	<i>nèe</i>	‘small’	<i>ree</i>	‘roast’
<i>náap</i>	‘stick (v.)’	<i>nàap</i>	‘liquid mucous’		
<i>nàk^húak</i>	‘deaf person’	<i>nàk^hùak</i>	‘be deaf’		

The labels given here should not be taken as literal phonetic realities: H is not simply a high-pitched tone and L is not simply a low-pitched tone. The tones have characteristic pitch contours and also differ from one another in voice quality. The H tone rises to a peak about two-thirds through its duration before dropping slightly. It often has a tense voice quality. The L tone falls through its duration and sometimes starts higher in pitch than the terminus of a preceding L or M tone. L tone syllables are usually produced with lax or breathy voice quality. The pitch of M tone syllables falls gently before rising again. The voice quality is usually modal.

The contrastive load of tone in Sorbung is relatively low. This may be responsible for some of the difficulty we encountered in producing a complete analysis of the tone system. Our consultant was aware that there were tonal contrasts in Sorbung, and produced a number of minimal sets with little prompting (e.g. ‘rain cover’, ‘sand’, ‘crab’, ‘star’), but found it difficult to say whether two items were similar or different in tone.

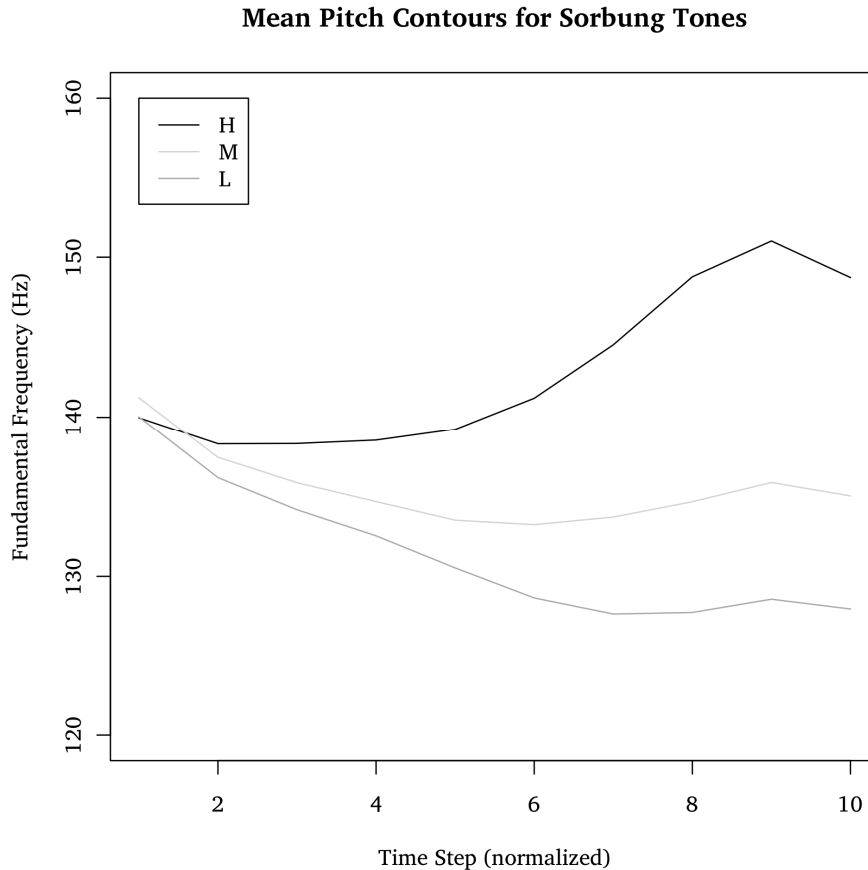


Figure 1: *Averaged pitch plots for the three Sorbung tones.*

2.5.2. Distribution

Tonal contrasts in Sorbung are subject to two major restrictions. First, tonal contrasts only occur on major syllables. The pitch of minor syllables is predictable from the pitch of surrounding major syllables. Second, there is only a two-way tonal contrast in stopped syllables (syllables with a stop coda) but a three-way contrast in major syllables of other types. Neither of these restrictions is unusual in the area.

2.5.3. Alternations

Given the current state of knowledge about Sorbung, it is clear that the language has a robust system of tonal alternations. However, the data are not sufficient to paint a clear and complete picture of these alternations. What follows, then, is only an introduction to some of the most obvious alternations.

Sorbung displays dissimilation of high tones. In smooth syllables, underlying H become M after H:

- | | | | | |
|------|----|--------------------|---------------------|--------------------|
| (24) | a. | <i>cíuu</i> | ‘liquid’ | ‘water; liquid’ |
| | b. | <i>kʰò-cíuu</i> | ‘bee’ + ‘liquid’ | ‘honey’ |
| | c. | <i>məcam-cíuu</i> | ‘saliva’ + ‘liquid’ | ‘saliva’ |
| | d. | <i>məkʰáan-cuu</i> | ‘sweat’ + ‘liquid’ | ‘sweat’ |
| | e. | <i>məkʰá-cuu</i> | ‘tear’ + ‘liquid’ | ‘tear’ |
| (25) | a. | <i>ɽə-ríuu</i> | PREFIX + ‘bone’ | ‘bone’ |
| | b. | <i>mai-ríuu</i> | ‘face’ + ‘bone’ | ‘cheekbone’ |
| | c. | <i>kòŋ-ríuu</i> | ‘back’ + ‘bone’ | ‘spine; backbone’ |
| | d. | <i>ɽəká-i-ruu</i> | ‘buttocks’ + ‘bone’ | ‘tailbone; coccyx’ |
| | e. | <i>mətóo-ruu</i> | ‘thigh’ + ‘bone’ | ‘thighbone; femur’ |
| | f. | <i>məkʰáa-ruu</i> | ‘jaw’ + ‘bone’ | ‘jawbone’ |

As can be seen from (24a-c) and (25a-c), the morphemes meaning ‘water; liquid’ and ‘bone’ are realized with a H tone word-initially and when preceded by M or L. However, as shown by (24d-e) and (25d-f), they are realized with a M tone when preceded by a H tone.

This process interacts with another alternation. Stopped syllables with an underlying H tone are realized as L except word-finally. Thus, /kút/ ‘hand; arm’ is realized as [-kút] when it occurs finally but as [-kút-] when it appears elsewhere:

- | | | | | |
|------|----|------------------|----------------------|-------------|
| (26) | a. | <i>kút</i> | ‘hand’ | ‘hand, arm’ |
| | b. | <i>kút-sáa</i> | ‘hand’ + ‘child’ | ‘finger’ |
| | c. | <i>kúd-bíuu</i> | ‘hand’ + AUG | ‘thumb’ |
| | d. | <i>kúd-məjáa</i> | ‘hand’ + ‘palm/sole’ | ‘palm’ |
| (27) | a. | <i>cíup</i> | ‘breast’ | ‘breast’ |
| | b. | <i>cúb-ráa</i> | ‘breast’ + ‘fruit’ | ‘nipple’ |
| | c. | <i>cúb-cuu</i> | ‘breast’ + ‘liquid’ | ‘milk’ |

In isolation, the morphemes meaning ‘hand; arm’ and ‘breast’ are realized with H tones. However, when followed by other syllables, within a word, the tone changes to L. As shown by example (27c), this process counter-bleeds the H → M / H___ process described above. Were this not the case, we would expect *[cúb-cúu], with a H tone on the second syllable.

Our Sorbung database is not currently large enough, and does not have enough morphologically complex forms, to allow us to confidently characterize all of the tonal alternations that we have observed. These two patterns are described here because they were exemplified in a relatively large number of data. Based upon these relatively limited findings, though, we believe it is safe to characterize the tone sandhi system of Sorbung. On the one hand, Sorbung clearly has more tonal alternations than Tangkhulic languages like Ukhrul, where tone is morphophonologically very stable. On the other hand, Sorbung cannot claim as pervasive a system of tone rules as many Kuki-Chin languages (Hyman & VanBik 2002; Hyman & VanBik 2004; Hyman 2007a; Hyman 2007b; Yip 2004).

3. Onsets

3.1. Inventory

The onset inventory of Sorbung includes plosives, fricatives, nasals, trills, and approximates. Sorbung has five contrasting places of articulation: labial, coronal, palatal, velar, and glottal. Plosives appear at all of these places and display a three-way distinction among voiceless, aspirated, and voiced. In word-initial position, the /b/-/p/ and /d/-/t/ contrasts are retained but the /g/-/k/ contrasts is neutralized. Neither voicing nor aspiration is contrastive at the palatal place. There is no voiced-voiceless contrast among fricatives and the only voiced fricative is [v], which is a conditioned variant of /w/. There are labial, coronal, and velar nasals but no palatal nasal. The inventory also includes the trill /r/ ([r] ~ [r̄]) and the approximates /l/, /w/, and /y/. The inventory is summarized in **Table 16** below.

Table 16: *Sorbung onsets. Onsets in parenthesis only occur as predictable variants of other underlying sounds.*

	LABIAL	CORONAL	PALATAL	VELAR	GLOTTAL
NASAL	m	n		ŋ	ʔ
PLOSIVE	p	t	c	k	ʔ
	p ^h	t ^h		k ^h	
	b	d		g	
FRICATIVE	(v)	s	ʃ		h
TRILL		r			
APPROXIMATES		l	j	w	

The language variably allows a few consonant clusters, comprising of /pl/ /p^hl/ /tr/ /k^hr/ and /kl/, (see section 0 above).

3.1.1. Examples

Table 17 provides examples of each of the onsets.

Table 17: *Examples of each onset.*

/m/	<i>mai</i>	‘face’	<i>mùit</i>	‘eye’	<i>mùu</i>	‘see’
/n/	<i>náa</i>	‘hurt’	<i>niiŋ</i>	‘mind’	<i>nèe</i>	‘small’
/ŋ/	<i>ŋaa</i>	‘fish’	<i>ŋia</i>	‘stand’	<i>ŋée</i>	‘white’
/p/	<i>paasàa</i>	‘man’	<i>pìi</i>	‘aunt’	<i>pée</i>	‘give’
/t/	<i>taa</i>	‘old’	<i>tətir</i>	‘crossbar’	<i>tee</i>	‘sister’
/c/	<i>càap</i>	‘cry’	<i>məcìi</i>	‘salt’	<i>cèm</i>	‘knife’
/k/	<i>kaa</i>	‘that’	<i>ʔəkíi</i>	‘corner’	<i>kèe</i>	‘leg’
/g/	<i>ʔəŋgúu</i>	‘child’	<i>cəgòo</i>	‘river’	<i>təgók</i>	‘pot’
/ʔ/	<i>ʔaa</i>	‘fowl’	<i>ʔiin</i>	‘sleep’	<i>ʔèk</i>	‘feces’
/p ^h /	<i>p^háaglàaŋ</i>	‘wall’	<i>p^hít</i>	‘hit’	<i>ʔəp^hút</i>	‘breath’
/t ^h /	<i>t^haaŋ</i>	‘dry’	<i>t^hii</i>	‘blood’	<i>t^hèe</i>	‘bow’
/k ^h /	<i>k^haa</i>	‘bitter’	<i>k^hii</i>	‘wind’	<i>k^hoo</i>	‘bee’
/b/	<i>ʔəbák</i>	‘bat’	<i>buum</i>	‘sit on eggs’	<i>ʔəbúu</i>	‘mole’
/d/	<i>dài</i>	‘cold’	<i>dùu</i>	‘sew’	<i>mədəe</i>	‘be born’
/s/	<i>sàa</i>	‘animal’	<i>lusìp</i>	‘crown’	<i>ʔəsóo</i>	‘spear’

/ʃ/	<i>fúuk</i>	‘exit’	<i>fír</i>	‘vagina’	<i>fóo</i>	‘long’
/h/	<i>haa</i>	‘tooth’	<i>maihiŋ</i>	‘freckle’	<i>máhéé</i>	‘charcoal’
/r/	<i>ʔaraa</i>	‘fruit’	<i>ríiŋ</i>	‘alive’	<i>rée</i>	‘war’
/l/	<i>laa</i>	‘song’	<i>málii</i>	‘four’	<i>lee</i>	‘medicine’
/j/	<i>jàa</i>	‘accept’	<i>ʔaji</i>	‘drive, chase’	<i>ʔajòok</i>	‘friend’
/w/	<i>ʔawaa</i>	‘bird’	<i>t^həwái</i>	‘mosquito’	<i>wór</i>	‘swell’
	<i>laavá</i>	‘this song’	<i>iiŋk^haavá</i>	‘this boundary’	<i>bélcàavá</i>	‘this spade’

Two of the onsets have visibly limited distributions.

- /ʃ/- never occurs before /ee/, /e/, or /aa/.
- /w/- only occurs before /a/ and /aa/.

While it would seem desirable to collapse these onsets with other phonemes, they are both in contrastive distribution with all of the possible candidates.

A more interesting case is the relationship between /g/ and /k/. Both of these occur in the onsets of major syllables following vowel-final minor syllables, establishing them as distinct phonemes:

- (28) a. *təgók* ‘pot’
 b. *cəgòo* ‘river’
- (29) a. *səkii* ‘deer’
 b. *məkée* ‘kidney’
 c. *kap k^hət* ‘length from thumb to finger’
 d. *cəkáap* ‘tongs’
 e. *səkúan* ‘horse’
 f. *ʔəkór* ‘peel’

However, while /k/ can occur word-initially (30), and word-internally after another major syllable (31), /g/ does not:

- (30) a. *kít* ‘arm’
 b. *kèe* ‘foot’
 c. *kuum* ‘year’
 d. *koo* ‘do’
- (31) a. *pakíuu* ‘younger paternal uncle’
 b. *kèkíuk* ‘crippled, lame, handicapped’

In this respect, /g/ is unlike the other voiced stops, /b/ and /d/. However, like them (and unlike the voiceless stops) /g/ can occur after nasal-final minor syllables:

- (32) a. *ʔəŋguu* ‘baby, child’
 b. *təŋguáp* ‘hat’
 c. *təŋgoon* ‘anteater’

A reviewer suggests that this might be a case of free [g ~ k] variation, as is found Jingpho. This is possible. However, within our corpus there are no clear instances of [g] varying with [k] in the same lexical item without a clear conditioning environment (see 0 below).

3.2. Alternations

As discussed above, CəC sequences vary with CC sequences in the same lexical item. That is to say, the epenthetic process that breaks up underlying initial consonant clusters to produce minor syllables is variable. The [ə] appears to be the only vowel permitted in the minor syllables. We note three types of cases where epenthesis sometimes fails to occur, depending crucially on the nature of the two initial consonants. First, in /s/-sonorant clusters, epenthesis variably fails yielding, e.g. [sə.m...] varying with [sm...], as in ‘cattle’:

(33) [səmùk] ~ [smùk] ‘cattle’

In this example, as in other cases with /s/+sonorant clusters, the /sm/ cluster is tautomorphemic (/s+muk/ ‘animal prefix’ + ‘cattle’). This variation is apparently sensitive to word length and speech rate. For example, /smùk/ ‘cattle’ was usually realized *with* epenthesis, but /smukpaasaa/ ‘bull’ was usually realized *without* epenthesis.

Most of the obstruent-sonorant onset clusters in Sorbung contain stops. These clusters consist of voiceless labial or velar stops followed by /l/ or /ɾ/:

(34) b. [pəlaaj] ~ [plaaj] ‘umbilical cord’
 c. [p^heláa] ~ [p^hláa] ‘wing (of a bird)’
 d. [k^həràan] ~ [k^hràan] ‘spider’

In none of these cases is there convincing evidence that the minor syllable is an independent morphological unit. It is notable that while we have recorded instances of clusters with /p/ and /p^h/, we have noted no clusters with /b/ even though the prefix /b-/ occurs before the root /ríi/. This leads us to believe that epenthesis between morphemes is more favored than epenthesis into morphemes.

The final, and probably most theoretically interesting, cluster of environments in which epenthesis sometimes fails to occur is between a glottal stop and a following consonant:

(35) a. [ʔək^hau] ~ [ʔk^hau] ‘grasshopper’
 b. [ʔək^húan] ~ [ʔk^húan] ‘voice’
 c. [ʔəkáí] ~ [ʔkáí] ‘hips’
 d. [ʔək^húk] ~ [ʔk^húk] ‘knee’
 e. [ʔərík] ~ [ʔrík] ‘louse’
 f. [ʔəpáat] ~ [ʔpáat] ‘soft, mushy’
 g. [ʔəŋaa] ~ [ʔŋaa] ‘fish’

As with the /s/+sonorant clusters, the /ʔ/+consonant clusters are all tautomorphemic. There are a number of different prefixes in Sorbung which are underlyingly /ʔ-/, usually with a following epenthetic schwa. However, all of them seem to pattern similarly with regard to

epenthesis. When there is no epenthetic schwa, the glottal stop is realized as a faint glottal release (before stops), increased duration, or preglottalization (on a following sonorant). The variation does not seem to be constrained by the manner of the second consonant and seems to be a function mostly of speech rate and higher-level prosodic influences.

There are also cases, in Sorbung, where the first mora of a diphthong is devocalized and syllabified as an onset. The best examples of this involve the proximal demonstrative enclitic, /ùá/. When the syllable to which /ùá/ cliticizes ends in a consonant, the clitic is realized as [ùá]:

- (36) a. [cèmùá] ‘this knife’
 b. [ʔəʃimùá] ‘this needle’
 c. [təgógùá] ‘this pot’
 d. [təŋguábuá] ‘this headdress’
 e. [təŋguàdùá] ‘this bowl’

If the syllable to which it cliticizes ends in a vowel other than /a/, /ùá/ is usually realized as [wá]:

- (37) a. [tʰèewá] ‘this bow’
 b. [lèewá] ‘this medicine’
 c. [réewá] ‘this war’
 d. [ròowá] ‘this axe’
 e. [ʔəsóowá] ‘this spear’

After high vowels, though, /ùá/ varies between [wá] and [vá]:

- (38) a. [məʃiíwá ~ məʃiívá] ‘this necklace’
 b. [ræuwá ~ ræuvá] ‘this rope’
 c. [meciiwá ~ meciivá] ‘this necklace’

Finally, and most curiously, after /a/, /ùá/ varies between [vá] and [βá]:

- (39) a. [bélcàavá] ‘this spade’
 b. [iink^haavá ~ iink^haaβá] ‘this door’
 c. [laavá] ‘this song’

It is not clear to us what the phonetic basis for this latter pattern is.²⁹

²⁹ A reviewer inquires about the fate of the initial low tone in /ùá/. When the /u/ is realized as a vowel, rather than a glide, it is produced with a distinct low tone and the following /a/ is produced with a high tone. When the /u/ is devocalized to [w], [v], etc., the low tone disappears so that in [məʃiíwá], the last two syllables are realized as high tones with no intervening low. This suggests that the low tone may be a default inserted on toneless TBUs.

4. Rhymes

4.1. Inventory

We analyze Sorbung as having five short-long pairs of vowel phonemes in major syllables:

/ii/	/i/	/uu/	/u/	/uu/	/u/
/ee/	/e/			/oo/	/o/
		/aa/	/a/		

Phonetically, the short-long distinction is realized through a number of different cues other than duration, including quality and dynamic quality. In general, long vowels have a more peripheral quality and short vowels a more centralized quality. For the high vowels, duration is the primary cue to length. The long mid-vowels /ee/ and /oo/ are phonetically the diphthongs [ej] and [ow] while their short counterparts are realized as the lower monophthongs [ɛ] and [ɔ]. The low vowels /aa/ and /a/ are phonetically [a:] and [ɐ]. A summary of the phonetic values of the vowel phonemes is given below:

/ii/ [i:]	/i/ [ɪ]	/uu/ [u:]	/u/ [ʊ]	/uu/ [u:]	/u/ [ɔ]
/ee/ [ej]	/e/ [ɛ]			/oo/ [ow]	/o/ [ɔ]
		/aa/ [a:]	/a/ [ɐ]		

We readily acknowledge that it would be possible to reanalyze this system in terms other than quantity (vowel length). The similar vowel systems of Northern Chin languages, for example, were analyzed by Button (2009) in purely qualitative terms. Our analysis provides a number of advantages. It allows us to easily characterize the stress-conditioned vowel alternations as alternations in a single phonological parameter (length), even though different phonetic parameters are involved depending on the vowel. By treating /ee/ and /oo/ as long vowels rather than phonemic diphthongs, we save the generalization that codas do not appear after phonemic falling-sonority diphthongs (diphthongs with a phonemic off-glide occupying the coda position). It is not simply the case, though, that analyzing the Sorbung vowel system into short-long pairs allows us to better capture the phonological patterns in the language than a quality-based analysis. It is also conceptually simpler and requires us to invoke fewer phonological parameters in order to characterize all of the contrasts in the system.

Aside from vowel length, one other aspect of our analysis of the vowels calls for comment. Although they are contrastive on the surface, it is possible to eliminate /uu/ and /u/ from the underlying vowel inventory if some abstractness in underlying representation is allowed. In stressed syllables, short [ʊ] only occurs before coronal codas (/n/ and /t/) and thus occurs in complementary distributions with [u], which never occurs in these environments. In isolation from other facts, [ʊ] could be reduced to an allophone of /u/. We treat it as distinct because [ʊ] can also surface as a result of stress-conditioned length alternations with [uu]. Compare, for example, *cíuu* ‘water’ and *cú-rée* ‘thirsty (water-thirst)’. Thus, the phonemic status of /u/ is dependent on the phonemic status of /uu/. As noted in Section 0 (above), Sorbung -uu corresponds to PKC *-uy and probably reflects earlier *-ui. As will be seen below, it also behaves like the falling-sonority diphthongs /ai/ and /au/ in that it never occurs with a coda. For this reason, it is tempting to consider [uu]

to be the realization of underlying /ui/. This would not be problematic, apart from the degree of abstraction involved, if there were not a number of lexical items in Sorbung with surface [ui]. All of these appear to be loanwords from Meithei or Ukhrul Tangkhul:

(40) LOANS FROM UKHRUL TANGKHUL

- a. *fúí* 'tempt'
- b. *lúí* 'finish'
- c. *ókət^hui* 'world'

(41) LOAN FROM MEITHEI

- cákúí* 'dance'

In principle, there is nothing to stop us from saying that there is a phonological process mapping /ui/ to [uɨ] that only affects the lexical stratum containing native vocabulary. Claims of the same type have been made for other languages (Itô & Mester 1995; Itô, Mester, & Padgett 1999). However, in the interest of descriptive neutrality, we have retained the more concrete analysis while highlighting some of its shortcomings.

Excluding a few loanwords which have final /l/, Sorbung allows major syllables to have one of seven coda consonants:

/p/ /t/ /k/
/m/ /n/ /ŋ/
/r/

With a few exceptions, these can occur after the long and short vowels and after the rising-sonority diphthongs /ia/, /ua/, and /iu/. They do not occur after the falling-sonority diphthongs, /ai/ and /au/, and after /uɨ/. **Table 18** shows which of the combinatorial possibilities between nuclei and rhymes are attested in our corpus. Examples are given in the subsequent tables.

Table 18: *Sorbung rhymes.*

–0	–r	–m	–n	–ŋ	–p	–t	–k
[a]	ar	am	an	aŋ	ap	at	ak
aa	aar	aam	aan	aaŋ	aap	aat	aak
[e]		em	en	eŋ		et	ek
ee				eeŋ			eek
[o]	or	om	on	oŋ	[op]		ok
oo		oom	oon	ooŋ	oop		ook
[i]	ir	im	in	iŋ	ip	it	ik
ii			iin	iiŋ	iip	iit	
[u]	ur	um	[ʉn]	uŋ	[up]	[ʉt]	
uu	uur	uum		uuŋ	uup		uuk
ʉʉ							
au							
ai							
ia	iar	iam		iaŋ			iak
ua	uar	uam	uan	uaŋ	uap	uat	uak
iu		ium					iuk

Table 19: *Examples of Sorbung open and r-final rhymes.*

	–0		–r	
a	<i>ciŋsiipá</i>	‘ant’	<i>ʔəhár</i>	‘new’
aa	<i>ʔaa</i>	‘fowl’	<i>ŋənáar</i>	‘snore’
e	<i>kəmiit</i>	‘ankle’	<i>cəm</i>	‘knife’
ee	<i>mələe</i>	‘tongue’		
o	<i>tətrò</i>	‘comb’	<i>mor</i>	‘mouth’
oo	<i>móo</i>	‘fire’	<i>caŋk^hoom</i>	‘milk’
i	<i>t^hiŋiuk</i>	‘bleed’	<i>mətír</i>	‘shrew’
ii	<i>t^hi</i>	‘die’		
u	<i>lukaan</i>	‘head’	<i>ʔəhúr</i>	‘frost’
uu	<i>məjúu</i>	‘mouse’	<i>júur</i>	‘nation’
ʉʉ	<i>habʉʉ</i>	‘molar’		
au	<i>jau</i>	‘sheep’		
ai	<i>mai</i>	‘face’		
ia	<i>ʔənfià</i>	‘enemy’	<i>t^hiár</i>	‘iron’
ua	<i>kəfuà</i>	‘rain’	<i>húar</i>	‘bright’
iu	<i>ciu</i>	‘dig’		

Table 20: *Examples of Sorbung nasal-final rhymes.*

	-m		-n		-ŋ	
a	<i>hambɨɨ</i>	‘tiger’	<i>wàn</i>	‘stomach’	<i>wáŋ</i>	‘fly (v.)’
aa	<i>səráam</i>	‘otter’	<i>laan</i>	‘err’	<i>lukaanŋ</i>	‘head’
e			<i>lèn</i>	‘close’	<i>seŋ</i>	‘clear’
ee					<i>kələeŋ</i>	‘loose’
o	<i>som</i>	‘ten’	<i>ʔítà kʰón</i>	‘my village’	<i>cóŋ</i>	‘speak’
oo			<i>məhɨtʰoon</i>	‘window’	<i>joŋ</i>	‘monkey’
i	<i>ʔənim</i>	‘shadow’	<i>mətín</i>	‘nail’	<i>maiɦìŋ</i>	‘freckle’
ii			<i>ʔiin</i>	‘drink’	<i>jiiŋ</i>	‘dark’
u	<i>kùmpʰék</i>	‘duck’	<i>kən</i>	‘straight’	<i>məlúŋ</i>	‘heart’
uu	<i>kʰuum</i>	‘sweet’			<i>fúuŋ</i>	‘cook’
ia	<i>səriám</i>	‘mithun’			<i>mətɨnliən</i>	‘eighth’
ua	<i>mùam</i>	‘hold’	<i>puan</i>	‘blanket’	<i>cəluáŋ</i>	‘to flow’
iu	<i>lækium</i>	‘bar’				

Table 21: *Examples of Sorbung stop-final rhymes.*

	-p		-t		-k	
a	<i>nàpkáap</i>	‘dry mucus’	<i>tʰàt</i>	‘kill’	<i>ʔəbák</i>	‘bat’
aa	<i>naap</i>	‘stick (v.)’	<i>kʰáat</i>	‘one’	<i>ŋàak</i>	‘wait’
e			<i>ʔarét</i>	‘eight’	<i>mooʔék</i>	‘hot coal’
ee					<i>ʔaméek</i>	‘meat’
o					<i>ʔacok</i>	‘frog’
oo	<i>kʰòop</i>	‘be satisfied’			<i>òok</i>	‘pig’
i	<i>məcíp</i>	‘frighten’	<i>mit</i>	‘squeeze’	<i>ʔarík</i>	‘louse’
ii	<i>lusìip</i>	‘top of head’	<i>mìit</i>	‘eye’		
u	<i>məjúp</i>	‘suck (v.)’	<i>kʰít</i>	‘arm’		
uu	<i>məjúup</i>	‘kiss (v.)’			<i>fúuk</i>	‘exit’
ia					<i>ʔəntiàk</i>	‘green’
ua	<i>ʔəciáp</i>	‘spleen’	<i>kʰuat</i>	‘scratch’	<i>ʔəkʰúak</i>	‘brain’
iu					<i>kèkiuk</i>	‘crippled’

4.1.2 Gaps

Two major nucleus-coda co-occurrence restrictions are found in Sorbung. The first applies to the falling-sonority diphthongs. As mentioned above, phonemic falling sonority diphthongs are never followed by coda consonants. Unlike /ia/, /ua/, and /iu/ that act as nuclei with codas, /ai/ and /au/ do not permit codas because the off-glide fills the coda position. In general, complex codas are not permitted in Sorbung. /ui/ follows the same pattern even though it is found only in loanwords from Ukhrul (Standard) Tangkhul and Meithei. Similarly, no codas appear after long [ɨɨ], which, as discussed above, is the modern reflex of historical *-ui.

- (42) a. *ʔə-sái* ‘elephant’ *sài-róoŋ* ‘slingshot’
 b. *jau* ‘sheep’ *ʔək^hauríaŋ* ‘grasshopper’
- (43) a. *ʔa-cùu* ‘chicken egg’ *mə-nuu* ‘laugh’
- (44) a. *ŋia* ‘stand’ *ʔən-tiàk* ‘green’
 b. *fíu* ‘hit’ *kè-kíuk* ‘crippled’
 c. *kà-fuà* ‘rain’ *fuan* ‘weak’
- (45) a. *ŋée* ‘white’ *kə-léŋ* ‘loose’
 b. *sóo* ‘run’ *òok* ‘pig’

As mentioned above, the coda restriction does not apply to the long vowels /oo/ and /ee/, though phonetically they are falling-sonority [ow] and [ej]. Our analysis of [ow] and [ej] as underlying long vowels, /oo/ and /ee/, explains why they function as monophthongal nuclei instead of falling-sonority diphthongs with regard to syllable structure.

4.2. Alternations

The phonological alternations affecting Sorbung rhymes include voicing alternations in stop codas and vowel shortening in unstressed syllables.

The voicing of stop codas is predictable from the phonological environment. When the following segment in a word is a vowel or voiced consonant, the coda is voiced. Elsewhere, when word-finally and before voiceless consonants, stops are then voiceless. (46a) demonstrates the changing voicing of the coronal stop in /miit/ ‘eye’. When this stem is compounded with /mii/ to form ‘eyebrow’ (literally ‘eye hair’), the final stop is voiced as it precedes a voiced segment. However, when followed by the voiceless consonant-initial /kor/, the stop remains voiceless. Other examples in (46) further demonstrate the same pattern:

- (46) a. *miit* ‘eye’ *mìd-mii* ‘eyebrow’
 mit-kor ‘eyelid’ *mìd-ék* ‘eye sand’
- b. *kút* ‘hand’ *kùd-mətín* ‘finger nail’
 kùt-k^hóom ‘handspan’
- c. *k^hòop* ‘full’ *k^hóob-maa* ‘full-question’
- d. *káap* ‘shoot’ *káab-òo* ‘shoot-ind’
- e. *ʔèk* ‘feces’ *ʔèg-len* ‘defecate’

This alternation cannot be attributed to a general process of voicing assimilation. Intervocalic voiceless stops occur frequently:

- (47) a. *lukaaŋ* ‘head’
 b. *kèkíuk* ‘crippled’
 c. *pátáabàa* ‘elder paternal uncle’
 d. *jáapee* ‘many’
 e. *k^hupú k^huncaaŋ* ‘storm’

The examples in (47) show that voiceless stop onsets are not subject to intervocalic voicing and therefore dismiss a general voicing assimilation rule. However, it is not clear whether the voicing alternation in Sorbung stop codas should be treated as the voicing of underlying voiceless stops before voiced segments, the devoicing of underlying voiced segments word-finally and before voiceless segments, or the “filling” of laryngeal specifications into segments underlyingly unspecified for voice. As both voiced and voiceless stops appear as onsets, intervocalically or otherwise, we cannot assume all stops are either voiced or voiceless. In Kom, an Old Kuki language also spoken in Manipur which shares many traits with Sorbung, we find a similar pattern of voicing (Grierson 1903:244). Word final voiceless stops become voiced when a vocalic suffix is added (48):

- (48) a. *kut* ‘hand’ *a-kud-a* ‘his hand on’
 b. *kaap* ‘shoot’ *a-kaab-a* ‘his shooting’

Like Sorbung, Kom also has examples of voiceless stops between vowels. While the available corpus of Kom data is not large enough to determine whether the alternations are identical, the similarity is suggestive.

In simple compounds consisting of two monosyllabic roots, the first syllable is consistently unstressed. If the first syllable’s vowel is short underlyingly, it will not be modified; if it is long, it will be shortened.

- (49) a. *kut* ‘hand’ *kùdbùu* ‘thumb’
 b. *cóŋ* ‘words’ *cóŋʔàak* ‘stutter’
- (50) c. *kʰoo* ‘bee’ *kʰocíu* ‘honey’
 d. *miit* ‘eye’ *mitkor* ‘eyelid’
 e. *haa* ‘tooth’ *habuu* ‘molar’
 f. *cuup* ‘breast’ *cupcuu* ‘milk’
 g. *tʰii* ‘blood’ *tʰiʃuuk* ‘bleed’
 h. *jaŋ* ‘penis’ *jaŋhun* ‘foreskin’
 i. *kèe* ‘leg, foot’ *kemíit* ‘ankle’

However, the constituent structure of compound formation alters stress and thus reduction. As expected, the vowel reduces in the unstressed position in (51b), but remains a long vowel in the stressed position in (51c). The reduction appears to only apply to instances of endocentric compounds of two simple, one-syllable roots.

- (51) a. *'kèe* ‘leg, foot’
 b. *ˌkè-'mii* ‘leg hair’
 c. *ˌkèe-mə'jáa* ‘footprint’
- (52) a. *mə'kʰáa* ‘jaw’
 b. *mə,kʰáa-'ruu* ‘jawbone’

In other complex compounds, especially those containing sesquisyllables, reduction does not apply in the same manner. (52b) displays a complex compound where the

nucleus of the head does not reduce. The length alternation follows almost entirely from the alternations in stress discussed in Section 0 above.

5. Conclusion

Leaving aside Brown's (1837) brief word list, this study is the first introduction of Sorbung/Southern Tangkhul to the scholarly community. We have introduced some of the significant properties of this language that are likely to be of interest to synchronic theorists, typologists, and comparativists. At the same time, a great deal of work remains to be done. Because our goals in this paper were descriptive, we have not ventured to provide a theoretical rationale for the patterns and alternations that we have reported. We have only ventured far enough into matters diachronic to promote our position that Sorbung is closer to Kuki-Chin than to Tangkhulic and that its phonology should be seen in this perspective. A detailed working-out of both the theoretical implications of the synchronic alternations in Sorbung phonology—cyclic stress assignment, opaque tone sandhi, morphologically conditioned place-assimilation, variable epenthesis, and problematic voicing alternations—will have to be left for further work. So, too, will detailed study of the historical development of Sorbung phonology. Likewise, we have scarcely touched upon the implications that this work has for understanding Tangkhul ethnogenesis and for plotting the historical relationship between the Tangkhuls of the Sorbung area and the rest of the Tangkhul ethnicity. What we do hope to have established is that Sorbung presents interesting data and problems to both the historical and synchronic linguist.

References

- Brown, Nathan. 1837. Comparison of Indo-Chinese languages. *Journal of the Asiatic Society of Bengal* 6(2). 1023-1038.
- Button, Christopher Thomas James. 2009. A Reconstruction of Proto Northern Chin in Old Burmese and Old Chinese Perspective. SOAS, University of London phd dissertation.
- Burling, Robbins. 2003. The Tibeto-Burman Languages of Northeast India. In G. Thurgood and R. H. LaPolla (eds), *The Sino-Tibetan Languages*, 169–192.
- Chelliah, Shobhana Lakshmi. 1997. *A Grammar of Meithei*. (Mouton Grammar Library 17). Berlin: Mouton de Gruyter.
- Grierson, George A. 1903. *Tibeto-Burman Family: Specimens of the Kuki-Chin and Burma Groups*. Vol. 3. (Linguistic Survey of India). Calcutta: Office of the Superintendent of Government Printing.
- Hartmann, Helga. 2001. Prenasalization and preglottalization in Daai Chin, with parallel examples from Mro and Maru. *Linguistics of the Tibeto-Burman Area* 24(2). 123–142.
- Hyman, Larry M. 2007a. Elicitation as experimental phonology: Thlantlang Lai tonology. *Experimental Approaches to Phonology*. 7-24.
- Hyman, Larry M. 2007b. Kuki-Thaadow: An African tone system in Southeast Asia. *UC Berkeley Phonology Lab Annual Report*.
- Hyman, Larry M. & Kenneth VanBik. 2002. Tone and stem2 formation in Hakha Lai. *Linguistics of the Tibeto-Burman Area* 25(1). 113-120.

- Hyman, Larry M. & Kenneth VanBik. 2004. Directional rule application and output problems in Hakha Lai tone. *Language and Linguistics* 5(4). 821–861.
- Itô, Junko & Armin Mester. 1995. Japanese phonology: constraint domains and structure preservation. In , *The handbook of phonological theory*, 817-847.
- Itô, Junko, Armin Mester & Jaye Padgett. 1999. Lexical classes in Japanese: a reply to Rice. *Phonology at Santa Cruz* 6. 39-46.
- La Polla, Randy J. 2006. The Role of Migration and Language Contact in the Development of the Sino-Tibetan Language Family. In Alexandra Y. Aikhenvald & R. M. W. Dixon (eds.), *Areal diffusion and genetic inheritance: problems in comparative linguistics*. Oxford: Oxford University Press.
- Matisoff, James A. 2003. *Handbook of Proto-Tibeto-Burman: System and Philosophy of Sino-Tibetan Reconstruction*. Berkeley: University of California Press.
- Mortensen, David R. 2003. *Comparative Tangkhul*. University of California, Berkeley, ms.
- Mortensen, David R. & James A. Miller. 2009a. A Reconstruction of Proto-Tangkhulic Rhymes. Paper presented at the SEALS 19, Ho Chi Minh City, Vietnam.
- Mortensen, David R. & James A. Miller. 2009b. Proto-Tangkhul Onsets in Comparative Perspective. Paper presented at the ICSTLL 42, Chiangmai Thailand.
- VanBik, Kenneth. 2009. *Proto-Kuki-Chin: A Reconstructed Ancestor of the Kuki-Chin Languages*. Vol. 8. (STEDT Monograph Series). Berkeley: STEDT.
- Yip, Moira. 2004. Phonological markedness and allomorph selection in Zahao. *Language and Linguistics* 5(4). 969-1001.

Appendix: Sorbung Word List

1dl	ʔərúkəniŋ
1pl	ʔənək ^{huo}
1sg	ʔoo
1sg-gen	ʔitə
1sg-refl	ʔóoməsaa
2dl	nàrúkəniŋ
2pl	naaŋnək ^{huo}
2sg	naaŋ
2sg-gen	nanta
2sg-refl	naaŋməsaa
3dl	mərúkəniŋ
3pl	maanək ^{huo}
3sg	maa
3sg-gen	mata
3sg-refl	maaməsaa
Gloss	NewTrans
abdomen (external bulge); belly	wan
able	kòomán
accept	jàa
accustomed	məfúhaa
adam's apple; larynx (throat)	kəráak
add together	mətùuk
adult	santáaria
afterbirth; placenta	ʔəlaam
alive; living, be	ríiŋ
animal	saajúur
animal	sàa
ankle	kəmíit
ant	ciŋfiipá
anteater (pangolin); crocodile	təŋgooŋ
antler (of deer)	ʔəkii
anus; rectum	ʔəkái
arm, upper	paambúm
arm; hand	kút
armpit; underarm	cùbláa
arrive	həŋguŋ
arrow	t ^h èe
artery; blood vessel; vein	məŋeet ^h áaruu
ascend; climb	háaŋ
ashamed; shy	jáak
asleep, be; sleep	ʔiin
aunt, elder paternal	ʔənnii
aunt, maternal	nugənuu
aunt, wife of elder paternal uncle	nútáanu

aunt, wife of maternal uncle	píi
aunt, wife of younger paternal uncle	núgənùu
aunt, younger paternal	ʔənnii
awake, be; wake up (v.i.)	ʔənt ^h èe
axe	ròo
back	kòoŋ
back (of something)	ʔənúuŋlám
back basket	məkóoŋ
back, lower	kòoŋcat
backbone; spine	kòoŋrúu
bad (morally)	saaleè
bamboo spoon for curry	mək ^h éè
bar	lækium
bat	ʔəbák
bathe	cəʃíi
beak; bill (of bird)	mòr
bear (Selenarctos thibetanus)	ʔəwám
bear; birth, give	mədəe
beard	mòrmíi
beautiful	ʔəloo
become old	taalen
bee	k ^h oo
behind	ʔənúuŋlám
belly; abdomen (external bulge)	wan
big	lòok
bile	məʃiŋk ^h aa
bill (of bird); beak	mòr
bird	ʔəwaa
birth, give; bear	mədəe
bite (v.)	ʔóo
bitter	k ^h aa
black	wám
blanket; cloth	púan
bleed (v.)	t ^h iʃúuk
blemish on skin	maihìŋ
blemish; mole	səmàamíi
blind	mìtcòo
blink (v.)	mìtk ^h əjùp
blood	t ^h ii
blood vessel; vein; artery	məŋeet ^h áaruu
blow (mouth)	səm
blow fire	mò-səm
body	məsàa
body, dead; corpse	t ^h iilenta məsàa
boil	bùt
bone	ʔərúu

bone, jaw; jawbone; mandible	mək ^h áaruu
bone, malar; cheekbone	mairúu
born, be	mədəe
borrow	màmbúk
bosom	məlúuŋ
boundary	t ^h ərii
bow	t ^h èe
bowels; intestines	ʔərii
bowl	téŋguàt
brain	ʔək ^h úak
break	mək ^h óo
break (glass, egg); split (a watermelon, other round object)	k ^h òo
breast	cúup
breath	ʔəp ^h üt
breathe (v.)	ʔəp ^h üt-lée
bright	húar
brother of female, elder	moo
brother of female, younger	ʔináupa
brother of male, elder	moo
brother of male, younger	ʔináupa
brother-in-law (of ego)	ʔinèeŋ
brow; forehead	məcèe
brush	ʔətáat
buffalo	səloo
bull; cattle, male	səmùk paasaa
burn off fields; slash and burn	lèe-hèe
burn; on fire, be	moo-còk
burnt	kaaŋlen
burst	mək ^h óo
butterfly	ʔəpàak
buttocks	ʔəkáí
buy	ʔəlóo
calf; shin	kèemərái
call	kèe
carry	p ^h ùa
carry on the head	tooŋ
carry with hand; lift up	k ^h aaŋ
cat (Felis domesticus)	k ^h aan
catch	ʔəráak
catch; drive; chase	ʔəjíi
cattle	səmuk
cattle, female; cow	səmùk mòosa
cattle, male; bull	səmùk paasaa
cave; hole (in rock)	luŋk ^h ua
charcoal	məhée

chase; catch; drive	ʔəjii
cheap	pai
cheek	mai
cheekbone; bone, malar	mairúu
chest (=liver)	mət ^h ín
chew (v.)	ʔóo
chicken egg	ʔacùu
chicken, jungle	ʔarip
chicken; fowl	ʔaa
chief	ʔəhuáŋbáa
child	ʔəŋgúu
chin; jaw	mək ^h áa
choose	kəplāaŋ
chop (at tree)	dùu
clan; surname	səkái
claw (of animal)	mətín
clay	lámph ^h u
clean	t ^h àar
clear	seŋ
clever; smart; wise	t ^h àaŋmee
climb; ascend	háaŋ
close; shut	lèn
cloth; blanket	púan
cloud	k ^h ám ^h báa
coccyx; tailbone	ʔəkáiruu
cock (intact male)	ʔak ^h óoŋ
cold, very; quiet, very	dàiriŋi
cold; quiet	dài
comb (of rooster)	tətrò
come	hóoŋ
come back (imperative)	hoóŋledee
conceive	mərái
cook	ʃúuŋ
copulate (v.); have intercourse	ʔətéŋná
cord, one; fathom, one	laam k ^h àt
corner	ʔəkíi
corpse; body, dead	t ^h iilenta məsàa
correct	kún
cosmos; world	ókət ^h ui
cough	mək ^h áa
cow; cattle, female	səmùk mòosa
crab	ʔəjii
crawl	ʔəwák
crayfish	tukòŋ(g)lá
creator of the universe	semnùu sempáa
crippled; lame; handicapped	kèkiuk

(whether or not legs are involved)	
crocodile; anteater (pangolin)	təŋgoŋ
crooked	mək ^h ái
cross	ʔəkaan
crossbar	tətír
crotch; fork of legs	kèemətóo, kèemədóo
crow (v.); sound (v.i., as a musical instrument)	ʔək ^h uàn
crown; pate; head, top of	lusìip
cry; weep	càap
cut (wood, vegetables)	táan
dance	cəkúí
dark	jíiŋ
darkness	ʔəjíiŋ
daughter	saa
daughter-in-law (of ego)	ʔihàa
day	ʔəʃun
dead, be; die	t ^h íi
deaf	nàk ^h uák
decay	məmǎaŋ
deceive	mináam
deep	t ^h ùuk
deer	səkíi
defecate (v.)	ʔèklen, ʔèglen
descend	júuŋ
desire; want	pam
devil, ghost, demon, evil spirit	rambúu
dew	ʔədáícuu
dhole	ràmʔúu
die; dead, be	t ^h íi
difficult	lùu
dig	ciu
digit	-məjúumràa
dirty	múat
dish out; scoop out	ʃuaglee
divide; split	mətáan
do	koo
dog (Canis familiaris)	ʔúu
door	iink ^h aa
dove; pigeon	ʔəʃúu
dream	máaŋ
drink (v.)	ʔíin
drive; chase; catch	ʔəjíi
drought	cəkáaŋ
dry	t ^h aaŋ
dry land (vs. water); earth; ground	ʔəloo

duck	kùmp ^h ék
dull (not sharp)	ŋáimai
dumb; mute	coŋjúuk
dust; powder	ʔəhút
eagle	bùumlík
ear	ʔənáa
ear canal	nák ^h ùur
eardrum	nájàaŋ
early; fast	t ^h àk
earring	ʔənáat ^h iŋ
earth; ground; dry land (vs. water)	ʔəloo
earthquake	ràmbuu háaŋ
earwax	ʔənáaʔék
easy	pai
eat	saa
egg	ʔəwaacúu
eggs, fish; roe	ʔəŋaacúu, ʔŋaacúu
eight	ʔərét
eighteen	som ʔərét
eighth	mətunliar
elephant	ʔəsái
eleven	som k ^h áat
ember; hot coal	moʔék
enemy	ʔənʃià
erase a chalkboard; extinguish	məmìt
err	laan
evening	k ^h əleelám
exist	ʔóm
exist, to	ʔòmna
exit	ʃúuk
exit, to	ʃùgná
expect	ʔədúuk
expensive	táaŋ
extinguish; erase a chalkboard	məmìt
extinguish; put out fire	məmìt
eye	mìit
eye sand	mìdék; mìtʔék
eye, white of	mìdèk, mìtʔèk
eyeball	mìit
eyebrow	mìdmíi
eyelash	mìdmii
eyelid	mìtkor
face	mai
fall	ʔəndía
family	ʔinʃúnjk ^h uar
far	ʃóo

fart	məcipát
fast; early	t ^h àk
fat	ʔət ^h áu
fat	t ^h áu
father (of ego)	ʔəmbaa
father-in-law (of ego)	ʔəmbúu
fathom, half	laam təŋk ^h ai
fathom, one; cord, one	laam k ^h àt
fear	ʔəncíi
feather (of bird)	ʔəmíi
feces; shit	ʔèk
femur; thigh bone	mətóoruu
few	kaajàŋ
field	lèe
fifteen	som rəŋáa
fifth born	t ^h oombáa
fifty	sòŋŋáa
fin; scale of fish	ʔəŋámíi
finger	kədməjúumràa, kətməjúumràa
finger width, one	kədməjumráa k ^h àt
finger, little	kùsáa
ingernail	kədmətín
finish	lúí
fire	móo
first born	moopaa
fish (general)	ʔəŋaa, ʔŋaa
fish species, light in color	ʔəŋk ^h au
fish, catfish	ʔəŋoŋ(g)lá
fish, flying	ŋámíi
fishing cat	səráam
fit	càa
five	rəŋáa
flea	ʔurík
flesh	ʔətáak
flood	cəlòk
flow (v.)	cəluáŋ
fly	t ^h əwái
fly (v.)	wáŋ
follow	ʔənúuŋ-aajíi
foot	kèe
footprint	kəlúuŋ
forehead; brow	məcèe
foreskin	jaŋhún
forget	mənii-lèn
fork of legs; crotch	kèemətóo, kèemədóo
four	məlíi

fourteen	som mǎlii
fourth born	miipáa
fowl, male; rooster	ʔak ^h óoŋ
fowl; chicken	ʔaa
fox	rámʔúu
freckle	maihiŋ
friend	ʔəjòok
frighten	məcíp
frog	ʔəcok
front (of something)	məŋaalam
front of, in	məŋaalam
frost	ʔəhúr
fruit	t ^h iŋraa
fry	t ^h àu-jáu
fry	jáu
full; satiated, be	k ^h òop
fur (of animal)	ʔəmíi
gape; open mouth (v.)	ʔàaŋ
get somebody up	mət ^h ée
get up; rise	ʔənt ^h ée
give	pée
go	jáu
goat	màŋgǎlée
goatee	mòrmíi
gold	sənáa
good	ʔəloo
grandchild (of ego); son's son	ʔisàata ʔisàa
grandfather, maternal (of ego)	ʔəmbúu
grandfather, paternal	ʔəpùu
grandfather, paternal (of ego)	ʔəmbúu
grandmother, maternal (of ego)	ʔəmbíi
grandmother, paternal (of ego)	ʔəmbíi
grasshopper	ʔək ^h aurúaŋ
grasshopper	ʔək ^h au, ʔk ^h au
grease (for cooking); oil	t ^h àu
green	ʔəntiàk
grind with pestle	p ^h ənaat
grobe (in the dark)	mət ^h áp
ground; dry land (vs. water); earth	ʔəloo
guest	məlóoŋ
gums	hákhíi
hail; sleet	ʔəriáráa
hair (general)	ʔəmíi
hair of head	sàam
hair, body	ʔəmíi
hair, facial	mòrmíi

hair, female pubic	ʃirmii
hair, leg	kəmii
hair, male pubic	jàŋmii
hair, underarm	cùbláamii
hand	kút
hand span	kùtk ^h óom
hand; arm	kút
handicapped (whether or not legs are involved); crippled; lame	kèkiúk
hard; tough	lùu
hat; headdress	təŋguáp
have intercourse; copulate (v.)	ʔətəŋná
head	lukaan
head, top of; crown; pate	lusii
headdress; hat	təŋguáp
hear (v.)	t ^h óo
heart	məlúŋ
heart; locus of anger	məlúuŋ
heart; locus of emotions	niiŋ
heartbeat (=breath)	ʔəp ^h ùt
heavy	riit
help	páan
hen	ʔaláa
hen, mother	ʔabúu
hiccup	səkái
hide	mət ^h úup
hide; leather (dried animal skin)	ʔəmèkhún
high	məsáan
hill	ciŋtòon
hips	ʔəkái, ʔkái
hit (with stick)	p ^h iit
hit (with the fist)	ʃiu
hold in mouth (v.)	mùam
hole	ʔək ^h uá
hole (in rock); cave	luŋk ^h ua
honey	k ^h òcúu
horn (of animal)	ʔəkii
horse	səkúan
host (house owner)	ʔinpu
hot	sáa
hot coal; ember	moʔék
house	ʔin
humble	t ^h òonáa
hunchback	kòonk ^h ùu
hundred	ʔəjaa k ^h áat
hungry, be	bùcám

hunt	ʔəjii
hurt	mət ^h ái
hurt; sick, be; ill, be	náa
husband	bəsée
ice	ʔəhúr
ill, be; hurt; sick, be	náa
insects	ʔəkúu ʔəkái
intestines; bowels	ʔərii
iron	t ^h íár
itch (v.); itchy, be	mət ^h àk
itchy, be; itch (v.)	mət ^h àk
jaw; chin	mək ^h áa
jawbone; mandible; bone, jaw	mək ^h áaruu
joint	mətèt ⁿ à
kick	məfùuk
kick; knead with the feet	kau
kidney	məkee
kill	t ^h àt
kiss (v.)	məjúup
knead with the feet; kick	kau
knee	ʔək ^h úk, ʔk ^h úk
knee cap; patella	k ^h úk
knife	cèm
lac insect	ʔəriip
ladder	ʔəláak
lame; handicapped (whether or not legs are involved); crippled	kèkiuk
langur; leaf-monkey	cəháa
lap	mətóo
large intestine	ʔəriilookpaa
larynx (throat); adam's apple	kəráak
late; slow	fùum
laugh	mənúu
lay egg	ʔəcúu-məcúu
lead; start	ʔənt ^h ée
leaf-monkey; langur	cəháa
leather (dried animal skin); hide	ʔəmèkhún
leech (land)	məhít
left hand	təmbuelàm
leg	kèe
lend	wàaj
length from thumb to forfinger	kap k ^h àt
leprous	ʔəhún-náa
lick (v.)	mələék
lift up; carry with hand	k ^h aaŋ
ligament (bone to bone)	ʔəlíg ^l ák

light (weight)	jaaŋ
lightning	k ^h umléeláap
like	pam
like, be; resemble	máan
lion	hambúu
lip	mòr
liver	mət ^h in
living, be; alive	riiŋ
locus of anger; heart	məlúuŋ
locus of emotions; heart	niiŋ
locus of thought; mind	niiŋ
long	ʃoo
look	ʔen
look (imperative)	ʔenee
loom	puank ^h ooŋnaa
loose	kələeŋ
loose; untie	tərət-kəlen
lost	máaŋ
louse	ʔərík, ʔrík
low	mənəm
man	paasaa
man, blind	mitcòbàa
mandible; bone, jaw; jawbone	mək ^h áaruu
mantis	ʔəsóm
many; very	jáapee
marrow	ʔəlíglák
meat	ʔəméek
medicine	lèe
milk	cùpcuu
milk	caŋk ^h oom
mind; locus of thought	niiŋ
mithun	səriám
mock	mənuyfá
molar	habuu
mole	ʔəbúu
mole; blemish	səmàamíi
monkey	joon
month	ʔək ^h áa
moon	ʔək ^h áa
morning	ʔək ^h uanlám
mosquito	t ^h əwái
moth	ʔəpàa
mother (of ego)	ʔənnúu
mother-in-law (of ego)	ʔənnii
mountain	cíiŋtòon
mouse	məjúu

moustache	mòrmíi
mouth	mor
mucus, dry	nàpkáaŋ
mucus, liquid	nàap
muscle (=flesh)	ʔəták
mushy (of rice, crushed banana); soft	ʔəpáat, ʔpáat
mute; dumb	coŋfúuk
my village	ʔitá k ^h ón
nail	mətín
name	ʔəmín
narrow	nèe
nation	júur
nationality	júur
navel	p(ə)lái
near	ʔənái
neck	riiŋ
necklace	məʃi
needle	ʔəʃim
nephew (of ego)	ʔitùsáa
nest	ʔəwaabúu
new	ʔət ^h ár
niece (of ego)	ʔitùsáa
night	ʔəján
nine	ʔəkúa
nineteen	som ʔəkúa
ninth	moolét
nipple	cùbráa
noon	ʔəʃun
nose	nàráa
nose bridge	nàráa
nostril	nàk ^h úur
nurse; suckle (v.) (=drink)	ʔiin
oil; grease (for cooking)	t ^h àu
old	ʔəlúu
old (age)	séndaa
old folks	ʔətaaria
on fire, be; burn	moo-còk
one	k ^h áat
open	məhóoŋ
open mouth (v.); gape	ʔàaŋ
orphan	səràasa
otter	səráam
owl	ʔəwaahambúu
palm	kùdməjáa
parrot	bəkíi
pate; head, top of; crown	lusiip

patella; knee cap	k ^h úk
paternal grandmother	ʔəpii
path; road	lambùu
path; this road	lambùuwá
peel	ʔəkór
penis	jaaŋ
person, blind	mitcòpà, mitconu
person, deaf	nàak ^h uàk
person, old	santáarià
perspiration; sweat	mək ^h áncuu
pestle	təŋfuk
pig	òok
pigeon; dove	ʔəʃúu
placenta; afterbirth	ʔəlaam
pond; pool	puk ^h rii
pool; pond	puk ^h rii
porcupine	səkúu
pot	təgók
pot, clay	lamp ^h ù təkòo, dəgòo
pound rice	ʔənp ^h ee-néŋ
powder; dust	ʔəhút
practitioner of witchcraft	lévárajuán
pregnant, be	mərái
price	ʔəmáan
pus	ʔənái
put	ləʃua
put out fire; extinguish	məmit
python	bəriilòokpáa
queen (no such office)	ʔəhuáŋnúu
quiet, very; cold, very	dàiriʃi
quiet; cold	dài
rain	kəʃua
rat	məjúu
rectum; anus	ʔəkái
red	ʔaaŋ
release	mək ^h áan-lèn
resemble; like, be	máan
respect	k ^h əjájáʃá
rib cage	wándà ʔəríu
right hand	ʔəncáaŋlám
rise; get up	ʔənt ^h ée
river	cəkòo, cəgòo
road; path	lambùu
roast	ree
roe; eggs, fish	ʔəŋaacúu, ʔŋaacúu
roll (v.i.)	laam

roll (v.t.)	màŋk(ə)liŋ
rooster; fowl, male	ʔak ^h óoŋ
rope	ræu
round	ʔəlúum
rub	tai
run	sóo
rust	ʔədáahàaŋ
saliva	məcámçàu
salt	məci
sambar; sambhur	səjáa
sambhur; sambar	səjáa
sand	ʔəʃi
satiated, be; full	k ^h òop
scab (hardened rice at bottom of pot)	ʔəkaaŋ
scale of fish; fin	ʔəŋámíi
scar	tàak
scoop out; dish out	ʃuaglee
scrape; scratch	ʔəhuát
scratch	k ^h uat
scratch; scrape	ʔəhuát
search; seek	jóoŋ
second born	k ^h oopaa
see (v.)	mùu
seek; search	jóoŋ
sell	júa
semen; sperm	jaŋçàu
seven	səríi
seventeen	som səríi
seventh	loopáa
sew	dùu
shadow	ʔənim
shadow	ʔəním
shallow	tàar-ləŋləŋ
sharp	ŋai
sharpen knife	cəm-ʔətát
sheep	jau
shin; calf	kèemərái
shit; feces	ʔèk
shiver; shudder; tremble	sái
shoot (v.)	káap
short (height)	tòo
short (length)	tòo
shoulder	lək ^h àa
shout	wàaw
shrew	mətír
shudder; tremble; shiver	sái

shut; close	lèn
shy; ashamed	jáak
sick, be; ill, be; hurt	náa
side (left)	tənpùelàm
side (right)	ʔəncáaŋlám
sinew; tendon (muscle to bone)	ʔəlíg-lák
sing	laasáa
sister of female, elder	tee
sister of female, younger	ʔinàunù
sister of male, elder	tee
sister of male, younger	ʔisaanù
sister-in-law (of ego)	ʔimèe
sit	ʔəŋjũŋ
sit on eggs	buum
six	kəruúk
sixteen	som kəruúk
sixth born	joombáa
skeleton	ʔəru-ráaŋ
skin	ʔəhún
skinny	kòŋ
skull	lukáaŋruu
skull	mairúu
sky	ʔəwaarám
slash and burn; burn off fields	lèe-hèe
sleep; asleep, be	ʔiin
sleet; hail	ʔəri-ráa
slingshot	sàiróŋ
slow; late	ʔum
small	nèe
small intestine	ʔəri-inèepaa
smart; wise; clever	tʰàaŋmee
smell; sniff (v.)	mənáam
smile (v.)	mənúu
smoke	mok ^h út
snail	càbrùulá, càbrùurá
snake	bərii
snake species, large	bəriis(ə)mùuk
sneeze	hát ^h ii
sniff (v.); smell	mənáam
snore	ŋənáar
snow	ʔəhúr
soft	mənée
soft; mushy (of rice, crushed banana)	ʔəpáat, ʔpáat
sole	kèeməjáa
son	saa
son's son; grandchild (of ego)	ʔisáata ʔisáa

son-in-law (of ego)	ʔihàa
song	laa
soul, spirit (Meithei)	t ^h əwái
sound (v.i., as a musical instrument); crow (v.)	ʔək ^h uàn
sour	t ^h uur
spade	bélcàa
speak (v.)	cóŋ
spear	ʔəsóo
speech; words	cóŋ
sperm; semen	jaŋcàu
spider	k ^h əràan, k ^h ràan
spin	məlám
spinal cord (=marrow)	ʔəlíglák
spine; backbone	kòŋrúu
spit (v.)	məcamcúu mət ^h úur
spit (v.)	məcamcúu məsàk
spleen	ʔəcúap
split (a watermelon, other round object); break (glass, egg)	k ^h òo
split (as with a candy bar)	ʔətáanjèe
split; divide	mətáan
spring	cəlám
squeeze	mit
squeeze; twist	mit
squirrel	ʔəlóo
squirrel, giant	ʔəlóo
stammer; stutter	cóŋʔàak
stand	ŋia
stand up (imperative)	ŋiaèe
star	ʔəʃii
start; lead	ʔənt ^h éee
stature	ʔəráan
steal	mərúu
stick (v.)	náap
stink	mənámʃoo
stomach	wan
stone	lùuŋ
storm	k ^h əpú k ^h uncaan
straight	kún
strong	məkàt
stupid	pàan, ʔəpàan
stutter; stammer	cóŋʔàak
suck (v.)	məjúp
suckle (v.) (=drink); nurse	ʔiin
sun	ʔəníi

sunrise	ʔənií-fuuk
sunset	ʔənií-lút
surname; clan	səkáí
swallow (v.)	məjúu
sweat	mək ^h áan-cuu
sweat; perspiration	mək ^h áncuu
sweet	k ^h uum
swell; swollen, be	wór
swollen, be; swell	wór
tail (of animal)	ʔəmóo
tailbone; coccyx	ʔəkáiruu
take	lée
tall	ʃoo
talon (of bird)	mətín
tear (n.)	mək ^h ácuu
tempt	ʃúi
ten	som
ten thousand	lèeʃiŋ som
tendon (muscle to bone); sinew	ʔəlígłák
termite; white ant	lùuksáa
testicle	jàŋráa
that	kaa
the muscle comes out	ʔəták ʃuugòo
the tail is long	ʔəmóo sóoòo
the tail is short	ʔəmóo tòoòo
thick	lòok
thigh	mətóo
thigh bone; femur	mətóoruu
thin	məmpáar
think	ʔədúuk
third born	teepáa
thirsty, be	curee
thirteen	som ʔənt ^h úum
thirty	somt ^h úum
this	wa
this axe	ròowá
this blanket; this cloth	puanùá
this boundary	t ^h əriiwá, t ^h əriivá
this bow	t ^h èewá
this cloth; this blanket	puanùá
this corner	ʔəkíiwá
this crossbar	tətirùá
this door	iij ^h aavá
this earring	ʔənáat ^h iijùá
this knife	cèmùá
this ladder	ʔəlágùá

this loom	puank ^h óoŋnaavá
this medicine	lèewá
this necklace	məʃiíwá
this needle	ʔəʃimùá
this pestle	təŋʃugùá
this pot	təgógùá
this price	ʔəmáanùá
this road; path	lambùuwá
this rope	rúuwá
this slingshot	sáiròoŋùá
this song	laavá
this spade	bélcàavá
this spear	ʔəsóowá
this wall	p ^h áaglànùá
this war	réewá
this window	məhʉt ^h oonùá
thousand	lèeʃiŋ k ^h áat
three	ʔənt ^h úum
throat; uvula	kəráak
thumb	kùdbúu
thunder	məcék kàa
tie	təràt
tiger	hambúu
tight	cin
to dry	məkaaŋ
toad	ʔəcok
toast	məhai
toe	kèeməjúumràa
toe, big	kèbúu
toe, little	kèsáa
toenail	kèemətín
tongs	cəkáap
tongue	mələée
tooth	haa
torso; trunk	məsàa
touch	còo
tough; hard	lùu
trachea; windpipe	kəráak
tremble; shiver; shudder	sái
trunk; torso	məsàa
twelve	som kəníiŋ
twenty	semnii
twist; squeeze	mit
two	kəniŋ
udder (of cow, goat)	cúup
ugly	cáakʃa

umbilical cord	p(ə)láí
uncle, elder paternal	pátáabàa
uncle, husband of elder paternal aunt	pùu
uncle, husband of maternal aunt	pakùu
uncle, husband of younger paternal aunt	pùu
uncle, maternal	pùu
uncle, younger paternal	pakùu
underarm; armpit	cùbláa
untie; loose	tərət-kəlen
urinate (v.)	məcée-len
urine	məcécècùu, məcécècuu
uvula; throat	kəráak
vagina	ʃír
valley (river valley)	təmbàak
vein; artery; blood vessel	məŋeet ^h áaruu
vertebrae (of spine)	kòonrúu
very; many	jáapee
village	ʔək ^h uan
voice	ʔək ^h úan, ʔk ^h úan
vomit (v.)	məlít
wait	ŋàak
wait (imperative)	ŋàagèe
wake up (v.i.); awake, be	ʔənt ^h èe
walk	jáu
wall	p ^h áaglàn
want; desire	pam
war	rée
wart (= skin disease)	ʔəhúnáa
wash (clothes)	ʔəʃúu
wash (hands, vegetables, cars)	p ^h ərii
water	cúu
weak	ʃuan
weep; cry	càap
well, be	ʔəloo
wet	juu
whiskers (of animal)	mòrmíi
whistle	məhúu
white	ŋée
white ant; termite	lùuksáa
whole	kəlúktaee
widow	nəmoomək ^h úu
wife	mənamóo
wind	k ^h ii
window	məhut ^h oon
windpipe; trachea	kəráak

wing (of bird)	p ^h əláa, p ^h láa
wise; clever; smart	t ^h àaŋmee
wolf	ràmʔúu
woman	moosáa
woman, blind	mìtcònuu
woodpecker	ʔəwaat ^h ìŋhútpá
words; speech	cóŋ
world; cosmos	ókət ^h ui
yawn (v.)	məhàm
year	kuum
yellow (color of curry)	jaiŋàaŋmæcuu
young (age)	mənée

FLOATING QUANTIFIERS IN BURMESE AND THAI

Andrew Simpson

University of Southern California
<andrew.simpson@usc.edu>

Abstract

This paper considers patterns of floating quantifiers in Thai and Burmese and attempts to provide answers to three general questions. First, syntactically how are nouns and floating quantifiers displaced from each other? Second, what pragmatic/functional motivations underlie the optional use of floating quantifier structures? Third, are patterns of Q-float uniform across languages, or are there differences, and how might these be accounted for? It is argued that Q-float differences found in Thai and Burmese are due to the interaction of universal principals shaping linear word order (information structure) and language particular syntactic organisation (the head-initial/final parameter).

Keywords: quantifiers, topicalization, numerals

1. Introduction

In many languages certain elements which quantify over noun-phrases/NPs, such as numerals and quantifiers equivalent to English ‘all’ and ‘both’, may most frequently occur adjacent to the NP they modify, but at other times occur separated from the NP they relate to, when they are commonly referred to as ‘floating quantifiers’. Typical alternations between the NP-adjacent and floating occurrences of nominal quantifiers are illustrated in examples (2-4) from English, Thai and Burmese following the linear schematization in (1) (Q = quantifier).³⁰

- (1) non-floated patterns: NP Q..... or Q NP....’
common floated pattern: NP Q.....
less common floated pattern: Q NP....

ENGLISH

- (2) a. All the students have arrived. *non-floated*
b. The students have all arrived. *floated*
NP = [the students]
Q = [all]

³⁰ The Thai and Burmese data presented in the paper was mostly gathered in a series of interviews carried out in 2004 with multiple native speakers of Thai and Burmese who were journalists regularly writing/broadcasting in Thai and Burmese, or teachers of one of the two languages. A preliminary version of the paper was subsequently presented at the SEALS conference in Chulalongkorn University in 2005. Sincere thanks are due to the language consultants who have helped with clarification of the patterns reported in the paper, and to two anonymous reviewers of the paper.

- THAI
- (3) a. **nāksùksáa thūk-khon** maa lēew *non-floated*
 student every-CL come ASP
 ‘All the students have arrived.’
- b. **nāksùksáa** maa lēew **thūk-khon** *floated*
 student come ASP every-CL
 NP = [nāksùksáa] ‘student’
 Q = [thūk-khon] ‘every’
- BURMESE
- (4) a. cāun-ko **sīqthāa** **avāuq** **hnayaa** laa-thwāa-te *non-floated*
 school-to soldier CL 200 come-ASP-REAL
 ‘200 soldiers came to the school.’
- b. **sīqthāa** cāun-ko **avāuq** **hnayaa** laa-thwāa-te *floated*
 soldier school-to CL 200 come-ASP-REAL
 NP = [sīqthāa] ‘soldier’
 Q = [avāuq hna-yaa] ‘200’

Floating quantifier patterns are linguistically interesting because they appear to show that it is possible to pull apart and separate two parts of a single grammatical unit (for example, a subject or object) into a discontinuous sequence of elements which are nevertheless still interpreted together, the quantifier being understood as modifying the reference of the NP. This gives rise to three important, general questions. First, how does this syntactic separation occur - what is the structure of non-floated NP/Q constituents, and how does it allow for the NP and the quantifier to be separated? Second, why are the NP and quantifier sometimes separated from each other? When elements are distanced from the position they would otherwise normally occur in, this may make sentences more difficult for hearers to parse and process. What functional purpose and benefit might there be in sometimes splitting apart NP and quantifier constituents? Third, there is a general typological question of whether floating quantifier patterns are indeed the same across different languages, and if not, what are the relevant differences and how might these be accounted for? Is it possible to identify any cross-linguistically shared properties in floating quantifier constructions?

This comparative study of floating quantifiers in Burmese and Thai sets out to investigate these issues and provide initial answers to these three questions which will hopefully be of use in the continued study of floating quantifiers. Concerning the particular choice of languages focused on in the present the study, it is interesting and potentially revealing to compare floating quantifier phenomena in Burmese and Thai because although the structure of NPs is similar in many ways in the two languages, the basic word order of Burmese and Thai is fundamentally different, Burmese being an SOV head-final language, while Thai is a language with very typical SVO patterns. In the course of the paper, it will be suggested that differences in basic word order may indeed impact on the way that the phenomenon of floating quantifiers occurs in a language. In addition to Burmese and Thai, the paper will also make comparative reference to certain other

languages in which significant work on floating quantifiers has been carried out, in particular English, and also Japanese and Korean.

**2. Previous syntactic characterizations of floating quantifiers:
the ‘movement’ analysis**


Within generative grammar approaches to linguistics, it has been common to analyse floating quantifiers as being transformationally derived from non-floating structures via an operation of constituent displacement/movement. It is posited that quantifiers and their associated NPs are regularly formed as single continuous syntactic units, and then in certain instances separated from each other resulting in discontinuous sequences of quantifier and NP. Such a hypothesis is argued to provide a principled account of various properties of floating quantifier constructions, for example the observation that a quantifier can frequently only occur in a floated position if it can also optionally be positioned adjacent to its NP associate, as illustrated in (5) and (6)

- (5) [both] [the students] → ✓[the students]....[both]...
Both the students left → *The students have both left.*
 [all] [the students] → ✓[the students]....[all]...
All the students left. → *The students have all left.*

Where a quantifier cannot be inserted directly adjacent to an associated NP (without other supporting words), a floating form also seems to be unavailable, as shown in (6), suggesting a systematic relation between floated and non-floated forms that can be described in terms of a movement transformation converting non-floated sequences into forms with NP and quantifier separated from each other.

- (6) *[few] [the students] → *[the students]....[few]...
 **Few the students left.* → **The students have few left.*
 *[some] [the students] → *[the students]....[some]...
 **Some the students left.* → **The students have some left.*

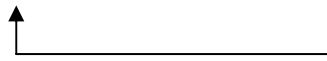
With regard to the question of which of the two elements NP/quantifier is understood to be displaced and moved away from the other, there are two logical possibilities. First, it might be hypothesized that the quantifier floats rightward away from an NP in subject position, as schematized in (7) (with strike-through representing the original, underlying posited position of the quantifier):

- (7) [~~All~~ **the students**] have arrived. → [~~All~~ **the students**] have **all** arrived.
- 

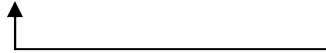
This possibility is commonly rejected as an analysis of floating quantifiers in English for theory-internal reasons, as the hypothesized movement would have to be analyzed as repositioning the quantifier in some structurally lower position in the syntactic configuration. Movement transformations are otherwise uniformly assumed to reposition elements in higher structural positions (which “c-command” the position moved from – Radford 1988, Haegeman 1991). A second possible analysis of quantifier float in languages such as English is the quantifier itself is optionally left behind or “stranded”

when an NP moves from a lower VP-internal position to the regular, surface subject position preceding auxiliary verbs, as represented in (8):

- (8) a. [**All the students**] have [~~all the students~~] arrived.



- b. [**The students**] have [**all** [~~the students~~]] arrived.



As there is other cross-linguistic evidence suggesting that subjects may originate in lower VP-internal positions (Ouhalla 1994), and the hypothesized movement in (8) is to a higher position in the syntactic structure, the analysis of quantifier floating in (8) has been widely adopted in transformational approaches to language. Such an analysis can be used to account for a range of phenomena, for example the observation that a quantifier such as ‘all’ can only occur floating between auxiliary verbs if there has been movement of the object to subject position, as in passive sentences such as (9), and ‘all’ may not simply be inserted in a floating position if no object-to-subject movement has occurred, as in parallel active transitive forms such as (10b). In (9), it is suggested that passive movement of the patient argument of the verb from object to subject position transits through an intermediate position between the auxiliary verbs, where the quantifier ‘all’ can be optionally stranded, giving rise to the floating pattern. This is schematized in representation (11):

- (9) **The criminals** have **all** been arrested.

- (10) a. ✓He has arrested **all the criminals**.
b. *He has **all** arrested **the criminals**.

- (11) [**The criminals**] have [**all** ~~the criminals~~] been arrested [~~all the criminals~~].



The fuller documentation of floating quantifier patterns in English has also noted certain other regular properties constraining the distribution of NP/quantifier separation (Bobaljik 2003). First, although passive structures permit quantifier float, as seen in example (9), other constructions which are regularly analyzed as involving movement transformations in English such as topicalization and relative clause formation do not appear to permit the stranding of quantifiers associated with topicalized and relativized NPs, as shown in (12) and (13):

- (12) a. ***The students**, I took **all** to the show.
(cf. All the students, I took to the show).
b. ***The reports**, I recently **both** responded to.
(cf. Both the reports, I recently responded to.)

- (13) ***The students** (who) I have **all** met today are very nice.
(cf. All the students (who) I have met today are very nice.)

Second, there is a definiteness restriction which applies to the NP in floating quantifier constructions in English – stranding and floating of a quantifier may only occur if its associated NP in subject position is definite (marked by ‘the’/‘those’):

- (14) a. **All/Both participants** have now arrived.
b. ***Participants** have **both/all** now arrived.
c. **The/Those participants** have now **both/all** arrived.

Such a restriction may account for the fact that numerals may not be stranded as floating quantifiers in English, as in cases such as (15) because the associated NP is indefinite:

- (15) a. **Three students** have now arrived.
b. ***Students** have **three** now arrived.

In summary, floating quantifiers in English are assumed to be stranded by movement of an NP to the subject position of a sentence, and NPs raised to such a position must be definite in reference (preceded by ‘the’ or a demonstrative).

3. Burmese

Turning now to consider floating quantifier patterns in the southeast Asian language Burmese (Tibeto-Burman), it can be observed that there are two common differences between quantifier-float/Q-float in English and Burmese, as well as other similarities. The first clear difference between the two languages is that floated quantifiers in Burmese are very frequently *numerals* (coupled with classifiers). As noted above (example 15b), numerals do not occur as floating quantifiers in English. Secondly, the NPs which occur separated from their modifying quantifiers in Burmese frequently do not occur in the subject position of the sentence as in English, and are instead often located in sentence-initial topic positions. Both such properties are illustrated in example (16):³¹

- (16) a. Daw Khin-Khin-kə **khētān ngāa-caun** wε-te *non-floated*
Daw Khin-Khin-NOM pencil 5-CL buy-REAL
‘Daw Khin-Khin bought 5 pencils.’
- b. **khētān** canaw **ngāa-sε-daun** wε-te *floated*
pencil I 50-even buy-REAL
‘I bought as many as 50 pencils.’

Separation of the NP and quantifier elements in Burmese can be naturally analyzed as occurring via stranding of the quantifier when there is movement of the NP to a higher (topic-like) position. The quantifier element is naturally located in the position that the NP would occur in if there were to be no splitting and separation of the NP and quantifier. The alternation in (16) can be schematized as in (17). (16a) and (17a) are neutral S-O-V-Aux

³¹ Here and in subsequent examples, CL stands for classifier, and REAL for realis.

word orders. The sequences in (16b) and (17b) can be analysed as arising when the NP raises out of the neutral object position to the pre-subject topic position, stranding its associated quantifier Q in the object position:

- (17) a. [Subject Daw Khin-Khin-kə] [Object [NP **khētān**] [Q **ngāa-caun**]] wε-tε
 b. [NP **khētān**] [Subject Daw Khin-Khin-kə] [Object [NP **khētān**] [Q **ngāa-caun**]] wε-tε
-

An analysis of movement of the NP stranding the quantifier is supported by other patterns and restrictions in Burmese. A first restriction is that an NP separated from its associated quantifier must structurally “c-command” the quantifier.³² Example (18) below is ungrammatical because the NP *meet-swee* ‘friend’ inside the subject *meet-swee ye seq-bein* ‘friend’s bicycle’ does not c-command the quantifier it relates to. This indicates that floating quantifiers cannot simply be inserted into a sentence in a random way but are subject to clear syntactic restrictions. The ungrammaticality of examples such as (17) can be simply explained if floating quantifiers can only result from the movement of an NP to a position which c-commands the quantifier.

- (17) *[mēet-swee-yε sēqbēin]-kə thōun-yāuq akhōo-khan-ya-tε
 friend-GEN bicycle-NOM 3-CL were-stolen
 Intended interpretation: ‘Three friends’ bicycles were stolen.’

A second, movement-related restriction is that an NP cannot be associated with a floating quantifier that occurs inside a syntactic “island” (constituents such as relative clauses, embedded questions, adjunct clauses – Ross 1967). As syntactic islands regularly disallow movement of a constituent from within the island to a position external to the island, this restriction on NP–quantifier relations in Burmese is again simply explained if the NP moves away from the position of the quantifier in sentences with floating quantifiers. Example (17) illustrates the ungrammaticality of a floating quantifier inside a relative clause island when its associated NP is external to the relative clause:

- (18) ***wiisakii-ayēq** canaw [manε-kə **thōun-palīn** wε tε meinmə-ko] thī-pa-tε
 whisky I yesterday 3-CL bought woman-ACC know
 *‘Whisky, I know the woman who bought three bottles yesterday.’

A third restriction found in Burmese is that a subject NP cannot relate to a quantifier which follows the object in a sentence, as schematized in (19) and illustrated in example (20):

- (19) *NP_{subject-k} NP_{object} Q-k V

³² The structural notion of c-command as it applies in syntactic tree structures is defined as follows: ‘A node X c-commands a node Y if the first branching node dominating X dominates Y, and X itself does not dominate Y.’ (adapted from Radford 1988 p.115).

- (20). ***cāun-thāa(-ka)** htamīn-caw **hna-vāuq** hmāa-tɛ
 student(-NOM) fried-rice 2-CL ordered
 Intended meaning: ‘Two students ordered fried rice.’

The ungrammaticality of examples such as (20) can be accounted for if floating quantifiers arise when there is a movement of the associated NP away from the position of the quantifier. In (20) it can be suggested that the subject NP will never have occupied any position to the right of the object (given the SOV base word order of Burmese), and so will not be able to strand a quantifier in such a position. Such patterns therefore again support a movement and stranding analysis of floating quantifiers in Burmese. Note that other adverbs can occur to the right of the object, as shown in (21), which provides an argument against treating floating quantifiers as simple adverbial elements, and favors the movement-stranding account:

- (21) canaw htamīn-caw khaṅ-khaṅ sāa-tɛ
 I fried-rice often eat
 ‘I often eat fried rice.’

A final restriction which can be noted here is that NPs cannot relate to floating quantifiers located inside postpositional phrases/PPs. Similarly, it is not possible for an NP contained within a PP to relate to a quantifier outside that PP. The generalization is therefore that if an NP and a quantifier occur together inside a PP, there is no way that such elements can be separated from each other. This follows from an analysis in which NPs relate to floating quantifiers via separation and movement of the NP from a position adjacent to the quantifier. Examples (22-24) illustrate such patterns with a range of different postpositions and PPs:

PP_{LOCATION}

- (23) a. turīt-twee-kā [pp[**hotɛ** **thōun-khu**] hmaa] tē-ne-tɛ
 tourist-PL-NOM hotel 3-CL in stay- ASP-REAL
 ‘The tourists were staying in 3 hotels.’
 b. ***hotɛ** turīt-twee-kā [pp**thōun-khu** hmaa] tē-ne-tɛ
 hotel tourist-PL-NOM 3-CL in stay-ASP-REAL
 c. *[pp**hotɛ** hmaa] turīt-twee-kā **thōun-khu** tē-ne-tɛ
 hotel in tourist-PL-NOM 3-CL stay-ASP-REAL

PP_{COMITATIVE}

- (24) a. canaw Suulee-payāa-ko [pp[**pōunjii** **thōun-paa**]nɛ] thwāa-tɛ
 I Sulee-temple-ACC monk 3-CL with go-REAL
 ‘I went to Sulee Temple with 3 monks.’
 b. ***pōunjii** canaw Suulee-payāa-ko [pp[**thōun-paa**] nɛ] thwāa-tɛ
 monk I Sulee-temple-ACC 3-CL with go-REAL
 c. ***pōunjii** **nɛ** canaw Suulee-payāa-ko **thōun-paa** thwāa-tɛ
 monk with I Sulee-temple-ACC 3-CL go-REAL

PP_{BENEFACTIVE}

- (25) a. sayaa-kə [pp]lē-tān-cāunthāa hna-vāuq atwεq] saa
 teacher-NOM 4-year-student 2-CL for letter
 ta-saun yēe-pēe-te
 1-CL write-give-REAL
 ‘The teacher wrote a letter to/for 2 fourth-year students.’
- b. * lē-tān-cāunthāa sayaa-kə [pp][hna-vāuq] atwεq] saa
 4-year-student teacher-NOM 2-CL for letter
 ta-saun yēe-pēe-te
 1-CL write-give-REAL

As PPs are known to be constituents which disallow extraction in many languages, the restriction here can again be suggested to support an analysis of floating quantifiers as being derived by stranding and movement.

Summarizing what has been presented in this section, it can be noted that the distribution of floating quantifiers in Burmese is highly compatible with an analysis of stranding as the result of movement of an NP, which is commonly to a topic-like position, and that floating quantifiers in Burmese are frequently comprised of a numeral and an appropriate classifier.

4. Thai

Standard Thai (Tai-Kadai) is the second southeast Asian language to be considered in detail in this paper. As noted earlier, the internal linear organisation of nominal expressions in Thai is similar to Burmese, with nouns/NPs being followed by numerals/other quantifiers and classifiers: [NP Numeral/Quantifier Classifier]. However, the neutral ordering of clause-level constituents in Thai is significantly different to Burmese, with Thai being an [S Aux V O] language and Burmese having neutral [S O V Aux] order. It is therefore interesting to consider whether issues of basic clausal word order might possibly affect the way floating quantifiers are distributed in the two languages.

4.1 A movement analysis of quantifier float in Thai?

In section 4 above, it was argued that floating quantifiers in Burmese arise via the stranding of a quantifier following movement of an NP to a higher position, as is frequently assumed for English. Shifting our focus to examine Thai, now, a natural comparative question is whether the basic mechanisms of movement and stranding are also responsible for creating floating quantifiers in this language? Consider first the occurrence of floating quantifiers in passive sentences in Thai, such as (26):

- (26) rōt-Mercedes thūuk khəmóy sīisiphāa-khan
 car-Mercedes PASS steal 45-CL
 ‘45 Mercedes were stolen.’

Such examples would appear to be compatible with the assumption that the numeral and classifier sequence *seesiphaa-khan* is stranded by movement of the theme NP *rot-Mercedes* from object to subject position during the derivation of the passive sentence,

which in turn might support the conclusion that Thai floating quantifier constructions are fundamentally similar to those in Burmese and English. However, further data involving floating quantifiers in Thai show that there are clear differences between Thai, Burmese and English which indicate that floating quantifiers in Thai do not have the same syntactic derivation as those in Burmese and English.

A first difference with regard to English is that, despite sharing a similar SVO word order, Thai does not permit the occurrence of floating quantifiers in positions between auxiliary and main verbs, unlike patterns common in English:

- (27) a. **The children** will **all** have arrived by now.
 b. **The children** will have **all** arrived by now.
- (28) ***phuak-dək** aat-ca **thuk-khon** maa lɛɛw
 children may every-CL come ASP
- (29) **The cars** were **all** stolen.
- (30) ***rōt-Mercedes** thùuk **sìisìphàa-khan** khəmóoy
 car-Mercedes PASS 45-CL steal

A second comparative difference is that in English floating quantifiers actually do not occur in post-verbal object position, unlike Thai (as illustrated in (26)):

- (31) ***The cars** were stolen **all**.

Thirdly, and very importantly, when the position of floating quantifiers is considered more closely in Thai, it actually does not correspond to a potential “stranding” position which an associated NP could have moved from. This can be seen in a closer examination of passive sentences with floating quantifiers such as (32) below. Here the floating quantifier does not occur in the object-of-verb position following the verb ‘steal’ *khāmooi*, and is instead located in sentence-final position following the adjunct of location ‘from a factory in Stuttgart’. The floating quantifier is therefore not located in a position from which the associated NP *rot-Mercedes* might have been moved, arguing against a simple movement-and-stranding analysis in Thai:

- (32) **rōt-Mercedes** thùuk khəmóoy cǎak rōong-ngaán
 car-Mercedes PASS steal from factory
 nay mǎəng Stuttgart **sìisìphàa-khan**
 in town Stuttgart 45-CL
 ‘45 Mercedes were stolen from a factory in Stuttgart.’

Floating quantifiers in Thai also occur sentence-finally when relating to subject NPs in (active) sentences where the subject cannot be suggested to have raised from sentence-final position under any standard (transformational) analysis:

- (33) mii phùuyíng maa háa khun khon-nùng
 be woman come find you 1-CL
 ‘A woman came looking for you.’
- (34) nāksùksáa àan böt-nī lēew kùəp thūk-khon
 student read chapter-this ASP almost every-CL
 ‘Almost every student has read this chapter.’

Similarly, floating quantifiers associated with objects can occur clause-finally distanced from the latter in positions which could never be occupied by direct object NPs:

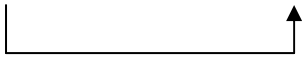
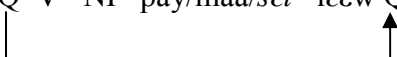
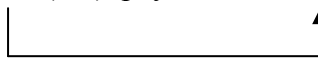
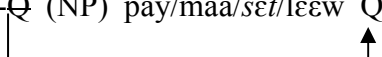
- (35) a. kháw háy ngən kəp phóm sóongrɔ̄y-bàat
 he give money to me 200-Baht
 ‘He gave me 200 Baht.’
- b. *kháw háy kəp phóm ngən
 he give to me money

This occurrence of floating quantifiers in positions that are not possible positions for object NPs can also clearly be seen when the object is separated from the floating quantifier by aspect-marking elements such as *yùu*, *maa*, *pai*, *sət* and *lēew*. In none of the examples in (36-39) below could the object NP occur in the position occupied by the associated floating quantifier.

- (36) phóm mii kaangkeng dii-dii yùu khèe tua-diaw
 I have trouser good-good ASP only CL-single
 ‘I only have one really good pair of trousers.’
- (37) kháw sūə nāngsúə maa sóong-ləm
 he buy book ASP 2-CL
 ‘He bought two books.’
- (38) mùawāannī tamrùat jəp nāksùksáa pay sóong-khon nay təlàat
 yesterday police arrest student ASP 2-CL in market
 ‘Yesterday the police arrested two students in the market.’
- (39) kháw kin khàaw sət sóong-vàang lēew
 he eat food ASP 2-CL ASP
 ‘He finished eating two of the dishes.’


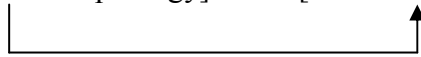
The general observation resulting from such data is that floating quantifiers in Thai very frequently occur in positions which their associated NPs could not have previously occupied or have been moved from under any standard transformational analysis incorporating the notion of syntactic movement/displacement. This results in the conclusion that floating quantifiers in Thai, unlike English and Burmese, cannot be analysed as resulting from stranding following the movement/displacement of an associated NP.

This being so, the question remains as to how floating quantifiers in Thai may occur in different types of positions to those in English and Burmese? If it is supposed that some form of movement/displacement operation is still potentially involved in separating quantifiers from NPs in Thai, as assumed in other languages, the question then becomes what kind of rather different movement/displacement operation could apply to create the structures found in Thai? Arguably, the only obvious way to analyse the data in Thai in terms of movement is to assume that in Thai separation of NPs and their quantifiers is achieved by movement of the quantifier-element rather than movement of the NP, and that the quantifier-element is displaced to some rightward position in the clause, most commonly following other argument NPs and the aspect markers *pay/maa/sèt/yüu* and either preceding or following the aspect marker *l̄εεw*, as schematized in (40):

- (40) a. NP-Q V NP *pay/maa/sèt* Q *l̄εεw*

 b. NP-Q V NP *pay/maa/sèt* *l̄εεw* Q

 c. NP V NP-Q (NP) *pay/maa/sèt* Q *l̄εεw*

 d. NP V NP-Q (NP) *pay/maa/sèt/l̄εεw* Q


The movement hypothesized above is quite different from the movement assumed in English and Burmese floating quantifier constructions. In English and Burmese the movement of the NP is leftward and can be suggested to occur for reasons of case (English) or topicalization (Burmese). In Thai the movement of the quantifier is to the right, and for (as yet) unclear reasons/motivations.

Such a working hypothesis naturally leads to the question of whether cross-linguistically it is possible to identify other instances of ‘rightward movement’ which might support a rightward movement analysis of floating quantifiers in Thai? The answer to this question is certainly ‘yes’, with instances of extraposition and ‘Heavy NP Shift’ being two common occurrences of the apparent displacement of syntactic constituents to the right of a clause, as illustrated with English (41) and (42) below:³³

- (41) Someone [~~who would change our lives forever~~] then entered the room [~~who would~~ change our lives forever].

 (42) John recently sent [~~a book about Polish morphology~~] to me [a book about Polish morphology].


³³ Heavy NP Shift characteristically involves the rightward displacement of heavy/long objects in double object constructions, as in (42). Instances of extraposition include (but are not limited to) the rightward displacement of PPs and relative clauses from subjects in English, as in (41).

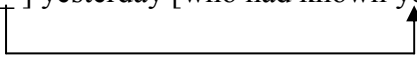
Note that similar to NP/floating quantifier pairs, in cases of extraposition such as (41) there is a syntactic constituent (a relative clause) which would normally occur attached to an NP, but is here separated from that NP. Furthermore, as with Thai floating quantifier structures, the NP in sentences with extraposition such as (41) cannot itself occur in the position where the associated relative clause is found (as in: *Then entered the room someone.’), forcing the conclusion that the relative clause has not been stranded by movement of the NP from some clause-final position but has been moved rightwards away from the NP in subject position.

A further restriction found to characterize occurrences of extraposition is that the NP which relates to an extraposed relative clause must be interpreted as non-specific and indefinite.³⁴ The NPs which are associated with floating quantifiers also commonly have the property of being non-specific, which suggests a further parallel between extraposition and Thai rightward floating quantifiers. In analyses of extraposition, the specificity constraint is understood to be a restriction on movement and extraction (barring extraction from specific NPs). If a similar restriction holds of NP-floating quantifier relations, this may add further support for a movement analysis of floating quantifiers in Thai. Additionally, and potentially relating to the issue of non-specificity, functionally the rightward extraposition of PPs and relative clauses in English is commonly used when an NP is introduced for the first time into the action described in a discourse situation, as for example in (43) and (44):

(43) [A review _] appeared in the Times [of a new book about Roosevelt].



(44) I met [a man _] yesterday [who had known your father in the 1960s]



As will later be discussed, floating quantifiers in Thai are also used frequently in presentational situations, increasing the parallels between extraposition and floating quantifier constructions.

Finally, the hypothesis that rightward syntactic *movement* is involved in the distribution of floating quantifiers in Thai is supported by the interaction of quantifier float and standard configurational restrictions on movement, i.e. ‘island phenomena’. As in Burmese, there are patterns indicating that the relation between an NP and a floating quantifier in Thai is regularly restricted by the occurrence of ‘island’ constituents, and it is not possible for an NP inside a constituent such as a relative clause, adjunct clause or other island type to relate to a floating quantifier located outside such a constituent in clause-final position, as illustrated in (45). This example is only acceptable with the continuation in (a) in which the floating quantifier is associated with the noun ‘man’ which is external to the relative clause island, and is not acceptable if relating to the noun ‘Rolls Royce’ inside the relative clause as in (b) (where the switch in classifier from the classifier for people *khon* to the classifier for vehicles *khan* makes the intended meaning clear, though unacceptable):

³⁴ An NP is non-specific if its identity is unknown by both the hearer and the speaker prior to the action described in the sentence.

- (45) phómkhəy jəə [phùuchaaythìi mii rōt Rolls-Royce maa lēəw]
 I ASP meet man REL have car Rolls-Royce ASP ASP
- a. ✓thúng sǐp-kwàa-khon
 as-many-as 10-above-CL_[people]
- b. ...*??thúng sǐp-kwàa-khan_[vehicles]
 as-many-as 10-above CL

ONLY: a. 'I have met more than 10 men who have owned a Rolls Royce.'

NOT: b. 'I have met a man who has owned more than 10 Rolls Royce cars.'

If the sensitivity of a syntactic dependency to island phenomena is an indication that such a dependency is the result of movement, then the presence of island restrictions in floating quantifier constructions in Thai clearly suggests that movement of the floating quantifier occurs, separating it from the NP.

4.2 Consideration of a non-movement analysis: floating quantifiers as adverbs

While the distribution of floating quantifiers in Thai seems highly compatible with a rightward movement analysis similar to extraposition, this being supported by island phenomena and other parallels with extraposition, one might also consider an alternative potential analysis of floating quantifiers, that they are not moved to their surface position from any other underlying location but are inserted directly into their clause-final position as (or, rather, like) adverbs.³⁵ Such a non-movement approach to floating quantifiers might be supported by the following range of observations.

First, VP-level adverbs occur in similar clause-final positions to those occupied by floating quantifiers in Thai:

³⁵ A reviewer of the paper notes that there are patterns in both Thai and Burmese which might suggest that even when a numeral-classifier pair are adjacent to a noun/NP, the numeral-classifier pair might not be part of the DP nominal constituent, hence might always be adverbial in nature. Specifically, it is observed that pronouns in Thai and Burmese can be followed by numeral-classifier pairs, as illustrated in (i) and (ii). If it were to be assumed that the pronoun replaces the full DP, one might have to conclude that numeral-classifier pairs are, at least in some instances of adjacency, not necessarily part of a DP:

- (i) kháw sóng-khon (Thai) (ii) thu-tō hna-yāuq (Burmese)
 they 2-CL they 2-CL
 'the two of them' 'the two of them'

Parallel forms also occur in Mandarin Chinese, as shown in (iii) from Li (1999):

- (iii) tāmen liang-ge (Mandarin)
 they 2-CL
 'the two of them'

Li (1999:83) argues convincingly that the pronoun in such sequences is in the D⁰ position (actually a fairly common syntactic assumption about pronominal elements), and that the numeral and classifier occur in the same post-D⁰ DP-internal positions that they occupy in other instances. There is consequently no need to conclude that the post-pronominal occurrence of numeral classifier pairs in examples such as (i-iii) indicates that such elements are DP-external and adverbial in nature. Furthermore, Simpson (2005) shows that elements which close off DPs, such as case-markers in Burmese, and demonstratives in Thai, can occur following post-nominal numeral-classifier pairs, confirming that such sequences are DP-internal elements, at least in certain occurrences, and hence would not be naturally analyzed as adverbial elements in all instances.

- (46) *kháw tǝp khamtháam yǎangchalàat*
 he answer question cleverly
 ‘He answered the question cleverly.’

Second, there are modifiers constructed from classifier bases which do occur in adverb-like ways in clause-final position. In examples (47-48), the bolded, underlined elements containing classifiers are not floating quantifiers as they either do not contain a quantifier or do not relate to any overt NP in the sentence, and yet they are licensed to occur in similar positions to other floating quantifiers.

- (47) a. *khǎng-thiilǝw kháw maa [khon-diaw]*
 time-last he come CL-single
 ‘Last time he came alone.’

- b. *khóo pay dùay [khon] nǎ*
 request go together CL PRT
 ‘Can I go too?’

- (48) *lǝn klǎay pen thiiprǔksáa khóong kháw pay thǔk-rǔəng lǝw*
 she change be advisor of he ASP every-CL ASP
 ‘She became his advisor in everything.’

Third, there may be some final floating quantifiers which can *not* occur adjacent to the NP they modify, hence which do not alternate with a non-floated form. For example, in (49) below, it is not possible for the floating quantifier sequence *sám-sǝp* to occur adjacent to its NP associate *phùak-nǎn* as shown in (49b). The absence of such an adjacent NP quantifier sequence may suggest that the floated form is not derived from an NP quantifier unit via movement of the quantifier.³⁶

- (49) a. *kháw yǐng phùak-nǎn taay lǝw sám-sǝp*
 they shoot group-that die ASP 3-CL
 ‘They shot three of them dead.’

- b. *??*kháw yǐng phùak-nǎn sám-sǝp taay lǝw*
 they shoot group-that 3-CL die ASP

It can also be noted that Thai floating quantifier patterns are different to those in Burmese (and Japanese, Korean; Kang 2002) in the patterning of sentences with prepositional phrases/Ps. In Thai it appears to be quite possible for a clause-final floating quantifier to relate to an NP located inside a PP. In other languages such as Korean, Japanese and Bengali (Simpson and Bhattacharya 2008) where floating quantifiers may not relate to an NP inside a PP this has been suggested to be because PPs in many languages may disallow extraction/movement. If this is the correct interpretation of PP-related data,

³⁶ The complication in this example is that *sǝp* is the classifier for dead bodies and only seems to allow for use in counting corpses once it has been established that the relevant people are dead – hence in floated clause-final position after ‘died’ *taay lǝw*.

it might weaken the case for a movement analysis of floating quantifiers in Thai. Examples (50-52) show that quantifier units floated in clause-final position are free to associate with NPs contained in a range of PP types:

- (50) *kháw kèp ngøn [ppcàak nāksùksáa] maa lēew kùəp thūk-khon*
 he collect money from student ASP ASP almost every-CL
 ‘He has collected money from almost all of the students.’
- (51) *chán khuy [ppkăp khèək] maa lēew sóong-sám-khon*
 I chat with guest ASP ASP 2-3-CL
 ‘I’ve talked with about 2 or 3 guests already.’
- (52) *kháw sóon-nāngsúu [ppthii maháawitavalaay] maa lēew s̄i-kwàa hēeng*
 he teach in university ASP ASP 4-over-CL
 ‘He has taught in more than 4 universities.’

4.3 The partitivity issue

Data relating to interpretations of partitivity provide further potential clues as to the derivation of floating quantifier structures. Sentences with floating quantifiers sometimes have different interpretations from those where numeral-classifier pairs are not floated. Consider examples (53) and (54), where quantifier float occurs in (54) (but not (53)):

- (53) *dèk sám-khon taay lēew*
 child 3-CL die ASP
 ‘The three children have died.’
- (54) *dèk taay (pay) sám-khon lēew (or: dèk taay lēew sám-khon)*
 child die ASP 3-CL ASP child die ASP 3-CL
 ‘Three of the children have died (so far/already)’

(53) is commonly described (by speakers of Thai) as having the meaning that some group of three children known to the speaker and hearer (i.e. a definite group of three children) had died. (54), by way of contrast, is suggested to mean that three children from some group known to the speaker and hearer have died, and to imply that there are still other children from that group who may be in danger. This corresponds to a “partitive” interpretation in which the numeral+classifier quantifies over a definite set.

A similar difference in interpretation is found to occur when the NP is located in object position as in (55) and (56) (quantifier float takes place in (56)):

- (55) *kháw kin kàpkhàaw sóong-vàang s̄t lēew*
 he eat dishes 2-CL ASP ASP
 ‘He has finished eating the two dishes.’
- (56) *kháw kin kàpkhàaw s̄t sóong-vàang lēew*
 he eat dishes ASP 2-CL ASP
 ‘He has finished eating two of the dishes.’

(55) is characterized as meaning that there are only two dishes in total (and they are now consumed), whereas (56) implies that there is more food on the table and has the partitive-like interpretation ‘two of the dishes’.

The potential significance of this data is the following. If clause-final floating quantifiers are supposed to be derived by movement of a numeral+classifier from a position adjacent to the NP, it might not be expected that this movement would affect the meaning/interpretation of the numeral+classifier in such a clear way. In other words, why would a partitive interpretation be present with floating quantifiers but not non-floated numeral+classifier sequences if the former are simply derived from the latter? This difference in interpretation could be taken as an argument against analyzing floating quantifiers as the result of a movement transformation, as syntactic movement is regularly understood to preserve rather than alter meaning.

In trying to make sense of the alternations here, it is useful to reflect on how partitive interpretations generally may arise from syntactic structures. In English and many other languages, partitive interpretations occur when a numeral (or a universal quantifier such as ‘all’) is positioned external to the “DP” unit created by the addition of a determiner to an NP, as represented in (57):

(57) [_{QP} two/all of [_{DP} the [_{NP} students]]]

This contrasts with the interpretation which arises when a numeral occurs inside the DP, following ‘the’:

(58) [_{DP} the two students]

The sequence in (57) implies there are more students who are part of a group familiar to both speaker and hearer, while (58) refers to a group composed of just two students known to speaker and hearer, and there is no implication that other students belong to this particular group. This allows for the statement of a simple partitive generalization relating to the syntactic structuring of partitive phrases:

(59) PARTITIVE GENERALIZATION

A numeral which is external to and quantifies over a definite DP/NP gives rise to a partitive interpretation.

A numeral which is internal to a definite DP/NP does not give rise to a partitive interpretation.

Now, if adjacent sequences of NP + numeral-classifier in Thai such as *dek saam-khon* (ex. 53) and *kap-khaaw soong-yaang* (ex. 55) do not give rise to partitive interpretations, this suggests that the numeral+classifier in such sequences has to be interpreted as being ‘internal’ to any definite DP and that Thai does not have a second possible position for numeral+classifiers equivalent to the position of the ‘outer’, DP-external numeral in English (57). Linearly adjacent NP + numeral-classifier sequences would therefore be assumed to always have the structure [_{DP} NP quantifier.classifier] and not the structure [[_{DP} NP] quantifier classifier] (which would be expected to license partitive interpretations). Finally, as movement operations are assumed to conserve

fundamental aspects of meaning, if Thai only makes available a DP-internal non-partitive position for numeral+classifier pairs, it might be concluded that clause-final floating quantifiers associated with partitive interpretations cannot be moved from positions adjacent to the NPs they modify.

However, the fuller patterning of partitivity with nominal expressions is actually more complex still than the patterns seen above, and there is clear evidence that Thai in fact allows for two different positions of NP-adjacent numeral+classifier pairs, one of which appears to be an ‘outer’ position which can give rise to partitive readings. This is seen in examples (59) and (60) when possessor-phrases and relative clauses co-occur with numeral-classifier pairs in two alternating orders. In the first (a) order, the numeral and classifier are directly adjacent to the noun and there is no partitive interpretation, whereas in the second (b) order the numeral-classifier pair occurs further to the right, separated from the noun by the possessor phrase/relative clause, and partitive interpretations naturally arise:

- (60) a. [bàan **sáam-láng** khóng phóm]
house 3-CL of I
‘my three houses’
- b. [bàan khóng phóm **sáam-láng**]
house of I 3-CL
‘three of my houses’
- (61) a. [bàan **sóng-láng** [thìi phóm sū# nai Amerikaa]]
house 2-CL which I buy in America
‘(the) two houses which I bought in America’
- b. [bàan [thìi phóm sū# nai Amerikaa] **sóng-láng**]
house which I buy in America 2-CL
‘two (of the) houses which I bought in America’

If the (b) forms of show that there is a second ‘outer’ position for numeral+classifier pairs in nominal expressions and this position can give rise to partitive interpretations, it should clearly be possible for a clause-final floating quantifier to move from such an outer position, maintaining the partitive interpretation which is made available by its initial occurrence in the outer position. Considered further, therefore, the availability of partitive interpretations with floated numeral+classifier pairs ultimately does not provide a clear argument against the assumption that they are moved to their surface position from a position adjacent to the NP associate.

What does still require some explanation, however, is why partitive interpretations of linearly adjacent NP numeral+classifier sequences do not appear not to be possible, i.e. why (53) and (55) do not seem to easily allow a partitive reading. If there is indeed a second structural NP-adjacent position for numeral+classifier pairs which will allow for a partitive interpretation (given the patterns in (60b) and (61b)), why is it the case that the numeral+classifier pairs in (53) and (55) cannot be interpreted in such a position, giving rise to partitive meanings? A possible explanation for the lack of a (now) expected

partitive interpretation in (53) and (55) may be to attribute this to parsing preferences and the cross-linguistic phenomenon of ‘Local Association’ (aka ‘Late Closure’ Frazier 1979). Local Association is a preference in parsing to keep adjacent words analyzed as being close together in the underlying syntactic structure constructed by hearers, and such a preference principle may restrict the analyses and associated interpretations that hearers find it easy to mentally construct. For example, both of the English sentences in (62) and (63) below are ambiguous, but Local Association strongly leads hearers to make the interpretation in (a), as this involves mentally analysing the adverb ‘yesterday’ and the relative clause ‘who was on the balcony’ as modifying the nearest available unit:

(62) John said that Mary left yesterday.

(a) strong parsing preference: ‘yesterday’ modifies ‘left’

(b) less naturally available: ‘yesterday’ modifies ‘said’

(63) Someone shot the servant of the actress who was on the balcony.

(a) strong parsing preference: ‘who was on the balcony’ modifies ‘the actress’

(b) less naturally available: ‘who was on the balcony’ modifies ‘the servant’

Applied to the apparent lack of a natural partitive interpretation in examples such as (53) and (55), it could be suggested that Local Association enforces the parsing attachment of the numeral+classifier in the first mentally available position - the syntactically lower internal position of numeral+classifier pairs - and that hearers are only able to make a higher attachment (in the outer position) when there is material intervening between the NP and the numeral classifier as in (60b) and (61b), which enforces high attachment of the numeral+classifier in the outer position and the ensuing partitive interpretation.

4.4 Mid-way conclusions and partial summary

Although it has been seen that there are two potential analyses of the derivation of clause-final floating quantifiers in Thai, it can be suggested that the balance of the evidence may favor an analysis of movement, with repositioning of floating quantifiers from a position adjacent to an associated NP to some clause/VP-final location. While an adverbial analysis of floating quantifiers is theoretically possible, it is perhaps not strongly supported, and is clearly challenged by the occurrence of island restrictions. Moving forward with such assumptions, we can now partially summarize the differences which seem to exist in floating quantifier patterns in Burmese, Thai and English, as established by the answers to five questions probing significant parameters of variation in floating quantifier phenomena.

QUESTION 1: *Where does the NP associated with the floating quantifier occur?*

In English, the NP occurs in subject position. In Thai, the NP associate occurs in regular subject, object, indirect object, or object of preposition positions (hence the NP itself is not moved to any special position). In Burmese, the NP commonly appears in a topic-like position and often seems to have been displaced from a regular subject or object position.

QUESTION 2: *Where does the floating quantifier/FQ occur?*

In English, the FQ commonly occurs between auxiliary verbs and the main verb, before the VP. In Burmese, FQs commonly occur in the regular position of the associated NP (e.g. in

regular subject or object position). In Thai, FQs occur in clause/VP-final positions which often do not correspond to the regular position of the associated NP, nor any position which the NP would have moved through.

QUESTION 3: *What is the definite/indefinite status of the “NP”?*

In English, the “NP” has to be definite (and is therefore actually a DP; e.g. ‘the students’, and cannot be a bare indefinite NP e.g. ‘students’). In Burmese, the NP is commonly a bare indefinite NP (e.g. *khetan* ‘pencil’). In Thai, the NP is often a bare indefinite NP (e.g. *phuuying* ‘woman’), but can also be more complex and definite (e.g. *pheuan khoong phom* ‘my friend(s)’).

QUESTION 4: *How does the floating quantifier structure seem to be syntactically created?*

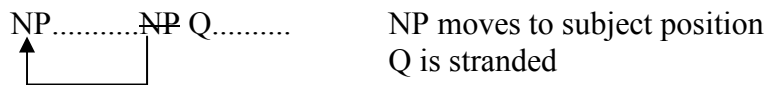
In English, stranding appears to occur - the NP leaves behind the FQ when it moves to a higher position (the subject position). In Burmese, stranding also appears to occur - the NP leaves behind the FQ when it moves to a higher topic position. In Thai, it may appear that rightward movement takes place - the FQ appears to be moved away from the NP to a position to the right of the clause, similar to extraposition movement in English.

QUESTION 5: *What elements occur as floating quantifiers?*

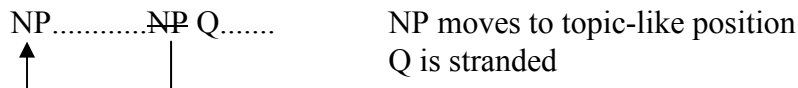
In English, only quantifiers such as ‘all’ and ‘both’, and not numerals occur as FQs. In Burmese, numerals (combined with classifiers) and other quantifiers (not combined with classifiers, e.g. ‘aa-loun’ ‘all’) appear as FQs. In Thai, numerals (combined with classifiers) and other quantifiers combined with classifiers (e.g. *thuk* ‘every’, *laai* ‘several’, *baang* ‘some’) function as FQs.

Generally, then, floating quantifier patterns are not fully uniform across different languages and may vary according to a range of distinct properties. The hypotheses developed here concerning the syntactic mechanisms which result in separation of NP and FQ (i.e. question 4) are schematized below in (64). In section 5, we go on to consider what may functionally be responsible for this separation of NP and FQ.

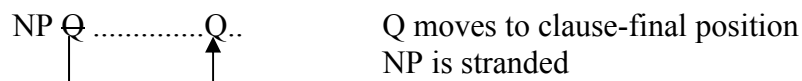
(64) a. ENGLISH



b. BURMESE



c. THAI



5. Why does separation of the NP and quantifier take place?

Having considered some of the structural properties and differences in floating quantifier constructions in Burmese and Thai (and English), we should now ask why this kind of splitting and separation of NP and quantifier ever occurs. The splitting of a constituent into two separate, distanced parts imposes considerable extra processing costs, as the reference value of a subject or object or other event participant can only be computed fully once both NP and quantifier are encountered in the processing of a sentence, and in instances where there is separation of an NP and an associated quantifier, hearers have to mentally ‘store’ the partial information provided by the NP until it can be combined with that of the associated quantifier and a full referential value for the subject/object etc can be arrived at. Given such extra processing costs, it is natural to ask what benefits may result from the use of floating quantifiers as compensation for the processing burden. In addition to such a functional question, one might also wonder whether there is any principled way to explain the syntactic differences observed above, or does it have to be concluded that the cross-linguistic variation found in floating quantifier constructions is simply random and fully unpredictable? In the remainder of the paper, it will be suggested that, in employing floating quantifier constructions, languages may frequently be trying to achieve the same basic effects in terms of information structure, but are constrained by differently configured local resources, and it is language-specific properties and restrictions which result in the range of differences noted above. FQ constructions will therefore be suggested to potentially contain elements of the universal interacting with elements of the language-specific, the universal here being the linear ordering of elements in information structure, and the language-specific being variation in basic word order type: V-O vs. O-V (i.e. head-initial vs. head-final). As a way to approach these issues, we will begin by considering how and when floating quantifier constructions are commonly made use of.

5.1 The functional use of FQs in Thai

In Thai, there are two particularly common contexts which give rise to the use of FQ constructions. The first of these is presentational sentences - floating quantifiers frequently occur when new referents are being introduced in a discourse, often with the existential verb *mii* or as the object of a verb, as illustrated in (65-66):

- (65) *mii* **dèk** *maa ngaanpaatii* **raw** **sìisìp-kwàa-khon**
 be child come party about 40-above-CL
 ‘More than forty children/young people came to the party.’

- (66) *phómphòng* *sū#* **nángsú#** *maa* **sóong-lè#m**
 I just buy book ASP 2-CL
 ‘I just bought two books.’

The second context where FQs occur with significant regularity in Thai is in instances of ‘re-presentation’ and partitivity. Splitting and separation of NP and quantifier occurs in instances where the NP is not new information – the NP is definite in reference and already familiar to hearer and speaker – and splitting results in partitive interpretations with a frequent focus on what characterizes a certain number of the set represented by the NP, as seen in (67-68):

- (67) lùuk khóong phùæn phóm taay lɛɛw sóong khon
 child of friend I die ASP 2-CL
 ‘Two of my friend’s children have died.’

- (68) bangalo kò wàang yùu sóong-sáam-láng
 bungalow PRT vacant ASP 2/3-CL
 ‘Two or three of the bungalows are free.’

The main difference between the two common contexts for FQs cases is the referential familiarity of the NP - the NP is either old-familiar information being re-presented in a sentence, or is new information being presented for the first time. A generalization which unites the two contexts is that in both cases the quantifier itself represents important, new information.

- (69) GENERALIZATION ONE (Thai)
 The quantifier in FQ constructions represents important, new information.

Because of (69), it is quite unnatural for demonstratives to occur floated in final position, as demonstratives regularly do not encode new information:

- (70) kháw sã# nángsú# maa ✓ sóong-lè#/*??lè#-nī lɛɛw
 he buy book ASP 2-CL/CL-this ASP
 ‘He has bought two books/this book.’

The unacceptability of ‘floating demonstratives’ here is similar to the unnaturalness of repositioning a demonstrative-marked NP to the right of a clause in English Heavy NP Shift constructions:

- (71) I gave _ to Mary ✓[a book about elves]/*??[this book].

A second functional generalization which characterizes FQ patterns in Thai is that use of floating quantifiers is often felt to sound more natural when the quantifier is accompanied by some other qualifying/focus particle such as the following: *khæ* ‘only’, *tang* ‘as many as’, *thawnan* ‘only’, *keuap* ‘almost’, *raaw* ‘approximately’. FQs are also judged to be natural-sounding when the numeral which occurs has a remarkable or high value:

- (72) mii khon maa tàng-hàa-sìp-khon
 be people come as-many-as-50-CL
 ‘As many as 50 people came.’

- (73) GENERALIZATION TWO (Thai)
 FQ constructions often involve the occurrence of an additional focus or qualifying particle or remarkable/high-valued numerals.

These two observations in (69) and (73) support the view that floating quantifiers instantiate *focused information*. It can therefore be suggested that when there is natural pressure to stress the focal salience of new information represented by a quantifier, this may be achieved by positioning the quantifier away from its associated NP in clause-final position, where new information is most naturally positioned in a very wide range of languages.³⁷

Continuing to examine such a functional view of floating quantifiers, in Thai in the frequent instances where there is presentation of a fully new quantified referent, it can be noted that there are actually two pieces of new, important information which occur in floating quantifier structures: (a) the noun/NP - i.e. the identity of the type of the referent: 'student', 'book' etc, and (b) the quantifier - the amount of the N: 'two books', 'fifty students' etc. In such instances it can be hypothesized that the splitting and separation of a constituent into two components (NP and FQ) may serve to highlight the two, separate parts of the constituent, and splitting may be a particularly useful solution, where syntactically available, to situations in which there are two pieces of adjacent information both of which are informationally new and in focus. If the two components are separated and certain linear space is created between them, this may potentially serve to increase the salience/prominence of both items. Elsewhere in studies of language there is much evidence that perceptual salience may be at a maximum at the beginnings and ends (the edges) of units. For example, it is known that the beginnings and ends of words and syllables are perceptually more salient than the internal parts of such units. In a similar way, it can be suggested that the splitting of an adjacent NP Q sequence (in which both NP and Q are new information) into a spatially separated NP.....Q order may serve to create a structure in which the perceptual salience of both parts is usefully heightened. In this regard, there may be similarities with other common splitting/separation constructions. As noted earlier, extraposition structures such as (74) are naturally used in presentational situations, where a new referent is introduced. Here the noun 'man' encodes information about the basic type of the new referent, and the extraposed relative clause adds further new information about this basic type. In such cases, splitting of the NP into two parts may functionally occur to enhance focal salience on both pieces of new information:

(74) **A man** entered the room **who was wearing a black hat**.

A key property of separation and splitting may therefore be to establish a certain distance between two new units of information, both of which are in need of emphasis. Whereas some languages have considerably flexibility in stress placement and the manipulation of stress for informational purposes, other languages (in particular tone

³⁷ In many languages, the cross-linguistic tendency for new information to be introduced in sentence/clause-final location causes the occurrence of non-canonical word order patterns, sometimes with the re-ordering of major argument constituents. For example, the neutral SV(O) word order in languages such as Italian is regularly reordered as VS if the subject encodes new information, as in (i):

(i) È arrivato Gianni.
has arrived Gianni

'Gianni has arrived' (a natural answer to the question: 'Who has arrived?')

The use of FQ constructions can therefore be seen as another manifestation of non-canonical word order to highlight new information in clause-final position.

languages, such as Thai and Burmese) have less flexibility, and may need to make use of special syntactic structures and movement/repositioning of elements to achieve similar functional ends. The occurrence of split, floating quantifier structures may consequently be the result of situations in which adjacent focal elements cannot both be naturally stressed and so constituents are split in two to allow for both parts to maximize their focal prominence.

5.2 The functional use of FQs in Burmese

When Burmese floating quantifier constructions are considered from a pragmatic, functional point of view, they are interestingly found to show similar focus properties to those observable in Thai, and floating quantifiers in Burmese occur very naturally with focus-type particles (e.g. *-taun* ‘as many as’). Indeed, various configurations involving floating quantifiers which speakers categorize as unacceptable/highly unnatural or even ungrammatical can be ‘rescued’ and made perfectly acceptable by the appropriate use of focus particles. This is an observation which has also been made about similar patterns in Japanese and Korean floating quantifier constructions (Kang 2002, Miyagawa and Arikawa 2007). In various earlier works on Japanese and Korean (e.g. Miyagawa 1989) the linear sequencing of a floating quantifier associated with a subject but following an object, as schematized in (75) and illustrated with Korean (76) was categorized as ungrammatical. However, if a focus particle is added to the floating quantifier, and/or a numeral quantifier is made into a large ‘remarkable’ number (hence inherently focused) as in (77), it has been noticed that the configuration in fact becomes perfectly acceptable (Kang 2002, Miyagawa and Arikawa 2007):

(75) NP_{Subject-k} NP_{Object} Q_{-k} V

(76) ***hakseyng-i** khempywuthe-lul **twu-myeng** sassta
 student-NOM computer-ACC 2-CL bought
 ‘2 students bought a computer.’ (Kang 2002)

(77) **hakseyng-i** khempywuthe-lul **twu-myeng-ina/-man** sassta
 student-NOM computer-ACC 2-CL-as-many-as/only bought
 ‘As many as/only two students bought computers.’ (Kang 2002)

Alternatively, if it is ensured that the object NP in sequences such as (75) is not interpreted as new information (which might potentially distract attention away from the intended focus on the new information of the floated quantifier), such a strategy will also ‘save’ structures with the form in (75). Again, this confirms the required focal properties of floating quantifiers. While (78) is regularly judged as deviant in Japanese, if the object *sake-o* ‘wine’ is pronominalized as *sore-o* ‘that’ and so encodes old/given information as in (79), the sentence is accepted as well-formed and natural (Nakanishi 2008):

(78) ?*gakusei-ga sake-o san-nin non-da
 student-NOM wine-ACC 3-CL drank
 ‘3 students drank wine.’

- (79) kinoo-wa gakusei-ga sore-o san-nin non-da
 yesterday-TOP student-NOM that-ACC 3-CL drank
 ‘3 students drank it.’

Similar patterns occur in Burmese and reinforce the assumption that floating quantifiers are focused information in Burmese, as in Thai, and occur in final pre-verbal position in order to heighten their focal prominence, the pre-verbal position in Burmese being the position that other focused elements naturally occur in, as in many SOV type languages (e.g. Turkish, Hindi, Bangla).

Although the focused interpretation of floating quantifiers can thus be characterized as similar in Burmese and Thai (and Japanese and Korean), it can be noted that there is also a difference in the interpretation of the associate NP which frequently occurs in Burmese but not in Thai. Speakers of Burmese often note that there seems to be a natural sense of *contrast* implied in many cases of splitting and separation of the NP and a quantifier, and the implication of ‘lists’ in which items are compared and contrasted against each other. This is illustrated in (80). When presented with such sentences, speakers report that there is a natural implication that the subject also bought (different) quantities of other items too

- (80) dāqhkēh-ko candaw lēe-lōun wε-de
 battery-ACC I 4-CL buy-REAL
 ‘I bought batteries.’
 implication: I bought different quantities of other items too.

The interpretation of the NP in sentences such as (80) is therefore that of a *contrastive topic*. Contrastive topics are frequently both old and new in informational terms: their identity is generally known/familiar, but there is new information present in the fact that they are contrasted with other members of a particular set, as illustrated in (81) (Lee 1999):

- (81) [[These]_{Focus} examples]_{Topic} I found [in Gundel].
 ‘these’=focal/contrastive
 ‘these examples’=old/known information

Burmese separation of NPs and associated quantifiers therefore involves both new information focus on the pre-verbal floating quantifier and frequent contrastive topic-like interpretation of the NP. In this patterning, floating quantifier separation constructions are similar to splitting constructions in languages such as German (also Polish, Russian). In German (Fanselow and Cavar 2001), the two parts of a single NP unit can be split apart as in (82),

- (82) Autos besitzt er (nur) schnelle.
 cars owns he only fast
 ‘He owns only fast cars.’
 ‘As far as cars are concerned, he only has fast ones. As for motorcycles,...’

The same kind of contrastive interpretations that are often felt in Burmese floating quantifier constructions are common in such splitting, and may be a frequent property of many splitting constructions. However, they do not seem to be a common interpretation in Thai floating quantifier constructions, and this accordingly is an instance of some difference in the patterning of floating quantifiers in the two languages (to be returned to below).

Reflecting on the commonalities found with floating quantifiers in Thai and Burmese, and the functional question of why splitting of NPs and quantifiers occurs in the two languages, a general conclusion which it seems plausible to adopt is that splitting takes place in order to focus the quantifier and its new information in a prominent, final focus position as stated in (83).

- (83) Functional generalization on FQ constructions in Thai and Burmese:
Splitting and distancing of NP and quantifiers coincides with and is appropriate for the encoding of (new information) focus on the quantifier.

We have now attempted to provide at least partial answers to the three questions in (84) we began this paper with, with special reference to Burmese and Thai, and comparisons made with patterns already reported in English, Japanese and Korean:

- (84) i. How are the NP and quantifier related to each other in floating quantifier constructions?
ii. Why does separation of the NP and quantifier occur?
iii. Are floating quantifier constructions cross-linguistically uniform?


The answer to question (84iii) has been that there are in fact a number of differences in floating quantifier patterns across languages, even when one considers just the two languages Burmese and Thai. For example, it appears that in FQ constructions in Thai the quantifier is repositioned to the right, stranding its associated NP in situ, while in Burmese the NP is moved to the left, stranding the associated quantifier. An interesting question which we can now ask is whether such differences might in any way be predicted or accounted for by the answer to the other questions (i) and (ii)? The answer here may be ‘yes, quite possibly so’. Specifically, it will be suggested that differences such as those observed with Thai and Burmese may result from the interaction of ‘universal’ and language-specific properties. The ‘universal’ property relevant here is the observation drawn from general studies of information structure that there is a pervasive cross-linguistic tendency for elements representing new information to occur focused in clause-final positions (hence new referents are commonly introduced in object rather than subject positions). The language specific property we will consider here is the difference in basic, neutral word order in Burmese and Thai.

First, looking at patterns in Thai, let us consider the case of a subject which is represented by an NP and a quantifier whose content the speaker wishes to focus. Due to the basic SVO word order in Thai, if the quantifier remains adjacent to the NP, this will result in the linear sequencing in (85):

(85) [NP Q_{focused}]_{Subject} V NP_{Object}

In this neutral word order, the quantifier which is to be focused is located far away from the clause-final position which new information focus most naturally occurs in. In order for the quantifier to occur in such a position, it is regularly moved/relocated to its right, stranding the NP, as in (86):

(86) [NP Q_{focused}]_{Subject} V NP_{Object} Q



The stranded subject NP is not in a position which has a special focal status in the information structure of the sentence and simply receives the regular interpretation of an NP in subject position, hence there is no necessary/common interpretation of the NP as being a contrastive topic (unlike the NP frequently in Burmese FQ constructions). Displacement of the quantifier to the right occurs simply in order to position the quantifier in the clause-final, new information focus position.

Now turning to Burmese, let us consider the case of an object NP with a quantifier whose content a speaker wishes to focus. Given the SOV basic word order in Burmese, in cases where there is no NP-quantifier separation, and the NP and quantifier remain adjacent to each other, this will result in the linear sequencing depicted in (87):

(87) NP_{Subject} [NP Q_{focused}]_{Object} V

In Burmese, as in many other SOV languages, the ‘clause-final’ position associated with new information focus is actually not fully clause-final, but the position immediately preceding the verb. Hence in the set of constituents which can be re-ordered (this not including the verb), a focused argument or adverbial is commonly placed in final position and preceded by other old and backgrounded information. In the neutral word order configuration in (87), the quantifier associated with the NP already naturally occurs in the pre-verbal focus position, so there is a natural convergence of position and information structure status for the quantifier. Where quantifier float patterns do occur, and result in a further heightened focus effect on the quantifier due to the splitting and separation effect, this results in a splitting away of the NP stranding the quantifier in focus position and placement of the NP further forward in the sentence, in topic position. As the NP which is regularly displaced leftwards to topic position here is a commonly a bare noun, such a bare, indefinite noun/NP is interpreted generically as representing the type of the noun/NP, and this in turn results in its interpretation as a contrastive topic, as non-contrastive topic interpretations are restricted to entities that are definite in reference (hence the oddness of sentences such as: ‘?A book, I bought yesterday.’).

Consequently it can be suggested that the different word orders of Thai and Burmese interacting with cross-linguistic pressures to place focused, new information in final positions conspire to cause the major observable differences between floating quantifier constructions in the languages, namely: (1) in Thai, the quantifier undergoes movement, while in Burmese it is the NP which is regularly moved away stranding the quantifier, and (2) in Burmese: there is frequently a contrastive topic interpretation of the NP, while in Thai: no special interpretation of the NP occurs in floating quantifier constructions.

A further common property of Q-float constructions which appears to be shared widely across languages is the patterning that when NP and quantifier are separated, it is common for the NP to linearly precede the quantifier (...NP.....Q....), and the opposite sequencing of quantifier preceding NP is quite uncommon (...Q.....NP....), though sometimes suggested to be grammatically possible in languages such as Japanese.³⁸ This common linear distribution of NP and quantifier can arguably also be attributed to aspects of information structure, and the strong cross-linguistic tendency for new information to be sequenced following old information. In instances where an NP and a quantifier are split apart in Q-float constructions/configurations, it is much more likely that the generic reference value of the noun/NP can be assumed to be familiar, contextually retrievable and more available as a topic-like center of interest than the value of the quantifier, which will frequently resist any topic-like licensing in the information structure of a sentence, as illustrated in the contrast in (88):

- (88) a. As for apples, I want three.
 b. ??As for three, I want apples.

General principles of information structure interacting with parametrizable properties of languages therefore results in a range of quite predictable and understandable variation in the distribution of floating quantifiers and their associated NPs.

A final question we will consider here relating to the issue of word order, floating quantifiers and information status is the special occurrence of post-verbal elements in SOV languages and the syntactic status of ‘afterthought information’. Above it was mentioned that the basic word order of languages such as Burmese, Japanese and Korean is SOV, with the verb in clause-final position. In written forms of these languages, subjects, objects, obliques and other adjuncts may occur in a range of different orders preceding the verb, depending on their contextual information status as definite/indefinite, new/old, focused/topical material, but the verb is regularly final in its clause and does not participate

³⁸ A striking illustration of the pressures to conform to a linear NP > Q ordering can be noted from Mandarin Chinese, where the quantifier *dou* ‘all’ has grammaticalized in a fixed pre-verbal position, and is actually never combined with an associated NP in a single syntactic constituent. Whenever *dou* quantifies over an object NP which would normally follow the verb in the basic SVO word order of Mandarin, the object NP is actually forced to undergo repositioning into some position to the left of *dou* resulting in a linear NP>Q sequence. This may result in the NP object being immediately adjacent to *dou* as in (c) below, or further to the left in pre-subject topic position (d). Examples (a) and (b) show that *dou* cannot be combined with the NP object in post-verbal position, and that the object may not remain in situ following the verb if associated with *dou*:

- (a) *wǒ kàn-le dōu shū
 I read-ASP all book
 (b) *wǒ dōu kàn-le shū *..Q...NP..
 I all read-ASP book
 (c) wǒ shū dōu kàn-le ✓..NP...Q..
 I book all read-ASP
 (d) shū wǒ dōu kàn-le ✓..NP...Q..
 book I all read-ASP
 ‘I read all the books.’

in any linear re-ordering with arguments and adjuncts (hence Burmese, Japanese and Korean are often referred to as ‘verb-final’ languages). In spoken forms of these languages, however, certain non-verbal elements such as subject and object NPs are found to optionally occur following the verb. This being the case, an interesting question is whether it is possible for floating quantifiers to be positioned following the verb resulting in a separation and splitting more similar to that in Thai, with floating quantifiers often occurring in sentence-final positions which do not correspond to regular locations of the associated NPs.

In Burmese and Japanese, it is in fact possible for floating quantifiers to occur following the verb in a clause, as illustrated in (89) and (90):

- (89) manṅeḱa thuu zēe-hmaa **sa-ōuq** wε-tε, **thōun-ōuq**
 yesterday he market-in book buy-REAL 3-CL
 ‘Yesterday he bought books in the market, ..three to be precise’
- (90) Taroo-wa Kinokuniya-de **hon-o** katta, **san-satsu**
 Taroo-TOP Kinokuniya-in book-ACC bought 3-CL
 ‘Taroo bought books in Kinokuniya, ..three it was.’

However, such post-verbal elements in Burmese and Japanese are commonly interpreted in a particular way as ‘afterthoughts’ - information which a speaker adds to a sentence often in the way of further clarification – and it is often assumed that, as afterthoughts, post-verbal elements in verb-final languages such as Japanese are not syntactically integrated into the preceding clause (Kuno 1978, Sells 1999, Soshi and Hagiwara. 2004). The existence of such postposed ‘afterthought’ quantifiers in Burmese and Japanese raises a question about Thai where floating quantifiers regularly occur in sentence-final positions. Specifically, we may ask whether the patterns found in Thai are really different from those in (88) and (89) and whether clause/sentence-final floating quantifiers in Thai might be ‘afterthought’ additions to the sentence like post-verbal quantifiers in Burmese and Japanese? The answer to the latter question is ‘no’. Thai floating quantifiers are indeed clearly integrated into the syntactic structure of the sentence and are not just added on as ‘afterthoughts’ in sentence-final position. First, Thai floating quantifiers can in fact (optionally) precede sentence-final particles such as *laew* (see ex.39), showing that they occur within the main syntactic structure of the clause. Second, the prosodic break/intonational pause between verb and floating quantifier which characterizes the occurrence of post-verbal quantifiers in Burmese and Japanese, setting the quantifier off from the rest of preceding clause is not present with clause-final floating quantifiers in Thai, and these are intonationally integrated into the sentence without any separating pause. Third, although Burmese and Japanese post-verbal quantifiers may represent new information in some sense, they are not interpretable as obviously focal new information, unlike the situation in Thai. Because of this, there are clearly different restrictions on what kinds of quantifiers can occur post-verbally in Burmese/Japanese and what may occur

clause-finally in Thai. Significantly, in Burmese and Japanese, focused and interrogative quantifiers are unacceptable in post-verbal position:³⁹

- (91) *thuu zēe-hmaa saōuq wε-th-lē, behna-ōuq?
 he market-in book buy-REAL-Q how-many-CL
 Intended interpretation: ‘How many books did he buy in the market?’
- (92) a. *kare-wa hon-o katta no, nan-satsu?
 he-TOP book-ACC bought Q how.many-CL
 Intended: ‘How many books did he buy?’
 b. *kare-wa hon-o kawanakatta san-satsu-shika
 he-TOP book-ACC bought-NEG 3-CL-only
 Intended: ‘He only bought three books.’

This contrasts with patterns in Thai, where interrogative and focused quantifiers in final position are natural and frequent in their occurrence:

- (93) jon-bāad-nī mii nāksùksáa maa lēw/thúng kīi-khon?
 up-until-now be student come ASP/arrive how.many-CL
 ‘How many students have arrived so far?’
- (94) kháw hày ngən Dæng khèε sóong-bàat(thàwnān)
 he give money Daeng only 2-Baht (only)
 ‘He gave Daeng only two Baht.’

Thai is therefore clearly distinctive in this patterning, not only from SOV Burmese and Japanese, but also from SVO Chinese which permits ‘afterthought’-type VP-final floating quantifiers, as illustrated in (95), but not focused or interrogative floating quantifiers, as seen in (96):

- (95) yǒu rén lái zhǎo nǐ, sān-ge
 be person come find you 3-CL
 ‘There were some people who came looking for you, three people.’
- (96) *Zhāngsan xǐang-zhīdào yǒu rén lái zhǎo tā jǐ-ge
 Zhangsan want-to-know be person come find him how-many-CL
 Intended: ‘Zhangsan wanted to know how many people came looking for him.’

The conclusion from such contrasts is that Thai clause-final floating quantifiers are not simple afterthought elements and are integrated into the syntactic structure of the sentence in a way quite different to afterthoughts, which may simply be linear concatenations not syntactically connected to what precedes them as parts of a single sentence (Sells 1999, Soshi and Hagiwara 2004).

³⁹ Note that if the focused and interrogative quantifier-classifier pairs in (91) and (92) are positioned before the verb, these sentences are grammatical. There is just a special restriction on these elements when they occur in the post-verbal ‘afterthought’ position.

Having thus considered how aspects of the functional use of floating quantifier structures may bear on their distribution within the clause, and how certain interpretations are associated with quantifiers floated in various positions, we now close the paper with a brief summary of this exploration of Q-float phenomena in Thai and Burmese.

6. Summary

This paper set out to document and analyze patterns involving displaced, ‘floating’ quantifiers (and classifiers) in two neighboring languages of southeast Asia, Thai and Burmese, as a way to provide further potential insight into three general questions associated with floating quantifier constructions. First, what syntactic mechanisms result in the separation and linear distancing of a noun/NP and its associated quantifier? Second, what pragmatic/functional motivations might underlie the optional use of Q-float forms, licensing their occurrence? Third, from a comparative, typological viewpoint, do the mechanics and triggers of Q-float show signs of patterning in a uniform way across languages, or are there differences, and how might these be accounted for? Thai and Burmese were selected as the two principal languages of this micro-study from southeast Asia as both languages clearly exhibit the phenomenon of Q-float in appropriate contexts and allow for the separation of quantifier-classifier pairs from similar adjacent linear sequences of [NP quantifier classifier], hence seem to display parallel base resources in this regard. With regard to word order patterns at the clausal level, however, Thai and Burmese show significant differences, Thai being S-Aux-V-O and Burmese S-O-V-Aux in neutral sentences. One intended focus of the study was to look at how this difference in basic word order might potentially influence the way Q-float is manifested in a language. As the investigation of Thai and Burmese proceeded, it was concluded that both languages separate nouns/NPs from modifying quantifiers by mechanisms of movement (hence subject to syntactic restrictions commonly associated with movement, such as island constraints), and that Q-float in both languages regularly appears to be linked to the expression of focused new information. However, Thai and Burmese were seen to differ with regard to the element of the NP/quantifier pair that undergoes movement/displacement in Q-float constructions, in Thai the quantifier shifting rightwards to a clause-final position in a way similar to extraposition or Heavy NP Shift in English, while in Burmese it is the NP which undergoes a leftwards displacement in sentences exhibiting Q-float. This major difference between Thai and Burmese was attributed to a tension between principals determining linear word order - in this instance, information structure - and those regulating hierarchical syntactic structure - here, the syntactic organisation of clauses in a head-initial or head-final way, resulting in SVO and SOV type languages. Both languages attempt to achieve an optimal ordering of separated NP and quantifier for the purposes of information structure and focus, and bring this about through the displacement of different elements in opposite directions in the clause, (new) focused constituents cross-linguistically favoring a rightward, clause-final position. The comparison of Q-float patterns in Thai and Burmese therefore illustrates how linear strings with an important shared property (...NP...Q...) may be produced in distinct ways by languages with different syntactic properties, underlining the fundamental importance of linear sequencing for (certain) aspects of interpretation despite the dominance of hierarchical structure in other areas of syntax and construal. The paper also provides the first substantial description of floating quantifier patterns in Thai and Burmese and so extends the available empirical coverage of this area of syntax in Asian languages, adding

it as a resource to previous insightful studies of Japanese and Korean and the issues discussed in these works (Kang 2002, Miyagawa 1989, Miyagawa and Arikawa 2007, Nakanishi 2008, Ko 2007 among others).

References

- Bobaljik, Jonathan. 2003. Floating Quantifiers: handle with care. In Lisa Cheng and Rint Sybesma, eds. *The Second Glot International State-of-the-Article Book*. Mouton de Gruyter, pp. 107-148.
- Fanselow, Gisbert and Damir Cavar. 2001. 'Distributed deletion', In Artemis Alexiadou (ed.), *Theoretical Approaches to Universals*, Amsterdam, Benjamins. 65- 97.
- Frazier, L. 1979. *On comprehending sentences: syntactic parsing strategies*. Bloomington Indiana: Indian Linguistics Club.
- Haegeman, Liliane. 1991. *An Introduction to Government and Binding Theory*. Blackwell: Oxford.
- Kang, B.M. (2002). Categories and meanings of Korean floating quantifiers – with some reference to Japanese. *Journal of East Asian Linguistics*, 11:375-398.
- Ko, Heejeong. 2007. Asymmetries in scrambling and cyclic linearization. *Linguistic Inquiry* 38:49-83.
- Kuno, Susumo. 1978. *Danwa no bunpoo* [Grammar of discourse]. Tokyo: Taishukan.
- Lee, Chungmin. 1999. 'Contrastive topic: a locus of the interface – evidence from Korean and English.' in K. Turner (ed) *The Semantics/Pragmatics Interface from Different Points of View*, Maryland Heights: Elsevier Science.
- Li, Y.-H. Audrey. 1999. Plurality in a classifier language. *Journal of East Asian Linguistics* 8:75-99.
- Miyagawa, Shigeru. 1989. *Structure and case-marking in Japanese*. Academic Press: San Diego.
- Miyagawa, Shigeru and Koji Arikawa. 2007. 'Locality in syntax and floated numeral quantifiers. *Linguistic Inquiry* 38:645-670.
- Nakanishi, Kimiko. 2008. 'The syntax and semantics of floating numeral quantifiers.' In S. Miyagawa and M. Saito *The Oxford Handbook of Japanese Linguistics*, 286-318, Oxford: Oxford University Press.
- Ouhalla, Jamal. 1994. *Introducing Transformational Grammar*. Arnold: London.
- Radford, Andrew. 1988. *Transformational Grammar*. Cambridge University Press: Cambridge.
- Ross, J.R. 1967. Constraints on variables in syntax. Ph.D. dissertation, MIT, published in *Infinite Syntax* Ablex: Norwood, 1986.
- Sells, Peter. 1999. Postposing in Japanese;. Ms. Stanford University.
- Simpson, Andrew. (2005). Classifiers and DP structure in southeast Asian languages. In *The Oxford Handbook of Comparative Syntax*, eds R. Kayne and G. Cinque, 806-38, Oxford University Press: Oxford).
- Simpson, Andrew and Tanmoy Bhattacharya. 2008. Ditransitives in Bangla/Bengali. Presentation made at the Conference on Ditransitives, Leipzig University.

Soshi, T. and H. Hagiwara. 2004. Asymmetry in linguistic dependency: Linguistic and Psychophysiological Studies of Japanese Right Dislocation. *English Linguistics* 21: 409-453.

THE EVOLUTION OF POLYFUNCTIONALITY OF *DÂY* CONSTRUCTION IN THAI: SPLIT PATTERNS OF POSSIBILITY-RELATED MODAL CONCEPTS*

Kiyoko Takahashi
Kanda University of International Studies
<kiyoko@kanda.kuis.ac.jp>

0 Abstract

This study aims to investigate historical changes in Thai modal concepts marked by the irrealis marker *càʔ* (> *càk*), especially possibility-related concepts denoted by periphrastic constructions containing the morpheme *dây*. I have examined irrealis expressions in Thai inscription corpora from the end of the 13th century through the 20th century, and found that Thai modal concepts in the present have gradually emerged. In semantic extensions of Thai modals, an original modal concept does not disappear even after a newly derived modal concept has established itself. Rather, it is as if one single line split into two lines both of which would continue extending.⁴⁰ In particular, I have identified two directions of semantic changes involved in split patterns of possibility-related modal concepts in Thai: (1) less subjective > more subjective; (2) non-volitive (participant-external) > volitive (participant-internal). The latter direction is opposed to the direction that has been postulated in the literature of historical semantic change. I have also found split patterns with little change in subjectivity, to which most of previous studies have not paid due attention.

Keywords: modals, historical semantics, epigraphy

1 Introduction

This paper examines historical changes in the types of modal concepts marked by the irrealis marker *càʔ* (> *càk*) in Thai. Specifically, this paper analyses split patterns of the modal concepts by using discourse corpora of Thai inscriptions from the end of the 13th century through the 20th century (see the list of the inscription corpora at the end of this

* Earlier versions of this paper were presented at LSJ134 (the 134th General Meeting of the Linguistic Society of Japan), Chiba, June 16-17, 2007 and at ICHL18 (the 18th International Conference on Historical Linguistics), Montréal, August 6-11, 2007. I would like to thank Rumiko Shinzato, Atsuhiko Kato, Heiko Narrog, and Bruce Horton for their helpful comments and suggestions on the earlier versions of this paper. I am also grateful to an anonymous reviewer of JSEALS (the Journal of the Southeast Asian Linguistics Society) for constructive and helpful comments on a draft of this paper. Errors and shortcomings are all my own.

⁴⁰ This kind of historical semantic change is often called ‘layering’ (Hopper 1991: 22) in the study of grammaticalization, i.e., the process whereby lexical items and constructions come in certain linguistic contexts to serve grammatical functions, and, once grammaticalized, continue to develop new grammatical functions (Hopper & Traugott 1993: xv).

Kiyoko Takahashi. 2011. The Evolution of Polyfunctionality of *dây* Construction in Thai: Split Patterns of Possibility-Related Modal Concepts.

Journal of the Southeast Asian Linguistics Society 4.1:147-166.

Copyright vested in the author.

Received 19/10/2010, revised text accepted 13/1/2011

paper).⁴¹ I use the term ‘split patterns’ to mean patterns of semantic extension in which a more specific modal concept emerges out of a less specific modal concept, and the original concept does not readily disappear and typically remains long after the new concept becomes established, as if one single line split into two lines both of which keep extending (shown in Diagram 1 in Section 5).

In this paper, special consideration is given to split patterns of the possibility-related types of modal notions such as ‘ability (or participant-internal, volitive possibility)’ and ‘circumstantial possibility (or participant-external, non-volitive possibility)’. Those possibility-related concepts are expressed by specific constructions ending up with the morpheme *dây* (see examples (18) through (23) in Section 5). *Dây*’s original substantive meaning is supposed to be ‘quantity-emergence’, that is, coming into existence of a certain number, volume, distance, or duration of some entity (Takahashi & Shinzato 2005, Takahashi 2008b).⁴²

I have found that in Thai, ‘ability (participant-internal, volitive possibility)’ evolved out of ‘circumstantial possibility (participant-external, non-volitive possibility)’, and in turn ‘circumstantial possibility’ derived from ‘circumstantial impossibility’. The direction of the former split pattern (circumstantial possibility > ability) is opposed to the direction found in most of the existing studies on the evolution of modal concepts, namely ‘ability > circumstantial possibility (root possibility)’ or ‘participant-internal modality > participant-external modality’ (Bybee 1988, Bybee & Pagliuca 1985, Bybee et al. 1994, Heine & Kuteva 2002, Traugott 2006, Traugott & Dasher 2002, van der Auwera & Plungian 1998, *inter alia*).⁴³ Furthermore, the latter split pattern (circumstantial impossibility >

⁴¹ It is generally believed that the oldest Thai inscription (King Ram Khamhaeng Inscription) was made in 1292. There is an opposing view, however, that it was made in 1354-1376 (Prasithrathsint 2006: 129). Since I do not have knowledge of the inscription-dating, in this study I simply follow the former general view that the oldest one was made in 1292 (the end of the 13th century).

Note that the number of Thai inscriptions is limited to a few hundreds and each inscription is not long. Moreover, most of their contents is about the Buddhist religion or royalty-related matters such as politics and laws. Since the inscription data thus have the limitations in terms of quantity and genre-variety, it might be possible that some missing types of modal expressions with *càk* / *cà?* (and *dây*) in earlier periods were actually in use but unfortunately are not attested due to the limitations of the data. This study, however, uses the inscription data because the data nonetheless have a very good point. That is, the production years of most of the inscriptions have been estimated ranging from the end of the 13th century throughout the 20th century, which enables us to at least roughly see which types of Thai expressions were getting more and more (or less and less) common in a relatively long time span.

⁴² Matisoff (1991: 419-420) and Bisang (1996: 569-570) among others postulate that the original meaning of the Thai morpheme *dây* is ‘get’. Enfiled (2001: 279-280, 2004: 276) posits that the original meaning of the Lao morpheme *dāj* and the corresponding morpheme in other Tai languages is ‘come to have, acquire’.

⁴³ However, van der Auwera et al. (2009) examine semantic extension of so-called ‘acquisitive’ modals (e.g. ‘I get to watch TV tonight’) and acknowledge its plausible bidirectionality between ‘participant-external’ and ‘participant-internal’ possibility. It also deserves mentioning that Li (2004) and Shinzato (2008) report the details of historical semantic change from ‘participant-external’ to ‘participant-internal’ possibility observed in Chinese and Japanese, respectively. Li argues for the following semantic change of a Chinese modal for possibility *dé/de*: (a) ‘participant-external non-deontic’ > (b) ‘deontic/epistemic’ > (c) ‘participant-internal possibility’. In a similar vein, Shinzato discusses the following stages of semantic change of two

circumstantial possibility) has scarcely been discussed in the relevant literature.⁴⁴ I believe that the findings of this study shed some light on the theory of the directionality of the evolution of modal concepts.

This paper is organized as follows. Sections 2 and 3, respectively, explain the nature of the irrealis marker in Thai and the semantic map of modal concepts to be used for visualizing a seven-century course of gradual change in Thai modal concepts. Section 4 presents quantitative data regarding historical changes in Thai modal concepts. Section 5 accounts for split patterns of possibility-related modal concepts in Thai, some of which have hitherto been rarely documented. In Section 6, I will give concluding remarks.

2 Thai irrealis marker

Thai, an isolating language, has a morpheme indicating the irrealis status of the situation represented by the following verb phrase, namely *càʔ*. Thai lacks the obligatory coding of grammatical categories, and therefore *càʔ* is not an obligatory marker of irrealis. However, in present-day Thai *càʔ* is necessarily used to express counterfactuality (Srioutai 2004) and to form an irrealis complement with the complementizer or relativizer *thîi* (Diller (2001) calls *càʔ* in this usage ‘irrealis-complement formative’). Accordingly, I regard *càʔ* in modern Thai as the irrealis marker proper.

The irrealis marker *càʔ* derives from the verb *càk* meaning ‘to intend, consider’ (Diller 1988). The word form *càk* began to change into *càʔ* in the middle of the 14th century, and *càʔ* came to be commonly used in succeeding ages (Diller 1988, Takahashi 2007). Diller (2001) further hypothesizes that the two morphemes, *càk* and *càʔ*, have undergone the process of grammaticalization proceeding along the path indicated in (1), which is based on the attested grammaticalization path of the English auxiliary ‘will’ (Bybee et al. 1991: 26-29) as shown in (2).

- (1) Grammaticalization path of *càk* / *càʔ* in Thai (Diller 2001):
desire > intention > future > irrealis
- (2) Grammaticalization path of ‘will’ in English (Bybee et al. 1991):
desire > intention > future > probability, imperative

However, Diller (2001) did not present Thai diachronic data to verify this hypothesis. This study, on the other hand, uses a data-driven approach. Previously I have gathered modal expressions marked by the irrealis marker from the aforementioned inscription corpora (Takahashi 2007). With this diachronic corpus data, in this paper I will examine historical changes in the types of the modal concepts. I consider a new modal concept encoded by a certain conventionalized construction to be fully established when it becomes amenable to marking by irrealis. From the corpus data, I see that, over time, the number of specific types of modal concepts marked by the irrealis marker has gradually increased. Put

Japanese modals for possibility *naru* and *dekiru*: (a) ‘disabling conditions exist external to non-specific individuals’ > (b) ‘disabling conditions exist external to specific individuals temporarily’ > (c) ‘disabling conditions exist internal to specific individuals temporarily’ > (d) ‘disabling conditions exist internal to specific individuals permanently’ > (e) ‘enabling conditions exist internal to specific individuals permanently’.

⁴⁴ The semantic shift from negative to positive possibility of Japanese and Thai modals for possibility is reported in Takahashi & Shinzato (2005), Shinzato (2008) and Takahashi (2008b).

differently, specific modal concepts with clear semantic boundaries in the present have gradually emerged from inclusive and ambiguous modal concepts with fuzzy boundaries in the past. I have also found that the two morphemes, *càk* and *cà?*, have been compatible with quite a wide range of modal situations since as early as the 14th century (shown in Figure 3 in Section 4). This fact makes it clear that they had become an irrealis marker proper before that period. Therefore, if Diller's (2001) assumption about the grammaticalization path of the two morphemes (i.e. desire > intention > future > irrealis) holds, this process of grammaticalization must have been accomplished prior to the 14th century. In this paper, however, I am not able to validate his assumption because of the lack of relevant diachronic data in the corpora. What I can do instead is to reconstruct the most reasonable split patterns of Thai modal concepts marked by the full-fledged irrealis marker from the 14th century onward (see Section 5).

3 Semantic map for modal concepts

I follow Narrog's (2005: 683-690) idea that modal concepts can be classified in terms of the dimensions of (a) 'speaker-orientation' or 'subjectivity' and (b) 'volitivity'. Figure 1 illustrates a semantic map for modal concepts, where six representative modal concepts hold respective positions in terms of the two dimensions (cf. Figure 3 in Narrog 2005: 694).

<i>(Hearer-oriented, Intersubjective)</i>	Imperative	Conditional concessive
↑ <i>Speaker-oriented, Subjective</i>	Obligation	Prediction
↓ <i>Event-oriented, Objective</i>	Volition	Apparent imminence
	<i>Volitive</i>	<i>Non-volitive</i>

Figure 1. An illustrated semantic map for modal concepts

'Speaker-oriented' or 'subjective' situations are directly linked to the speaker's own modal judgment at the time of speech in the given speech situation. In contrast, 'event-oriented' or 'objective' situations are concerned with conditions on a participant of the described event, independent of the speaker and the present speech act. However, the distinction between subjective and objective situations is a gradient. For instance, 'volition' and 'apparent imminence' are less subjective than 'obligation' and 'prediction'. On the other hand, the distinction between 'volitive' and 'non-volitive' situations is a discrete one since they constitute a dichotomy as to whether or not the 'element of will' (Jespersen 1924: 320-321) is involved, or to put it another way, whether or not the person concerned has an 'interest' (Heine 1995: 29) or 'preference' (Givón 1990: 528-530) in an event occurring or not occurring. In this regard, it is evident that deontic (e.g. obligation) and boulomaic (e.g. volition) modal situations are volitive, while epistemic (e.g. prediction) and evidential (e.g. apparent imminence) modal situations are non-volitive.

It is known that languages with the irrealis-marking system may differ in the range of modal contexts in which the irrealis marker appears, that is, the types of modal concepts may cross-linguistically differ (Mithun 1995, Elliott 2000). Figure 2 below indicates Thai modal concepts that are attested to be compatible with the irrealis marker *càk* / *cà?*.

(Intersubj.)		1.1 Conditional concessive (e.g. (3))
Subj. ↑	2.1 Appropriateness (e.g. (12)), 2.2 Permission (e.g. (13)), 2.3 Obligation (e.g. (14))	1.2 Prediction (e.g. (4)), 1.3 Inevitability (e.g. (5)), 1.4 Conclusion (e.g. (6)(7))
	2.4 Ability (e.g. (15))	1.5 Circumstantial impossibility (e.g. (8)), 1.6 Circumstantial possibility (e.g. (9))
↓ Obj.	2.5 Volition (e.g. (16)), 2.6 Desire (e.g. (17))	1.7 Conditionals (e.g. (10)), 1.8 Apparent imminence (e.g. (11))
	<i>Volitive</i>	<i>Non-volitive</i>

Figure 2. *Semantic map for Thai modal concepts*

From Figure 2 we can see how close or distant each modal concept is to the others. Take the concepts of ‘ability’ (2.4) and ‘circumstantial possibility’ (1.6) for example. They show the same degree of subjectivity, but they are contrastive in their volitivity values; that is, ‘ability’ is volitive, while ‘circumstantial possibility’ is non-volitive.

It should be noted that although I accept the subjectivity and volitivity dimensions posited by Narrog (2005), the position of ‘ability’ in my semantic map of modals (Figure 2) is not the same as that in his semantic map of modals (cf. Figures 4 and 7 in Narrog 2005: 695, 702). He considers ‘ability’, like ‘apparent imminence’, as non-volitive and less subjective than ‘circumstantial possibility’, whereas I consider it to be volitive and no less subjective than ‘circumstantial possibility’. In my opinion, the concept of a human being’s physical or mental ‘ability’ entails his volition or desire to become able to do something, which is evidently distinct from the purely non-volitive concept of ‘circumstantial possibility’.⁴⁵ Moreover, I assume that the concept of ‘possibility’ in general, be it volitive or non-volitive, always involves the speaker’s evaluation of the possibility. That is to say, when mentioning some possibility (or impossibility), the speaker should have evaluated in what sense it is possible (or impossible). This is the reason why I regard ‘ability’, ‘circumstantial impossibility’ and ‘circumstantial possibility’ to be more subjective than ‘volition’, ‘desire’, ‘conditionals’ and ‘apparent imminence’ which are more objective and closer to mere propositional concepts.

Thai expressions for all the modal concepts listed in Figure 2 above, which are from the inscription corpus data, are given in examples (3) to (17) below.⁴⁶

⁴⁵ Narrog (personal communication) comments that although especially learnt or acquired ability might be subject to the person’s volition, there is good cross-linguistic evidence that ability also is non-volitive and many languages express it through a non-volitive ‘out-of-control’ (‘spontaneous’) marker.

⁴⁶ Sample expressions with *càk / càp* cited in this paper are transcribed into phonetic equivalents in present-day Thai. Abbreviations for functional morphemes in the English glosses are: IRR(ealis), NEG(ative), POSSI(bility), PART(icle), TOP(ic marker), REL(ativizer), CONJ(unction), CAUS(ative), PERF(ective), COP(ula), and NOM(inalizer).

1. Non-volitive

1.1 Conditional concessive ‘Even if ...’

- (3) suuu **càk** náp kôo læ mí? thûan
 even if IRR count CONJ NEG fully
Even if you would count them, you could not finish counting them. [5] (1361)⁴⁷

1.2 Prediction ‘... will ...’

- (4) **cà?** pay sùu ?abaayyathúk sǎ plàaw
 IRR go towards the way to hell PERF uselessly
They will go to hell in vain. [91] (1734)

1.3 Inevitability ‘... will ...; It is inevitable to ...’

- (5) sǎnkèet weelaa phrá?sǒŋ **cà?** loŋ sùat mon nay bòot
 notice time monk IRR go down recite a sutra in temple
They notice the time when monks will come to recite a sutra in the hall. [190]
 (1782-1925)

1.4 Negative conviction (Conclusion) ‘It is never possible to ...’

- (6) chây **cà?** nýom yindii lûiam sǎy látthí? sàatsanǎa
 NEG IRR favor be glad believe in ideology faith religion
 ?ùuun nòk càak phrá?phútthasàatsanǎa nán hǎa mí? dây
 other besides Buddhism TOP seek NEG POSSI
*It is impossible that we would willingly believe in other religions than the
 Buddhism.* [187] (1782-1925)

1.4 Ironical interrogative (Conclusion) ‘Wherever can one find ...?’

- (7) **cà?** hǎa mít mǔian câw thîi nǎy dây
 IRR seek friend be like he where POSSI
Wherever can I find a good friend like him? [278] (1925-1978)

1.5 Circumstantial impossibility ‘... is not achieved; it is not possible to ...’

- (8) **cà?** phannanaa bôo mí? dây læy
 IRR describe NEG POSSI PART
It is impossible to describe it. [86] (1528)

1.6 Circumstantial possibility ‘... is achieved; it is possible to ...’

- (9) mây mii pratau thîi **cà?** ?òk pay phaay nòk dây
 NEG there is door REL IRR exit go outside POSSI
There is no door through which we can go out. [146] (1782-1925)

⁴⁷ The number in the brackets attached to the end of each English translation (e.g. [5]) is the registration number for the identification of inscriptions that have been discovered in and around Thailand, and the number in the parentheses (e.g. (1361)) is the estimated production year of each source inscription.

1.7 Conditionals ‘If ...’

- (10) *phì?* **càk** *náp* *dûay* *duan* *dây* *yìp* *mùuun* *sii* *phan* *hòk* *sìp*
 if IRR count with month emerge 24,060
duan
 month
If we count by the month, it will be 24,060 months. [3] (1357)

1.8 Apparent imminence ‘... is about to ...’

- (11) *muaŋ* *sùkhõothay* *níi* *mii* *dàŋ* **càk** *tèek*
 city Sukhothai TOP noisy as if IRR be broken
The city of Sukhothai is so noisy as to be nearly broken. [1] (1292)

2. Volitive

2.1 Appropriateness ‘... should ...; It is appropriate to ...’

- (12) *khuan* **cà?** *pen* *thîi* *chúuun* *chom* *yindii*
 should IRR COP NOM love be glad
It should be loved. [178] (1782-1925)

2.2 Permission ‘It is possible for one to ...’

- (13) **cà?** *phùuk* *nay* *bân* *kõ* *dây*
 IRR perform a ceremony in home CONJ POSSI
One can perform a ceremony (of establishing the Sangha communion area) in the home. [193] (1925-1978)

2.3 Obligation ‘One must ...’

- (14) *phûu* *thîi* *tham* *chûa* **cà?** *tõŋ* *ráp* *thúk*
 person REL do bad IRR must receive sorrow
Those who commit a sin must suffer. [256] (1925-1978)

2.4 Ability ‘One is able to ...’

- (15) *ʔaat* *săamâat* **cà?** *ʔathíthăan* *dûay* *khaathăa* *bòt* *day* *bòt* *nùŋ*
 be able IRR pray with Pali verse a certain paragraph
hây *khrop* *dây*
 CAUS fully POSSI
One is able to pray by saying a certain full paragraph of a Pali verse. [257] (1925-1978)

2.5 Volitive ‘One will ...’

- (16) *phûa* **cà?** *hây* *ʔaanaaprachaarâatsàdõon* *tháj* *puan*
 in order to IRR CAUS the people in general
chom *lên*
 look at with admiration play
... to let the general public look at it respectfully. [187] (1782-1925)

2.6 Desire ‘One wants to ...’

- (17) mii sàtthaa **cà?** khrây sâaŋ ʔaaraam
 have faith IRR desire build temple
With faith in the Buddhism, they want to build a temple. [86] (1528)

4 Quantitative data

This section provides quantitative data on occurrences of *càk* and *cà?* in Thai inscriptions, which give an entire picture of gradual increase in the number of specific types of Thai modal concepts marked by *càk* / *cà?*.

Table 1 shows the number of occurrences of the two morphemes in different semantic contexts of the inscription discourses. For convenience’s sake, I divide the documented history of the Thai language into the following four periods according to different dynasties:

Period I: the Sukhothai period (1292-1438)

Period II: the Ayutthaya-Thonburi period (1438-1782)

Period III: the first half of the Rattanakosin period (1782-1925)

Period IV: the latter half of the Rattanakosin period (1925-1978)

A total of 635 tokens of *càk* and *cà?* were found in the inscriptions, though 15 of them were not decodable due to the unreadability of some inscriptions in Period I. The ratio comparing the two morphemes’ occurrences in each period is also indicated at the bottom of Table 1.

Table 1 reveals that before Period III the ratio of occurrences of the two morphemes in the volitive contexts exhibits higher values than that in the non-volitive contexts, but after Period III the two ratios do not differ very much. Note that *càk* was used as a verb in Period I. Since the number of occurrences of its verbal usage is too small (only four), I cannot tell exactly what verbal meaning it had. At any rate, the important point is that since Period I the two morphemes have been capable of functioning as an irrealis marker occurring not only in the contexts of ‘desire’, ‘volition (intention)’ and ‘prediction (future)’, which are named in (1) above, but also in the contexts of ‘conditional concessive’, ‘inevitability’, ‘conditionals’ and ‘apparent imminence’.

Table 1. Occurrences of *càk* and *cà?* in different semantic contexts of Thai inscription discourses⁴⁸

	Period I Sukhothai 1292-1438		Period II Ayutthaya & Thonburi 1438-1782		Period III Rattanakosin King Rama 1-6 1782-1925		Period IV Rattanakosin King Rama 7-9 1925-1978	
1. <i>Non-volitive</i>	50	38%	21	28%	73	53%	134	48%
1.1 Conditional concessive	6	5%	1	1%	11	8%	6	2%
1.2 Prediction, 1.3 Inevitability, 1.4 Conclusion (including Negative conviction, Ironical interrogative)	32	24%	14	18%	56	41%	109	39%
1.5 Circumstantial impossibility, 1.6 Circumstantial possibility	0		2	3%	3	2%	11	4%
1.7 Conditionals, 1.8 Apparent imminence	12	9%	4	5%	3	2%	8	3%
2. <i>Volitive</i>	77	59%	53	72%	64	47%	144	52%
2.1 Appropriateness, 2.2 Permission, 2.3 Obligation	0		4	5%	10	7%	45	16%
2.4 Ability	0		0		0		8	3%
2.5 Volition, 2.6 Desire	77	59%	49	66%	54	39%	91	33%
VERB <i>càk</i>	4	3%	0		0		0	
Total	131 +15	100%	74	100%	137	100%	278	100%
<i>càk</i> : <i>cà?</i>	119 : 27		44 : 30		1 : 136		4 : 274	

Figures 3 through 6 below are semantic maps designating a variety of Thai modal concepts marked by the irrealis marker *càk* / *cà?* in each of the four periods. From the first two Figures (Figures 3 and 4) we can see that the irrealis marker began to be used in the contexts of ‘appropriateness’ and ‘circumstantial impossibility’ in Period II (1438-1782).

⁴⁸ I have rounded off the decimal fractions of the percentages indicated in Table 1 and Figures 3 to 6, and so the total of the percentages in each part may be slightly under or over 100%.

	<p>Conditional concessive 5%</p> <p>Prediction 11%, Inevitability 14%</p> <p>Conditionals 8%, Apparent imminence 2%⁴⁹</p>
<p>Volition 38%, Desire 21%</p> <p><i>Volitive</i></p>	<p><i>Non-volitive</i></p>

Figure 3. *Semantic map for Thai modal concepts marked by the irrealis marker in Period I*

	<p>Conditional concessive 1%</p> <p>Prediction 9%, Inevitability 9%</p> <p>Circumstantial impossibility 3%</p> <p>Conditionals 5%</p>
<p>Appropriateness 5%</p> <p>Volition 54%, Desire 12%</p> <p><i>Volitive</i></p>	<p><i>Non-volitive</i></p>

Figure 4. *Semantic map for Thai modal concepts marked by the irrealis marker in Period II*

Figure 4 above and Figure 5 below show that the irrealis marker began to be used in the contexts of ‘permission’, ‘conclusion’ and ‘circumstantial possibility’ in Period III (1782-1925).

⁴⁹ Although the concept ‘apparent imminence’ disappears in the semantic maps of Period II-IV (Figures 4 to 6), I do not think that the irrealis marker was incompatible with the meaning of ‘apparent imminence’ during that time. The occurrence frequency of ‘apparent imminence’ expressions in the inscriptions is low, presumably because inscription discourses normally have nothing to do with situations characterized as ‘apparent imminence’. In fact, the irrealis marker in present-day Thai is still compatible with the meaning of ‘apparent imminence’. There are some entrenched forms with the irrealis marker for the concept (e.g. *kamlaŋ cà?* ‘be about to’, *klây cà?* ‘nearly’) which are used in oral discourses more frequently than in written discourses (Takahashi 2002).

<p>Appropriateness 7%, Permission 1%</p> <p>Volition 32%, Desire 7%</p>	<p>Conditional concessive 8%</p> <p>Prediction 13%, Inevitability 24%, Conclusion 2%</p> <p>Circumstantial impossi. 1%, Circumstantial possi. 1%</p> <p>Conditionals 2%</p>
<i>Volitive</i>	<i>Non-volitive</i>

Figure 5. *Semantic map for Thai modal concepts marked by the irrealis marker in Period III*

Figure 5 above and Figure 6 below show that the irrealis marker began to be used in the contexts of ‘obligation’ and ‘ability’ in Period IV (1925-1978).

<p>Approp. 14%, Perm. 1%, Obligation 1%</p> <p>Ability 3%</p> <p>Volition 24%, Desire 8%</p>	<p>Conditional concessive 2%</p> <p>Prediction 14%, Inevitability 24%, Conclusion 1%</p> <p>Circumstantial impossi. 2%, Circumstantial possi. 2%</p> <p>Conditionals 3%</p>
<i>Volitive</i>	<i>Non-volitive</i>

Figure 6. *Semantic map for Thai modal concepts marked by the irrealis marker in Period IV*

Thus, specific modal concepts in Thai have gradually emerged, at least, since the 15th century. In addition, the irrealis marker came to be less and less used in the contexts of ‘volition’ (38% → 54% → 32% → 24%) and ‘desire’ (21% → 12% → 7% → 8%), while it came to be more and more used in the contexts of ‘inevitability’ (14% → 9% → 24% → 24%), ‘prediction’ (11% → 9% → 13% → 14%) and ‘appropriateness’ (0% → 5% → 7% → 14%). It follows that formerly the irrealis marker was more likely to mark less subjective types of irrealis situations, but recently it has been used to mark more subjective types.

5 Split patterns of Thai modal concepts related to possibility

In this section, I will closely examine historical split patterns of the possibility-related types of modal concepts (such as ‘circumstantial possibility’ and ‘ability’) expressed in Thai. Here I limit the discussion to the issue of split patterns of possibility-related concepts because I could not find a sufficient number of tokens of expressions for modal concepts other than those related to possibility in the inscription data.

Examples (18) through (23) provide samples of Thai modal expressions of the possibility-related types which I have collected from the inscription corpora. To identify possible split patterns involving these possibility-related concepts, I have analysed similarities in their syntactic forms and semantics.

1.5 Circumstantial impossibility ‘... is not achieved; it is not possible to ...’

- (18)=(8) **cà?** phannanaa bõ mí? **dây** læy
 IRR describe NEG POSSI PART
It is impossible to describe it. [86] (1528)

1.4 Negative conviction (Conclusion) ‘It is never possible to ...’

- (19)=(6) **chây** **cà?** níyom yindii lûam sǎy látthí? sàatsanǎa
 NEG IRR favor be glad believe in ideology faith religion
 ?ũũũũ nōk càak phrá?phútthasàatsanǎa nán hǎa mí? **dây**
 other besides Buddhism TOP seek NEG POSSI
It is impossible that we would willingly believe in other religions than the Buddhism. [187] (1782-1925)

1.6 Circumstantial possibility ‘... is achieved; it is possible to ...’

- (20)=(9) **mây** mii pratuu thii **cà?** ?òk pay phaay nōk **dây**
 NEG there is door REL IRR exit go outside POSSI
There is no door through which we can go out. [146] (1782-1925)

2.2 Permission ‘It is possible for one to ...’

- (21)=(13) **cà?** phùuk nay bǎan kō **dây**
 IRR perform a ceremony in home CONJ POSSI
One can perform a ceremony (of establishing the Sangha communion area) in the home. [193] (1925-1978)

2.4 Ability ‘One is able to ...’

- (22)=(15) ?àat sǎamâat **cà?** ?athíthǎan dūay khaathǎa bõt day bõt nùŋ
 be able IRR pray with Pali verse a certain paragraph
 hây khróp **dây**
 CAUS fully POSSI
One is able to pray by saying a certain full paragraph of a Pali verse. [257] (1925-1978)

1.4 Ironical interrogative (Conclusion) ‘Wherever can one find ...?’

- (23)=(7) **cà?** hǎa mít mǔan câw thii nǎy **dây**
 IRR seek friend be like he where POSSI
Wherever can I find a good friend like him? [278] (1925-1978)

It is noteworthy that the syntactic forms for these possibility-related concepts in Thai all include the functional morpheme *dây* meaning ‘possibility’. Having analysed diachronic corpus data of *dây* expressions which I have compiled from Thai inscriptions (Takahashi 2005, 2006), I hypothesize *dây*’s grammaticalization from a verb into a modal, as follows (Takahashi & Shinzato 2005, Takahashi 2008b).⁵⁰ The verb *dây* originally expressed the meaning of ‘emergence’ in affirmative assertions and ‘non-emergence’ in

⁵⁰ Enfield (2001, 2003, 2004) has proposed different grammaticalization paths for so-called ‘acquire’ words which, he assumes, subsume the Thai verb *dây*.

negative assertions. When the meaning of ‘non-emergence’ was extended from the nominal domain (non-emergence of an object) into the verbal domain (non-achievement of an event) and the construction expressing the latter more abstract sense (VP NEG *dây*) became entrenched, *dây* gained the function of signalling ‘circumstantial impossibility’ (viz. it is not possible to do/be something due to certain circumstances). In the case of example (18) above, the entity that the writer wants to describe has very special properties in quality or quantity, and so he cannot perfectly describe its specialty (such as its wonderfulness or numerousness).

Plausible derivations of the other possibility-related constructions exemplified in (18) to (23) above are as the following.

(a) circumstantial impossibility (e.g. (18)) > negative conviction (e.g. (19)):

The verb *hăa* ‘seek’ came to be included in the construction for circumstantial impossibility (*hăa* NEG *dây* ‘seeking is not achieved; it is not possible to find out’), and another clause beginning with *chây* (*chây* (NP) VP ‘it is not ...’) came to precede the construction, which resulted in the construction for negative conviction (*chây* (NP) VP, *hăa* NEG *dây* ‘It is never possible to ...’).⁵¹

(b) circumstantial impossibility (e.g. (18)) > circumstantial possibility (e.g. (20)):

The negative in the construction for circumstantial impossibility came to be omitted, which gave rise to the construction for circumstantial possibility (VP *dây* ‘... is achieved; it is possible to ...’).

(c) circumstantial possibility (e.g. (20)) > permission (e.g. (21)):

The verb phrase in the construction for circumstantial possibility came to refer to an action to be engaged in by someone, which led to the construction for permission ((agent-NP) VP *dây* ‘It is possible for one to ...’).

(d) circumstantial possibility (e.g. (20)) > ability (e.g. (22)):

The two loanwords *ʔaat* ‘brave’ (from Khmer) and *săamâat* ‘hope, desire’ (from Pali) came to be added to the construction for circumstantial possibility. At the same time, the verb phrase in the construction came to refer to someone’s action. This resulted in the construction for ability ((agent-NP) *ʔaat* *săamâat* VP *dây* ‘One is able to ...’).

(e) circumstantial possibility (e.g. (20)) > ironical interrogative (e.g. (23)):

The interrogative *thîi năy* ‘where’ came to be included in the construction for circumstantial possibility, which yielded the construction for ironical interrogative (VP *thîi năy* *dây* ‘Wherever can one ...?’).

The derivational relationships among the five split patterns (a) to (e) is graphically shown in Diagram 1 below. In Period II, the concept of ‘circumstantial impossibility’ (e.g. (18)) was established. In Period III, ‘negative conviction (conclusion)’ (e.g. (19)) and

⁵¹ Although in present-day Thai the morpheme *chây* means ‘yes, that is correct’, originally it was a negative marker preceding a noun phrase (Takahashi 2008a). It came to be used as a constituent of some formulaic expressions like (19) before changing into the non-negative meaning in the present.

‘circumstantial possibility’ (e.g. (20)) derived from ‘circumstantial impossibility’, and also ‘permission’ (e.g. (21)) derived from ‘circumstantial possibility’. And in Period IV, ‘ability’ (e.g. (22)) and ‘ironical interrogative (conclusion)’ (e.g. (23)) derived from ‘circumstantial possibility’.

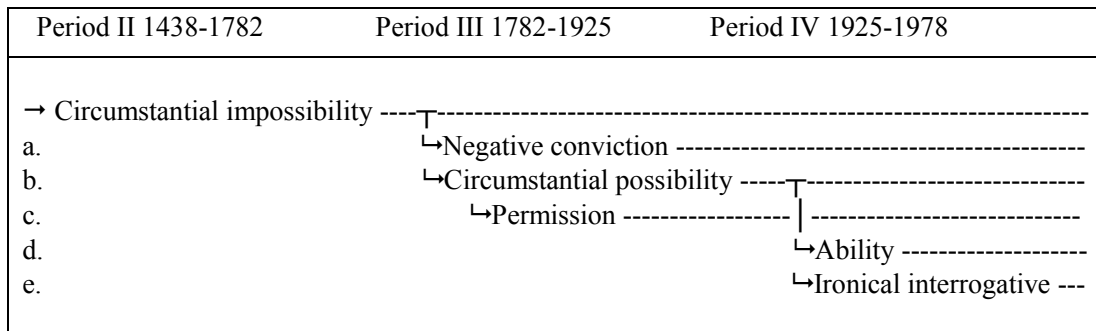


Diagram 1. *Split patterns of possibility-related modal concepts in Thai*

(24) and (25) summarize the directionalities of semantic changes involved in these split patterns occurring during Periods II-IV (1438-1978).

- (24) Directionality of changes in subjectivity of possibility-related concepts in Thai:
less subjective > more subjective
- (25) Directionality of changes in volitivity of possibility-related concepts in Thai:
non-volitive > volitive⁵²

Figures 7 to 10 below graphically depict different types of change in the values of subjectivity and volitivity associated with the split patterns (a) to (e) indicated in Diagram 1 above.

⁵² In her study on developments of Thai verbs into auxiliaries, Meesat (1997: 178) states that the verb *tâŋ* ‘hit, fit’ began to be used as an auxiliary expressing ‘obligation’ (i.e. volitive, deontic modality) in the reign of King Rama 1 (1782-1809), and then gained another auxiliary function to express ‘conclusion’ (i.e. non-volitive, epistemic modality) in the reign of King Rama 4 (1851-1868). Similarly, Diller (1988: 291) mentions that the Khmer-derived verb *ʔàat* ‘brave’ first developed into a modal for ‘ability’ (i.e. volitive, dynamic modality) and then came to function as a modal for ‘probability (possible conclusion)’ (i.e. non-volitive, epistemic modality). Unfortunately, however, my corpus data of irrealis expressions in the inscriptions dated in 1292-1978 do not include evident samples of the ‘conclusion’ usage of *tâŋ* (*câʔ tâŋ* VP ‘it is concluded by inference that ...’) and the ‘probability’ usage of *ʔàat* (*ʔàat câʔ* VP ‘it might appear that ...’), and therefore in this study, which deals with rather limited language data, I cannot attest the direction of these late semantic changes of Thai modals, viz., ‘volitive > non-volitive’.

Subj. ↑	Permission	Ironical interrogative	Negative conviction
↓ Obj.		↖ Circumstantial possibility (c)	↗ Circumstantial impossibility (a)
	<i>Volitive</i>	<i>Non-volitive</i>	

Figure 7. Semantic map for split patterns with change from ‘less subjective’ to ‘more subjective’

The three split patterns in Figure 7, (a) ‘circumstantial impossibility > negative conviction (conclusion)’, (c) ‘circumstantial possibility > permission’ and (e) ‘circumstantial possibility > ironical interrogative (conclusion)’ exhibit changes in subjectivity, namely ‘less subjective > more subjective’. This direction is consistent with the hypothesis called ‘unidirectionality of semantic change’ or ‘subjectification’ (Traugott 1982, 1989).

Subj. ↑	Ability	← Circumstantial possibility (d)	← Circumstantial impossibility (b)
↓ Obj.			
	<i>Volitive</i>	<i>Non-volitive</i>	

Figure 8. Semantic map for split patterns with little change in subjectivity

In contrast, the two split patterns in Figure 8, (b) ‘circumstantial impossibility > circumstantial possibility’ and (d) ‘circumstantial possibility > ability’ exhibit little change in subjectivity. It appears that judging to be possible is no more subjective than judging to be impossible; judging to be physically or mentally possible is no more subjective than judging to be circumstantially possible; and so on. So far the literature on semantic change has offered little in-depth analysis of these split patterns.⁵³

Subj. ↑	Permission	(c)
↓ Obj.	Ability	↖ Circumstantial possibility (d)
	<i>Volitive</i>	<i>Non-volitive</i>

Figure 9. Semantic map for split patterns with change from ‘non-volitive’ to ‘volitive’

The two split patterns in Figure 9, (c) ‘circumstantial possibility > permission’ and (d) ‘circumstantial possibility > ability’ exhibit shift in volitivity, namely ‘non-volitive >

⁵³ Shibuya (1993) and Shinzato (2008), which investigate the evolution of Japanese modals for possibility, are a couple of exceptions.

volitive'. This shift is opposed to the widespread view that non-volitive modal concepts arise from volitive (agent-oriented) ones (Bybee et al. 1994). However, my investigation into the Thai diachronic corpus data reveals that this view is not applicable to Thai. Aside from the recent semantic extensions of *tôŋ* (obligation > conclusion) and *ʔaat* (ability > probability) (see Footnote 52), semantic changes of Thai modals in the past are largely 'non-volitive > volitive'. This supports Narrog's (2005) claim that actually this direction of semantic change is pervasively found in many languages.

Subj. ↑		Ironical interrogative	Negative conviction
↓ Obj.		↑ Circumstantial possibility (e)	← Circumstantial impossibility → (b) (a)
	<i>Volitive</i>	<i>Non-volitive</i>	

Figure 10. *Semantic map for split patterns with no change in volitivity*

The three split patterns in Figure 10, (a) 'circumstantial impossibility > negative conviction (conclusion)', (b) 'circumstantial impossibility > circumstantial possibility' and (e) 'circumstantial possibility > ironical interrogative (conclusion)' exhibit no change in volitivity. Curiously, as far as Thai possibility-related concepts are concerned, non-volitive concepts more frequently split and became diversified than volitive concepts did.

6 Concluding remarks

In this study, I have examined actual discourses in Thai diachronic corpus data and found a number of plausible split patterns of Thai modal concepts related to possibility.

In order to explain the facts presented here, I suggest that the two opposing directions of semantic change between the two types of possibility-related concepts, i.e. 'ability > circumstantial possibility (root possibility)' and the other way around, may correspond to the typological dichotomy of the characteristic way of describing a situation, namely 'person- vs. situation-focus' (Hinds 1986; cf. also Teramura 1976 and Ikegami 1991). Person-focus languages (like English) tend to focus on the agent (i.e. conscious, willful and responsible person) in describing a situation, whereas situation-focus languages (like Japanese and Thai) tend to describe the whole event without placing a special focus upon the agent. It is reasonable to assume that in person-focus languages, the domain of volitive modal concepts, rather than the domain of non-volitive modal concepts, has rich split patterns, since lexical items expressing an agent's volitive action, ability, desire and the like are apt to evolve into modal markers, the process of which Langacker (1999: 297-315) calls 'attenuation'. In situation-focus languages, on the other hand, the domain of non-volitive modal concepts, rather than the domain of volitive modal concepts, has rich split patterns, since the lexical items for volition-related concepts are less likely to be employed for the expression of a modal concept. Whether this is true or not remains a matter of future research.

In conclusion, following Narrog (2005), I suggest that the directions of semantic change, except for the seemingly universal direction 'less subjective > more subjective', may vary across languages. I also suggest that variation in the directions may reflect

variation not only in the linguistic structures (morphosyntactic principles and lexical systems) but also in the speakers' preferred style of the description of a state of affairs.

References

- Bisang, Walter. 1996. Areal typology and grammaticalization: Processes of grammaticalization based on nouns and verbs in East and mainland South East Asian languages. *Studies in Language* 20.3: 519-597.
- Bybee, Joan L. 1988. Semantic substance vs. contrast in the development of grammatical meaning. *Proceedings of the 14th Annual Meeting of the Berkeley Linguistics Society*, 247-264.
- Bybee, Joan L. and William Pagliuca. 1985. Cross-linguistic comparison and the development of grammatical meaning. *Historical Semantics: Historical Word-Formation* ed. by Fisiak, Jacek, 59-83. Berlin, Mouton Publishers.
- Bybee, Joan L., William Pagliuca and Revere D. Perkins. 1991. Back to the future. *Approaches to Grammaticalization, Vol.2* ed. by Traugott, Elizabeth Closs and Bernd Haine, 17-58. Amsterdam, John Benjamins.
- Bybee, Joan, Revere Perkins and William Pagliuca. 1994. *The Evolution of Grammar: Tense, Aspect, and Modality in the Languages of the World*. Chicago, University of Chicago Press.
- Diller, Anthony V. N. 1988. Thai syntax and "national grammar". *Language Science* 10.2: 273-312.
- Diller, Anthony V. N. 2001. Grammaticalization and Tai syntactic change. *Essays in Tai Linguistics* ed. by Tingsabath, M.R. Kalaya and Arthur S. Abramson, 139-175. Bangkok, Chulalongkorn University Press.
- Elliott, Jennifer R. 2000. Realis and irrealis: Forms and concepts of the grammaticalisation of reality. *Linguistic Typology* 4: 55-90.
- Enfield, N. J. 2001. On genetic and areal linguistics in mainland South-East Asia: Parallel polyfunctionality of 'acquire'. *Areal Diffusion and Generic Inheritance: Problems in Comparative Linguistics* ed. by Aikhenvald, Alexandra Y. and R.M.W. Dixon, 255-290. Oxford, Oxford university Press.
- Enfield, N. J. 2003. *Linguistic Epidemiology: Semantics and Grammar of Language Contact in Mainland Southeast Asia*. London, RoutledgeCurzon.
- Enfield, N. J. 2004. Areal grammaticalization of postverbal 'acquire' in mainland Southeast Asia. *Papers from the Eleventh Annual Meeting of the Southeast Asian Linguistics Society 2001* ed. by Burusphat, Somsonge, 275-296. Tempe, Arizona State University.
- Givón, Talmy. 1990. *Syntax: A Functional-Typological Introduction, Vol.2*. Amsterdam, John Benjamins.
- Hinds, John. 1986. *Situation vs. Person Focus*. Tokyo, Kurosio.
- Heine, Bernd. 1995. Agent-oriented vs. epistemic modality: Some observations on German modals. *Modality in Grammar and Discourse* ed. by Bybee, Joan and Suzanne Fleischman, 17-53. Amsterdam, John Benjamins.
- Heine, Bernd and Tania Kuteva. 2002. *World Lexicon of Grammaticalization*. Cambridge, Cambridge University Press.

- Hopper, Paul J. 1991. On some principles of grammaticalization. *Approaches to Grammaticalization, Vol.1* ed. by Traugott, Elizabeth Closs and Bernd Heine, 17-35. Amsterdam, John Benjamins.
- Hopper, Paul J. and Elizabeth Closs Traugott. 1993. *Grammaticalization*. Cambridge, Cambridge University Press.
- Ikegami, Yoshihiko. 1991. 'DO-language' and 'BECOME-language': Two contrasting types of linguistic representation. *The Empire of Signs: Semiotic Essays on Japanese Culture* ed. by Ikegami, Yoshihiko, 285-326. Amsterdam, John Benjamins.
- Jespersen, Otto. 1924. *The Philosophy of Grammar*. London, George Allen & Unwin.
- Langacker, Ronald W. 1999. *Grammar and Conceptualization*. Berlin, Mouton de Gruyter.
- Li, Renzhi. 2004. *Modality in English and Chinese: A Typological Perspective*. Boca Raton, Dissertation.com.
- Matisoff, James A. 1991. Areal and universal dimensions of grammaticalization. *Approaches to Grammaticalization, Volume 2: Focus on Types of Grammatical Markers* ed. by Traugott, Elizabeth Closs and Bernd Heine, 383-453. Amsterdam, John Benjamins.
- Meesat, Paitaya. 1997. *A Study of Auxiliary Verbs Developed from Verbs in Thai*. Master's thesis, Chulalongkorn University.
- Mithun, Marianne. 1995. On the relativity of irrealis. *Modality in Grammar and Discourse* ed. by Bybee, Joan and Suzanne Fleischman, 367-388. Amsterdam, John Benjamins.
- Narrog, Heiko. 2005. Modality, mood and change of modal meanings: A new perspective. *Cognitive Linguistics* 16.4: 677-731.
- Prasithratsint, Amara. 2006. Development of *thùuk* passive marker in Thai. *Passivization and Typology* ed. by Abraham, Werner and Larisa Leisiö, 115-131. Amsterdam, John Benjamins.
- Shibuya, Katsumi. 1993. Aspects of Japanese potential expressions and their historical development in Japanese. *Memoirs of the Faculty of Letters, Osaka University* 33.1.
- Shinzato, Rumiko. 2008. From 'emergence' to 'ability': A case of Japanese *naru* and *dekiru*. *Proceedings of the 40th Annual Meeting of the Chicago Linguistic Society 2004*, 365-379.
- Srioutai, Jiranthara. 2004. The Thai *cà?*. A marker of tense or modality? *CamLing 2004: Proceedings of the University of Cambridge Second Postgraduate Conference in Language Research* ed. by Daskalaki, Evangelia et al., 273-280. Cambridge, The Cambridge Institute of Language Research.
- Takahashi, Kiyoko. 2002. Temporal proximity: An organizing factor of four major aspects in Thai. *Journal of Hokkaido Linguistics Society* 2: 1-17.
- Takahashi, Kiyoko. 2005. Uses of *dây* in Thai inscriptions. *Memoirs of Kanda University of International Studies* 17: 295-353.
- Takahashi, Kiyoko. 2006. Uses of *dây* in Thai inscriptions (2). *Memoirs of Kanda University of International Studies* 18: 427-448.
- Takahashi, Kiyoko. 2007. The irrealis marker in Thai. *Memoirs of Kanda University of International Studies* 19: 189-210.

- Takahashi, Kiyoko. 2008a. Negative markers in Thai. *Memoirs of Kanda University of International Studies* 20: 335-358.
- Takahashi, Kiyoko. 2008b. Grammaticalization paths of the Thai verb *dây*: A corpus-based study. *Papers from the 16th Annual Meeting of the Southeast Asian Linguistics Society 2006* ed. by Tadmor, Uri and Paul Sidwell, 122-132. Canberra, Pacific Linguistics.
- Takahashi, Kiyoko and Rumiko Shinzato. 2005. On grammaticalization of Japanese and Thai emergence verbs. *Proceedings of the 5th Meeting of the Japanese Cognitive Linguistics Association, Vol.5*, 197-207.
- Teramura, Hideo. 1976. BECOME vs. DO expressions: A comparison between voice-related expressions in Japanese and English [which is included in *The Collected Works of Hideo Teramura Vol.2: Linguistics and Teaching Japanese as a Foreign Language*, 213-232, published by Kurosio, Tokyo, in 1993].
- Traugott, Elizabeth Closs. 1982. From propositional to textual and expressive meanings: Some semantic-pragmatic aspects of grammaticalization. *Perspective on Historical Linguistics* ed. by Lehmann, Winfred P. and Yakov Malkiel, 245-271. Amsterdam, John Benjamins.
- Traugott, Elizabeth Closs. 1989. On the rise of epistemic meanings in English: An example of subjectification in semantic change. *Language* 65.1: 31-55.
- Traugott, Elizabeth Closs. 2006. Historical aspects of modality. *The Expression of Modality* ed. by Frawley, William, 107-139. Berlin, Mouton de Gruyter.
- Traugott, Elizabeth Closs and Richard B. Dasher. 2002. *Regularity in Semantic Change*. Cambridge, Cambridge University Press.
- van der Auwera, Johan and Vladimir A. Plungian. 1998. Modality's semantic map. *Linguistic Typology* 2: 79-124.
- van der Auwera, Johan, Petar Kehayov, and Alice Vittrant. 2009. Aquisitive modals. *Cross-linguistic Semantics of Tense, Aspect and Modality* ed. by Hogeweg, Lotte, Helen de Hoop, and A.L. Malchukov, 271-302. Amsterdam, John Benjamins.

Corpus data

- The Prime Minister's Secretariat, Thailand. 1924/1978. *Corpus of Thai Inscriptions, Installment 1*. Bangkok, The Prime Minister's Secretariat.
- The Prime Minister's Secretariat, Thailand. 1965. *Corpus of Thai Inscriptions, Installment 3*. Bangkok, The Prime Minister's Secretariat.
- The Prime Minister's Secretariat, Thailand. 1970. *Corpus of Thai Inscriptions, Installment 4*. Bangkok, The Prime Minister's Secretariat.
- The Prime Minister's Secretariat, Thailand. 1974. *Corpus of Thai Inscriptions, Installment 6.1*. Bangkok, The Prime Minister's Secretariat.
- The Prime Minister's Secretariat, Thailand. 1978. *Corpus of Thai Inscriptions, Installment 6.2*. Bangkok, The Prime Minister's Secretariat.
- Pongsripiian, Winai (ed.) 1991. *Corpus of Thai Inscriptions, Installment 7*. Bangkok, The Prime Minister's Secretariat.

JSEALS style guide for authors

Papers should be submitted electronically as .DOC or .RTF files. Initial drafts should be prepared with minimal formatting, and where possible with just one Unicode compliant font. After papers have been reviewed, revised and approved, authors are required to format their papers to style.

Download the template from the homepage at www.jseals.org, and type or paste text directly into this file, and then apply WORD STYLES according to the model.

Ensure that your pages settings are as follows:

Margins on A4 paper:

Top	2.5 cm
Bottom	4.0 cm
Left	2.5 cm
Right	2.5 cm
Header	1.25cm
Footer	3.0 cm
Gutter margin:	0.5 cm

Font – 12 point Times New Roman. If your version of Times New Roman does not accommodate the characters you require, you may substitute with DoulosIPA (freely available from sil.org). For non-roman fonts please consult with the Managing Editor first.

Line spacing: 14 pt exactly in normal text paragraphs

Mirror margins set ON

Image files for Tables or Figures should also be provided separately. Where possible, image files should be *greyscale* or *black & white*, .GIF or .TIF or .JPG, 300 DPI, 10cm wide.

Do not insert page numbers, or other header or footer text.

Do not link references, headings, tables or figure numbers in WORD.

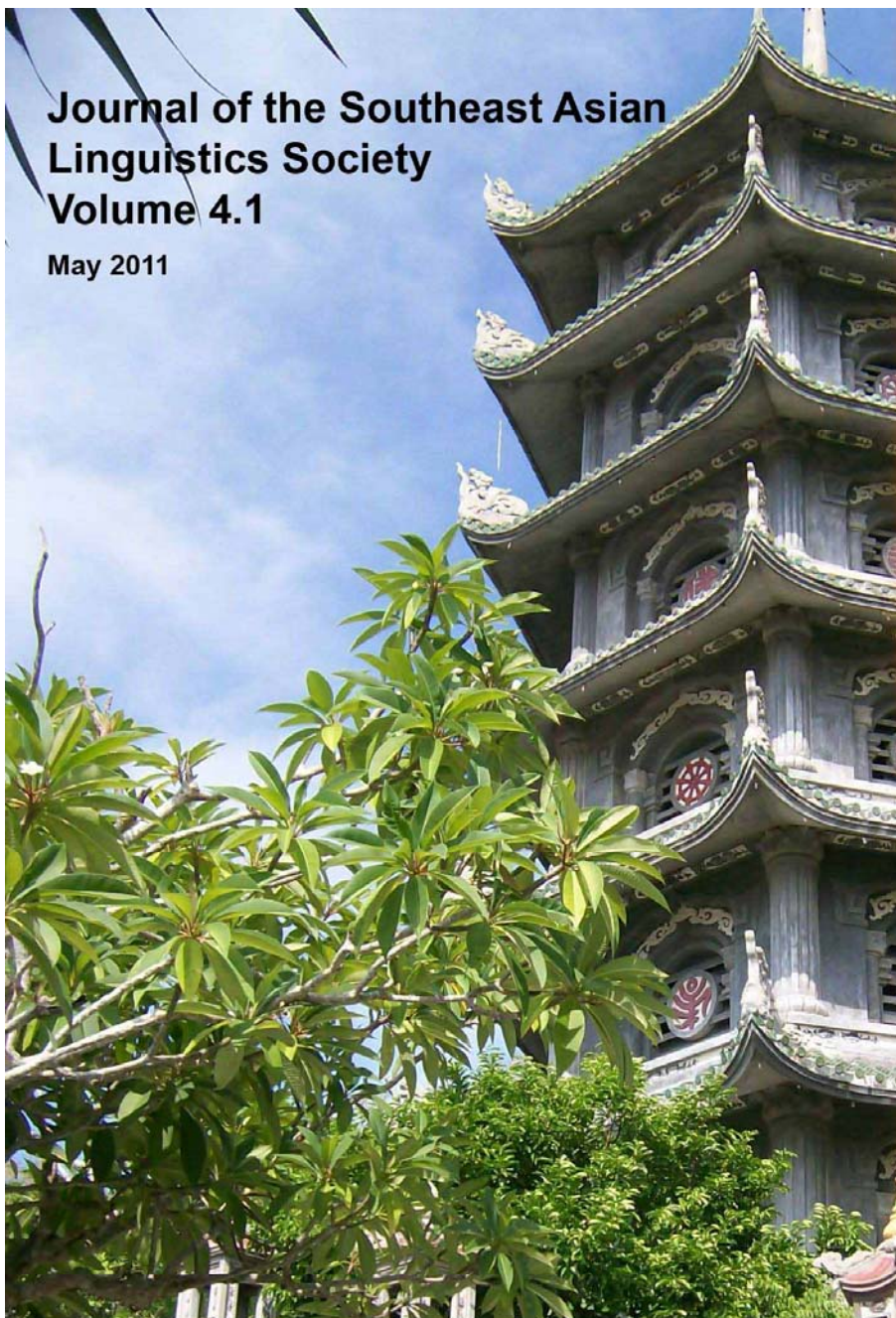
Use footnotes, not endnotes.

All papers must have an abstract of approximately 100 words, plus 3 keywords (preferably consistent with the keywords used at <http://sealang.net/sala/>).

If you have any quires, please consult the website at www.jseals.org or contact Paul Sidwell via <paulsidwell@yahoo.com> or <paul.sidwell@anu.edu.au>.

**Journal of the Southeast Asian
Linguistics Society
Volume 4.1**

May 2011



THE AUSTRALIAN NATIONAL UNIVERSITY



PACIFIC LINGUISTICS

ISSN 1836-6821

