

Basic word order in Tai Khamti: Language contact with Burmese

RIKKER DOCKUM

DEPARTMENT OF LINGUISTICS, YALE UNIVERSITY

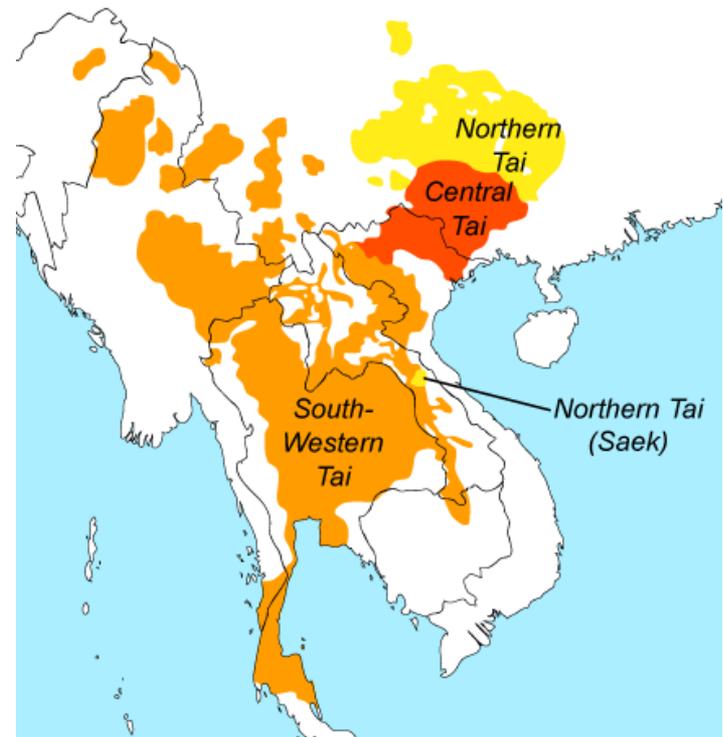
13TH INTERNATIONAL BURMA STUDIES CONFERENCE
BANGKOK, AUGUST 3-5, 2018

 BCS-1528386
Yale MACMILLAN CENTER

Kra-Dai and Southwestern Tai

- Kra-Dai (aka Tai-Kadai) has ~100 million speakers
- Southwestern Tai is the most populous subgroup

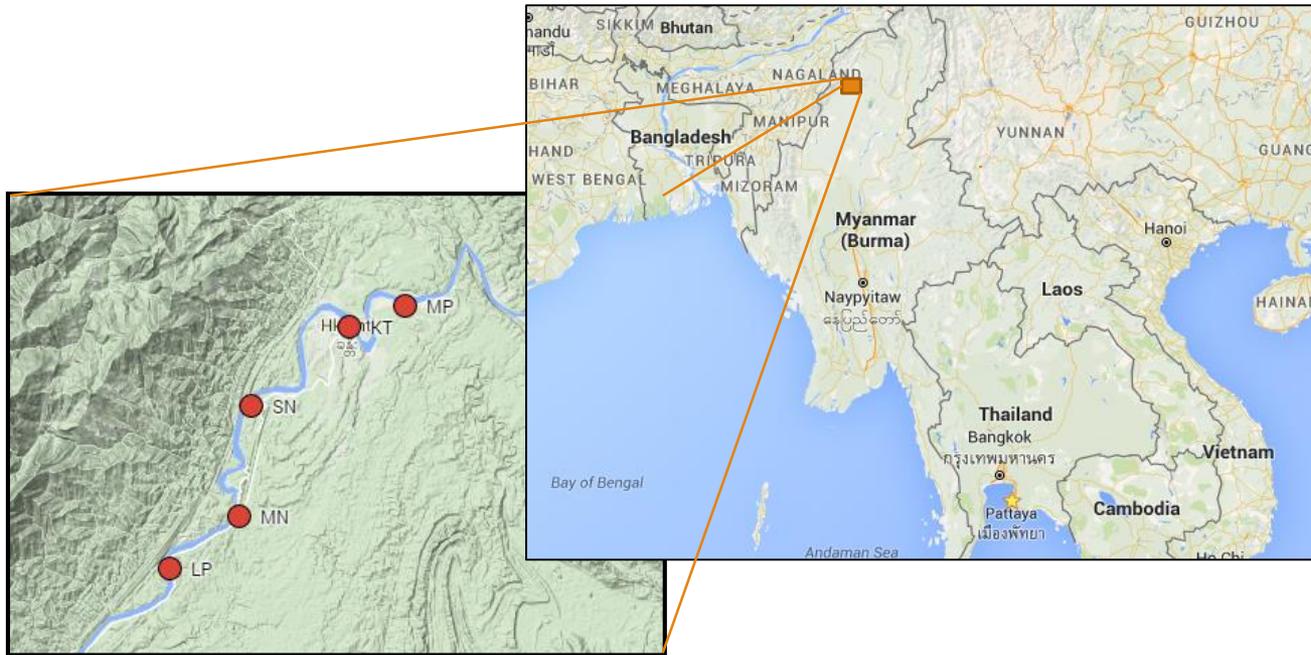
- Thailand 65 million
- Laos 3 million
- Myanmar 3 million
- China 1+ million
- Also India, Vietnam, Cambodia, Malaysia



Tai Khamti



Data gathering locations: Upper Chindwin river valley



Tai Khamti

Virtually all past work has been in Assam and Arunachal Pradesh, India:

- Robinson (1849), from materials of Rev. Nathan Brown
- Needham (1894), grammar and lexicon
- Grierson (1904), short text corpus
- Greenberg (1966), cites Khamti as exception to his Language Universal 4
- Harris (1976), brief notes and 700-word vocabulary
- Weidert (1977), 2000-word vocabulary

Still early days on Tai Khamti in Myanmar:

- Edmondson 2008 (from survey trips taken in mid-1990s)
- Inglis 2014 (first major work on Tai Khamti of Myanmar)
 - Speakers from Kachin state, resident in Thailand
- This project is the first extended fieldwork of Tai Khamti in Myanmar
 - First work on Tai Khamti of Chindwin River Valley, Khamti Township, Sagaing

SVO or SOV? What has been said

SUBJECT – VERB – OBJECT vs. SUBJECT – OBJECT – VERB

Needham 1894 – *Outline Grammar of Tai Khamti*

- “The order of words in a predicative sentence are **(1) subject, (2) direct object, (3) indirect object, and (4) verb**”

Grierson 1904 – *Linguistic Survey of India*

- “In Khamti (which at the present day stands isolated amid a sea of Tibeto-Burman languages), the order is as in them, **subject, object, verb...**”

Greenberg 1963 – *Some Universals of Grammar...*

- Khamti cited as an **exception** to Universal 4: “With overwhelmingly greater than chance frequency, languages with normal SOV order are postpositional”

SVO or SOV? What has been said

Wilaiwan 1986

- *“SOV is the dominant word order in Kamti while in other Tai dialects SVO is the dominant one.”*

Diller 1992

- *“The general impression...is of very ‘pragmatically’ controlled configuration”*

Morey 2005

- *“Both orders are still found [...] pragmatic factors are more important in determining the constituent orders than any ‘basic’ syntactic ordering.”*

Inglis 2014

- *“Khamti exhibits a basic SOV word order as generally demonstrated with the data in this dissertation.”*

Language Contact

- Morey 2005 – *“Areal influence, from a number of large verb-final languages [...] is undoubtedly playing a part in making verb final structures more acceptable to Tai speakers in Northeast India.”*
- Inglis 2014 – *“An SOV structure for Khamti is no doubt due to language contact with Tibeto-Burman languages...”*
- The Khamti case is interesting an interesting case because they are in contact with different SOV languages, and have been for 200 years.
 - If SOV order has become ‘basic’ Khamti dialects are following parallel trajectories.
 - If order is pragmatically conditioned, Khamti constituent order has probably been table for 100+ years.
- Why such change in Tai Khamti and not Phake or Aiton?

Sociolinguistic factors

Geography

- “Chindwin Khamti” speakers are lowland dwellers, wet rice farmers; villages border the Chindwin River
- Few traversable roads in Khamti District, travel primarily by river

Population

- ~7,000 residents in 12 Tai villages + Khamti Town (unofficial estimates)
- Many more Tais in neighboring Homalin township
- Referred to as Tai Laing, but difficult to distinguish due to convergence

Multilingualism and literacy

- Nearly all Chindwin Khamti speakers are also native speakers of Burmese
- Some intermarrying with T-B language speakers (both ‘Naga’ and Burmese)
- Local script is a recent adaptation of standard Shan script
 - (different from the modern Khamti script used in Kachin State/NE India)
- Low literacy in Khamti (a few hundred at most, <10 can read older scripts)

Sociolinguistic factors

Migration history

- Conquered modern Kachin State area (Mongkawng) from 14th century
- Migrations to Assam from 18th century
- Movement into the Upper Chindwin River Valley not long after that

Politics

- Upper Chindwin was ruled by Tai Khamti until the 1940s
- Until then the culturally dominant language of the immediate area
- Now many riverside villages populated with 'Naga' languages (diverse Tibeto-Burman languages), often with legacy Tai Khamti names
- In recent decades the main town is dominated by Burmans (civil servants, military, mining businesses, etc)

Data gathering

Tone Stimuli Response (TSR) Corpus (2015)

- Originally gathered to study tonal variation
- Question answering, targeting tones in certain words
- 37 speakers, ages 12-78, from 5 villages
- 16 target words x 3 questions per word
- 1,800 sentences (5,000+ including repetitions)
- 12,000 words (36,000+ including repetitions)

Corpus of other assorted texts (2014-2018)

- Folk tales, folk songs, 'pear story' and 'frog story' narratives, etc.
- 1,500+ sentences to date

Data gathering: TSR corpus

#	Form	Gloss	#	Form	Gloss	#	Form	Gloss	#	Form	Gloss
1	ma:¹	dog	2	kʰaw²	rice	3	pa:⁴	fish	4	kai⁶	chicken
5	mi¹	bear	6	ma:²	horse	7	kʰa:i⁴	buffalo	8	kʰa:⁶	galangal
9	pʰa:¹	wall	10	ɔi²	sugarcane	11	na:w⁴	star	12	taw⁶	turtle
13	sʰɿ¹	tiger	14	sa:ng²	elephant	15	nɿn⁴	moon	16	tʰo⁶	bean/nut
Tone 1 /45/ High rising			Tone 2 /21/ Low falling			Tone 4 /53/ High falling			Tone 6 /33/ Mid level		

Frame questions:

1. Have you ever seen / eaten / etc _____?
2. What kind of _____ have you seen / eaten / etc?
3. Where have you seen ____ / Where can _____ be found / etc?

SVO or SOV?

High-frequency transitive verbs

“Core transitive” verbs

- to see /han1/
- to eat /kyin4/

“Ambiguous transitive” verbs

- to like /thuk6 saw4/

Core transitive: /han1/ ‘to see’

- 495 instances of ‘see’ in the TSR Corpus
- In **sentence-final** position:
 - **354** /han1 kaw6/ ‘have ever seen’ **post-verbal modifier** (cf. pre-verbal Thai /khɯːj/)
 - **48** /han1 maa4/ ‘recently saw’ **post-verbal recent past marker**
 - **29** /han1 saaw2/ ‘see’ **matrix verb final (except for honorific particle)**
 - **17** /nai2 han1 u6/ ‘have seen’ **pre-verbal perfective, post-verbal present tense**

(6) k^hai6 kai6 nɛ2 ti2 hang4 kai6 **han1**
 maa4 saaw2
 egg chicken TOP LOC nest chicken **see**
 PAST HON

‘I saw the chicken eggs at the chicken coop.’

(7) kaw4-k^haa2 mi1 han1 kaw6 u6 saaw2
 1SG bear see before PRES HON

‘I have seen bears before.’

Core transitive: /kyin4/ ‘to eat’

- 218 instances of ‘eat’ in the TSR Corpus (212 /kyin4/, 6 /kin4/)
- In **sentence-final** position:
 - **105** /kyin4 kaw6/ ‘to have ever eaten’ **post-verbal modifier** (cf. pre-verbal Thai /khɤːj/)
 - **55** sentence-final /kyin4 maa4/ ‘eaten already’ (recent past)
 - **27** instances in sentence-final serial verb constructions (SVCs):

(8) mɔk6 k^haa6 nɛ2 tam4 s^hi1 kyin4 saaw2
bud galangal TOP pound CONJ eat HON
‘I pounded the galangal bud and ate it.’

(9) ɔi2 nɛ2 [..] puk6 wai2 s^hi1 kyin4 saaw2
sugarcane TOP plant ASP CONJ eat HON
‘I planted the sugarcane to eat.’

Ambiguous transitive: /tʰuk6 saw4/ 'to like'

- | | | | | |
|------|--|--|-------------------------------------|--------------|
| (10) | kaw4-kʰaa2
1SG-POLITE
S
'I like chicken eggs.' | kʰai6 kai6
egg chicken
O | tʰuk6 saw4
like
V | saaw2
HON |
| (11) | hə4-kʰaa2
1SG-POLITE
S
'I like (k.o.) fish.' | tʰuk6 saw4
like
V | paa4 kyit6
k.o. fish
O | saaw2
HON |
| (12) | kaw4-kʰaa2
1SG-POLITE
S
'I like (k.o.) beans.' | tʰuk6 saw4 nɛ2
like TOP
V | tʰo6 pʰɤk6
k.o. bean
O | saaw2
HON |
| (13) | tʰo6 nɛ2
bean TOP
O
'I like beans.' | tʰuk6 saw4 u6
like PRES
V | saaw2
HON | |

Remnants or active variation?

Inglis on SVO remnants (2014: 48-49):

- Generalized activity more likely to be encoded as VO
- Indefinite, non-specific

However, Chindwin Khamti does seem to have definite, specific VO usage:

(14) pɔ2 kaw4-k^haa2 njaa6 hɣn4 **het6** lɔŋg2 s^hi1 u6 saaw2
father 1SG-POLITE LOC home **do** **work** PROG PRES HON
'My father is working at home.' (sR Corpus)

Interestingly, VO and OV can intermingle in serial verb structures:

(15) pɔ2 nai2 p^hang1 nin4 **het6** s^hi1 kaa6 k^hut6 nin4
PURP PERF **bury** **ground** **do/CAUS** CONJ go dig ground
'In order to bury (the turtle), (they) dug the ground.' (Folktale Corpus)

Conclusions

- New corpora for Chindwin Khamti represent an additional 3,500 sentences or 50,000 words from a previously undocumented dialect
- Still difficult to tease apart what ‘basic’ constituent order means
- Certainly due to language contact but of unclear stability
- A more careful quantitative comparison of the two Myanmar Khamti corpora may reveal these subtle differences in distribution of constituent orders.
 - Inglis corpus from Kachin State ‘Khamti Lung’ ~ Big Khamti
 - Dockum corpus from Sagaing Division ‘Khamti On’ ~ Small Khamti
- Anyone who wants to study this issue in more detail is more than welcome to use my data!

Thank You!

References

- Diller, Anthony (1992). Tai Languages in Assam: Daughters or Ghosts? In Carol J. Compton & John F. Hartmann (eds.), *Papers on Tai Languages, Linguistics and Literatures*, 5–43. De Kalb, Ill.: Northern Illinois Centre for Southeast Asian Studies.
- Greenberg, Joseph (1963). Some universals of grammar with particular reference to the order of meaningful elements. In Joseph H. Greenberg (ed.), *Universals of Language*, 73–113. Cambridge, Mass.: MIT Press.
- Inglis, Douglas (2014). This here thing: Specifying Morphemes *an3*, *nai1*, and *mai2* in Tai Khamti Reference-point Constructions. Ph.D. Dissertation. University of Alberta.
- Khanittanan, Wilaiwan. (1986). Kamti Tai: from an SVO to an SOV language. *South Asian Languages: Structure, Convergence and Diglossia*, 174–178.
- Morey, Stephen. (2006). Constituent order change in the Tai languages of Assam. *Linguistic Typology*, 10(3). <https://doi.org/10.1515/LINGTY.2006.011>
- Needham, J. F. (1894). *Outline Grammar of the Tai (Khamti) Language*. Superintendent of Government Printing, Burma: Rangoon.



BCS-1528386

Yale MACMILLAN CENTER

Possible task effect?

- There is a possible task effect of asking people similar questions, priming them to use certain structures
- 1,811 unique sentences in Tone Stimuli Response (TSR) Corpus
- After filtering out sentence final TAM markers in the verb complex, and the sentence-final honorific particle /saaw2/, the three most common verbs are:
 - **301** have sentence-final /han1/ ‘to see’
 - **119** have sentence-final /kyin4/ ‘to eat’
 - **64** have sentence-final /t^huk6 saw4/ ‘to like’
 - **136** have sentence-final existential copula /jang4/ (also means ‘to have’)
- While there is some predictability and repetition of common structures, there is still a wide variety of sentence types

Key differences: Locative /ti2/ as accusative

The locative marker /ti2/ recruited as an accusative marker. Apparently optional, but this needs further study:

Basic locative usage:

- (1) s^hɣ1 nɛ2 **ti2** taa4 nam2 han1 kaw6 u6
tiger TOP **LOC** pier see EVER PRES
'(I) have seen a tiger at the pier before.'
- (2) maa1 **ti2** hɣn4 ling2 wai2 saaw2
dog **LOC** house raise
'(I) raise dogs at home.'

Innovative accusative usage:

- (3) kaw4-k^haa2**ti2** mi1 han1 u6 saaw2
1SG-POLITE**ACC** bear see PRES HON
'I see bears.'
- (4) kaw4-k^haa2**ti2** mɛ2 hak2 u6 saaw2
1SG-POLITE**ACC** mother love PRES HON
'I love my mother'

Key differences: Absence of /mai/

Other Khamti dialects, in both Northeast India and Kachin State, report a very high-frequency marker /mai/.

- Needham (1894) calls it a multipurpose case marker (“dative, accusative, and locative cases”)
- Wilaiwan (1986) refers to it as a postpositional object marker
- Diller (1992), using the Needham grammar as a corpus, notes that it also occurs with subjects, and also that it is not obligatory for marking case. Uncertain case-marking postposition, he concludes.
- Morey (2006) calls it an “anti-agentive” postposition.
- Inglis (2014) features /mai/ as a major component of his dissertation on Khamti specifiers, and teases apart its many uses.

Uniformly, everyone—from 1894 to 2014—note that it is one of the most frequent lexemes in Khamti. Of Inglis’s 90,000 word text corpus of Khamti spoken in Kachin State, Myanmar, /mai/ and the other two specifiers in his study make up 22% of all tokens (2014:13).

Note that /mai/ is also not used in Phake and Aiton, often considered Khamti’s closest relatives in Northeast India.

Key differences: /nai2/ vs /nɛ2/

Another high frequency form featured in Inglis 2014 is /nai/, which he treats as demonstrative. Chindwin Khamti does have a very high frequency postposed specifier /nɛ2/, which I am tentatively treating as a topic marker. Distinct in Chindwin Khamti from the demonstrative:

- /nai2/ ‘this’ /nan2/ ‘that’
- /thai2/ ‘here’ /than2/ or /han2/ ‘there’

/nai2/ is the likely lexical source for /nɛ2/ ~ /nə2/, but they are not one and the same. It cooccurs with the demonstratives, and sometimes appears to mark both noun and verb phrases.

(5)	kwaang4	an4	nai2	nə2	kwaang4	t ^h ɣk6	tsaaw2
	deer	CLF	this	TOP	deer	male	HON
	‘This deer was a male deer.’						