Internally-headed relative clauses in Shan*

Mary Moroney, *Cornell University*
mrm366@cornell.edu
@MaryTheLinguist
http://conf.ling.cornell.edu/mmoroney/

June 9, 2017

1 Introduction

- Shan is a Southwestern Tai language spoken in parts of Myanmar, Thailand, India, China, and Laos (Glick and Moeng 1991).
- Shan can form relative clauses in the same way as Thai, another Southwestern Tai language, as shown in (1-2).[1,2]

(1) [**Lik3** [an1 haw4 han1] naj5] man4 lEng1.
book COMP 1.SG see this 3.SG red
'A book that I see is red.'
(**Shan**; 2016.05.10.P.MM.07b)

(2) [**Naŋsɯ̆ɯ** [thîi raw hěn] níi] man dɛɛng.
book COMP 1.SG see this 3.SG red
'This book that I see is red.'
(**Thai**; 2016.09.28.MM.06)

- This is called a post-head relative clause, where the relative clause (*an1 haw4 han1* 'that I see') follows the head—the noun (*lik3* 'book') that it modifies.
- Unlike Thai, Shan also allows internally headed relative clauses (IHRCs), as shown in (3-4) where the head appears inside the relative clause itself.

*Thanks to Aye Twei Soe, Sireemas Maspong, and Ngampit Jagacinski who provided the Shan and Thai data. Thanks also to my committee members, John Whitman, Molly Diesing, and Sarah Murray, and the Spring 2017 Research Workshop for all their feedback. Any errors are my own.

[1] A note on the orthography for Shan: tʃ: c, ʔ: q, p^h: ph, t^h: th, $tʃ^h$: ch, k^h: kh, ɲ: nj, ŋ: ng, ɯ: y, ɯː: U, ʌ: v, ɛ: E, aː: aa, ɔ: O, rising tone: 1, low tone: 2, mid tone: 3, high tone: 4, falling tone: 5
Glossing conventions: 1: first person, 3: third person, ACC: accusative, ANIM: animal, ASP: aspect, CL: classifier, COMP: complementizer, GEN: genitive, HUM: human, IMPF: imperfect, IRR: irrealis, LOC: locative, NOM: nominative, PAST: past tense, PERF: perfect, PL: plural, ROUND: round, SG: singular, SM: scope marker, TOP: topic, VAL: validator

[2] The data on Shan comes primarily from my fieldwork on Shan working with a Shan speaker in Ithaca, NY from January 2016 to present. My Shan consultant is from Kayin (Karen) State in Myanmar and speaks the Yŭun Shan dialect —which is very different from the Taunggyi dialect. She also speaks Karen and received her education in Burmese. She has been in the United States for 5 years and speaks English, as well. Data was collected using a variety of elicitation methods: direct translation, grammaticality judgments, telling short stories, felicity judgments on grammatical sentences in specific contexts.

(3) [An1 haw4 han1 **lik3** naj5] man4 lEng1.
COMP 1.SG see book this 3.SG red
'A book that I see is red.' (**Shan**; 2016.11.15.MM.07)

(4) *[ʔan raw hěn **naŋsɯ̆ɯ** ʔan níi] man dɛɛng.
COMP 1.SG see book CL this 3.SG red
intended: 'This book that I see is red.'
(**Thai**; 2016.09.28.MM.02)

- This finding is interesting because internally headed relative clauses have not been reported in any other Tai languages (Warotamasikkhadit 1972; Morev and Moskalev 1979; Prasithrathsint and Yaowapat 2009).
- It is also somewhat typologically unusual because SVO languages are less likely to have internally headed relative clauses:
  - It had been assumed that a head-initial, SVO language like Shan could not have internally headed relative clauses because only OV languages were thought to allow IHRCs (Cole 1987).
  - This assumption has been called into question by verb initial languages like Seediq and Tagalog (Aldridge 2004) and SVO like Buli (Hiraiwa 2003).
  - 58/580 OV languages and 5/608 VO languages have IHRCs (Dryer and Haspelmath 2013).

Roadmap

2 Types of relative clauses

- Relative clauses are clauses used to modify a noun, called the **head** of the relative clause.
- These constructions are classified in terms of where the head appears with respect to the clause:
  - **Externally headed relative clauses (EHRCs)**: The head appears outside of the clause.
    - **Post-head**: The clause appears *after* the head.
    - **Pre-head**: The clause appears *before* the head.
  - **Internally headed relative clauses (IHRCs)**: The head appears inside of the clause.

(5) **Post-head relative, EHRC**

[**Naŋsɯ̆ɯ** [thîi raw hěn] níi] man dɛɛng.
book COMP 1.SG see this 3.SG red

'This book that I see is red.'
(**Thai**; 2016.09.28.MM.06)

(6) **Pre-head relative, EHRC**

Yoko-wa [[Taro-ga sara-no ue-ni oita] **keeki**]-o tabeta.
Yoko-TOP Taro-NOM plate-GEN on-LOC put cake-ACC eat

'Yoko ate a piece of cake which Taro put on a plate.'

(**Japanese**; Shimoyama 1999: (1))

(7) **Internally headed relative clause (IHRC)**

[[Nuna ishkay **bestya**-ta ranti-shqa-n]] alli bestya-m ka-rqo-n
man two horse-ACC buy-PERF-3 good horse-VAL be-PAST-3

'The two horses that the man bought were good horses.'

(**Quechua**; Grosu and Landman 1998: (93a)

Summary

Post-head RC	[	**head**	[	…	]	]
Pre-head RC	[		[	…	] **head**	]
Internally headed RC	[		[	… **head** …	]	]

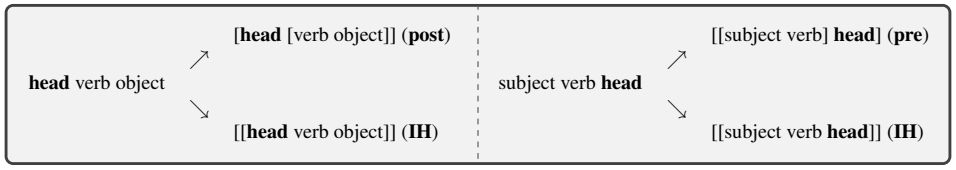
3 Shan relative clauses

3.1 Distinguishing between types of Shan relative clauses

- Shan has fairly strict SVO word order, as in (8).

(8) Haw4 han1 lik3.
1.SG see book
'I see a book.' (**Shan**; 2016.03.22.P.RH.05)

- In forming a relative clause, the head will appear in the position required by the relative clause construction rather than where it would appear within the relative clause if it were instead a matrix/independent clause.
- To determine the position of the head with respect to the clause—and thus classify the RC type—it is necessary to determine the boundaries of the clause.
- If the head is the *subject* of the RC, it would have the form ***head*** *verb object* and could be analyzed as [**head** [verb object]] (a post-head RC) or [[**head** verb object]] (an IHRC).
- If the head is the *object* of the RC, it would have the form *subject verb* ***head*** and could be analyzed as [[subject verb] **head**] (a pre-head RC) or [[subject verb **head**]] (an IHRC).



- To mark the left edge of the clause, I can use the complementizer *an1*—the equivalent of English *that* and Thai *thîi*.
- To mark the right edge of the clause, I use *wannaj5* ('today'), a word that should appear between the relative clause head and the right edge of the clause if the head is still inside the clause.[3]
- As shown in (9-11), *wan3naj5* can only appear at the beginning or end of a clause.

(9) Nan3 Li3 sUU4 kaj2 wan3naj5.
Nan Li buy chicken today
'Nan Li bought a chicken today.' (**Shan**; 2017.03.21.25)

(10) Wan3naj5 Nan3 Li3 sUU4 kaj2.
today Nan Li buy chicken
'Today, Nan Li bought a chicken.' (**Shan**; 2017.03.21.28)

(11) Nan3 Li3 (*wan3naj5) sUU4 (*wan3naj5) kaj2.
Nan Li today buy today chicken
intended: 'Nan Li bought a chicken today.' (**Shan**; 2017.03.21.26-27)

- In the following sections, these words are used to identify the types of Shan relative clauses.

Summary

head $\underset{\text{COMP}}{an1}$ verb object	→	[**head** [$\underset{\text{COMP}}{an1}$ verb object]] (**post**)
$\underset{\text{COMP}}{an1}$ **head** verb object	→	[[$\underset{\text{COMP}}{an1}$ **head** verb object]] (**IH**)
subject verb *wan3naj5* **head** (today)	→	[[subject verb *wan3naj5*] **head**] (today) (**pre**)
subject verb **head** *wan3naj5* (today)	→	[[subject verb **head** *wan3naj5*]] (today) (**IH**)

3.2 RC head as matrix subject

- For a translation of a relative clause with the head as the matrix subject, the first version offered was a internally headed relative, as in (12).
- The post-head relative clause, as in (13), was also acceptable.

(12) [[An1 Nan3 Li3 sUU4 **kaj2** wan3naj5]] man4 pen1 sii1 khaaw1.
COMP Nan Li bought chicken today 3.SG be color white
'A chicken Nan Li bought today was white.'

(**Shan IH**; 2017.03.07.MM.04a)

[3] Aspect aspect markers, like imperfective marker *ju3*, also appear on the right side of a clause:

(i) a. Haw4 cin1 tho2.
1.SG eat bean
'I eat beans.' (**Shan**; 2016.04.06.P.AC.09)

b. Haw4 cin1 tho2 ju3.
1.SG eat bean IMPF
'I'm eating beans.' (**Shan**; 2016.04.06.P.AC.10)

These were used to identify the right edge of the RC as an alternative to *wan3naj5* ('today').

(13) [**Kaj2**$_i$ [an1 Nan3 Li3 sUU4 t$_i$ wan3naj5]] man4 pen1 sii1 khaaw1.
chicken COMP Nan Li bought today 3.SG be color white
'A chicken Nan Li bought today was white.'
(**Shan Post**; 2017.03.07.MM.04b)

3.3 RC head as matrix object

- For a translation of a relative clause with the head as the matrix object, the first version offered was a post-head relative, as in (14).
- The internally headed relative clause, as in (15), was also grammatical.[4]

(14) Saj3 Kham3 aw3 [**kaj2**$_i$ [an1 Nan3 Li3 sUU4 maa3 t$_i$ wan3naj5]].
Saj3 Kham3 take chicken COMP Nan Li buy come today
'Saj Kham took a chicken Nan Li bought today.'
(**Shan Post**; 2017.03.07.MM.21a)

(15) Saj3 Kham3 aw3 [[an1 Nan3 Li3 sUU4 maa3 **kaj2** wan3naj5]].
Saj Kham take COMP Nan Li buy come chicken today
'Saj Kham took a chicken Nan Li bought today.'
(**Shan IH**; 2017.03.07.MM.21c)

- In addition to the elicited IHRCs, in (16) is an example of what is likely to be an IHRC from story told by my consultant.

(16) Luk3 kO5 kang2 naj5han3kO2 man4 sUU4 saw2 [an1 man4 cu4 **sU2kho4 haw2**]
child CL.HUM middle TOP he buy put COMP he like clothes PL
'The middle child, he bought and put the clothes that he liked,' (**Shan**; 2016.04.26.C.05)

Summary

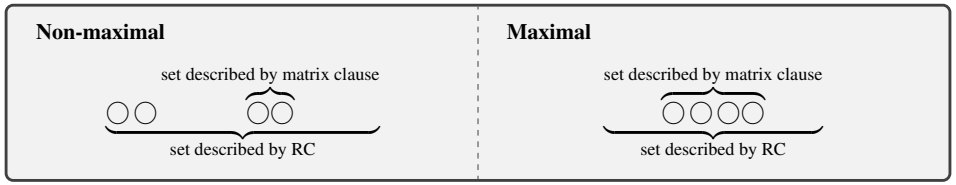
- Post-head and internally headed relatives are both allowed in Shan.

4 Maximality

- There are two broad patterns of interpretation of IHRCs:
  - Restrictive/intersective/non-maximal: Lakhota
  - Maximalizing: Quechua, Japanese

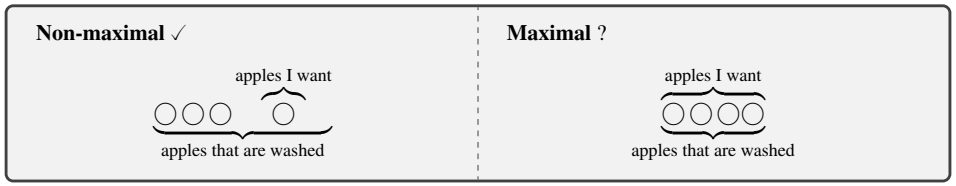
[4] The pre-head RC version of this sentence in (1) where the head of the clause is made specific/definite, was accepted in one elicitation session and rejected in another. Due to the inconsistently in acceptability of the pre-head relative, I will not pursue an analysis of them at this point.

(i) ?Saj3 Kham3 aw3 [[an1 Nan3 Li3 sUU4 maa3 t$_i$ wan3naj5] **kaj2**$_i$ to2 naj5].
Saj Kham take COMP Nan Li buy come today chicken CL.ANIM this
'Saj Kham took this chicken Nan Li bought today.'
(**Shan Pre**; 2017.03.07.MM.22)

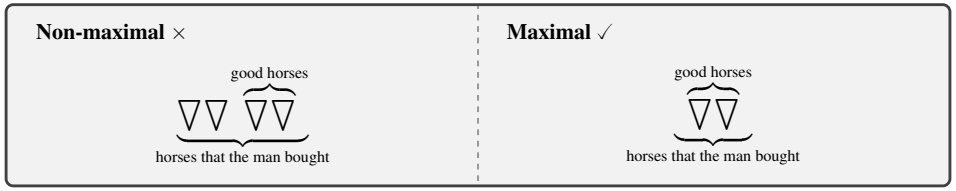


- Externally headed restrictive relative clauses have an intersective/non-maximal interpretation
- Grosu & Landman (1998) note that this difference correlates to the presence (Lakhota) or absence (Quechua/Japanese) of an overt D
- Shan, which lacks overt articles, would be expected to have maximalizing IHRCs.

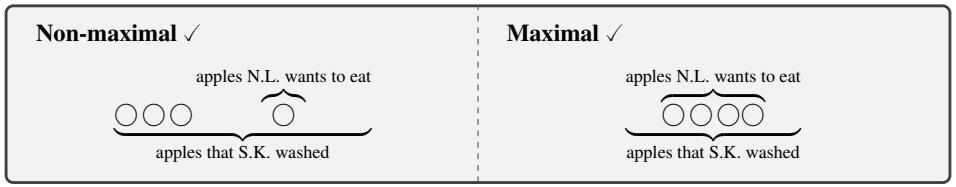
(17) [[**Thaspą** wazi taya yuzaza pi] cha] wachi
apple a-IRR well wash PL SM I-want
'I want an apple (nonspecific) that is well washed.'
(**Lakhota**; Grosu & Landman 1998: (92a))



(18) [Nuna **ishkay** **bestya**-ta ranti-shqa-n] alli bestya-m ka-rqo-n
man two horse-ACC buy-PERF-3 good horse-VAL be-PAST-3
'The two horses that the man bought were good horses.'
Unavailable interpretation: 'Two horses that the man bought were good horses.'
(**Quechua**; Grosu & Landman 1998: (93a)



(19) Nan3 Li3 khaj3 cin1 [an1 Saj3 Kham3 ti3 lang4 **mamO2** naj5]. Man4 khaj3 cin1
Nan Li want eat COMP Saj Kham will wash apple this 3.SG want eat
hwi4.
CL.ROUND
'Nan Li wants to eat an apple that Saj Kham will wash. She wants to eat one.'
(**Shan**; 2017.05.02.MM.30)



- Though Shan lacks overt determiners, its IHRCs appear to allow non-maximal interpretations, giving a counterexample to the correlation noted by Grosu & Landman (1998).

5 Analysis

- Since maximality is not something that my analysis needs to account for I can propose a simple raising analysis, often used for externally headed relative clauses.[5]
- The head of the clause is generated inside the relative clause and then moves to the SpecCP.
- In post-head relatives, this takes place overtly, and in IHRCs, this would take place at LF.
- This sort of analysis can get the intersective/non-maximal interpretation of relative clauses.
- The trees depicting this analysis can be seen in Appendix A.

5.1 Positive predictions

This head-raising analysis:

- Can capture the non-maximalizing, intersective interpretation of Shan relative clauses.
- Predicts that Shan IHRCs are subject to island constraints, which appears correct, as in (20).
- Predicts that internally headed relative clauses can have definite heads, as shown in (21-22).
  - This goes against what many people beginning with Williamson (1987) have argued, namely that relative clause heads in IHRCs are not allowed to be definite.

(20) *An1 Nan3 Li3 waa3 kam3 dahen3 an1 kong1 an2 lik3 naj5 man4 lEng1.
COMP Nan Li talk together with COMP person read book this 3.SG red
intended: 'The book that Nan Li talked to the person who read (it) is red.'
(**Shan**; 2017.05.30.MM.14)

(21) Saj3 Kham3 aw2 [an1 Nan3 Li3 boq5saw1 **maa1** to2 naj5 dahen3 maj5
Saj Kham take COMP Nan Li hit dog CL.ANIM this and/with stick
wan3naj5].
today
'Saj Kham took this dog that Nan Li hit with a stick today.' (**Shan**; 2017.03.07.MM.25a)

(22) [an1 **mamO5** an1 naj5 lang4 yaaw2 naj5] mii3 nuu3 lik3.
AN apple AN this wash PERF this exist on book
'This apple that is already washed is on the book.'
(**Shan**; 2017.05.02.MM.05)

5.2 Untested predictions

- This analysis also makes some predictions that might need to be ruled out:
  - Internally headed relative clauses can have base generated external quantifiers.
  - Externally headed relative clauses can strand CL + Dem in the relative clause:
    - This seems unlikely to be the case, perhaps because CL + Dem would be interpreted as a full DP.

[5] Shimoyama's (1999) E-type analysis is often used for languages with maximalizing IHRCs.

6 Conclusion

- Using certain words to identify the right and left edges of the relative clause, I have classified the types of relative clauses found in Shan.
- I have presented data on two types of Shan relative clauses: post-head and internally-headed relatives.
- The appearance of internally headed relative clauses is surprising since the construction is somewhat typologically unexpected and is not found in related Tai languages.
- Shan has non-maximalizing internally headed relative clauses, despite the fact that the language lacks overt articles.
- A head-raising analysis makes several predictions about Shan, some of which seem to be borne out and some of which need to be tested with future fieldwork.
- Other future work would be to investigate the connection between IHRC maximality and island sensitivity. Several accounts (e.g., Grosu 2002; Watanabe 2004) have connected the maximality of IHRCs to island sensitivity. Shan appears to be a counterexample.

References

ALDRIDGE, EDITH. 2004. Internally headed relative clauses in Austronesian languages. *Language and Linguistics* 5.99–129.

COLE, PETER. 1987. The structure of internally headed relative clauses. *Natural Language & Linguistic Theory* 5.277–302.

DRYER, MATTHEW S., and MARTIN HASPELMATH (eds.) 2013. *WALS online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Online: `http://wals.info/`.

GLICK, IRVING I., and SAO TERN MOENG. 1991. *Shan for English speakers*. Dunwoody Press.

GROSU, ALEXANDER. 2002. Strange relatives at the interface of two millennia. *Glot International* 6.145–167.

GROSU, ALEXANDER, and FRED LANDMAN. 1998. Strange relatives of the third kind. *Natural Language Semantics* 6.125–170.

HASTINGS, RACHEL ELIZABETH. 2004. *The syntax and semantics of relativization and quantification: The case of Quechua*. Cornell University dissertation.

HIRAIWA, KEN. 2003. Relativization in Buli. *Studies in Buli Grammar. Working Papers on Endangered and Less Familiar Languages* 4.45–84.

MOREV, LEV NIKOLAEVIC, and ALEKSEĬ ALEKSEEVICH MOSKALEV. 1979. The Lao language.

PRASITHRATHSINT, AMARA, and NATCHANAN YAOWAPAT. 2009. A typology of relative clauses in mainland southeast asian languages. *Mon-Khmer Studies* 38.1–23.

SHIMOYAMA, JUNKO. 1999. Internally headed relative clauses in Japanese and E-type anaphora. *Journal of East Asian Linguistics* 8.147–182.

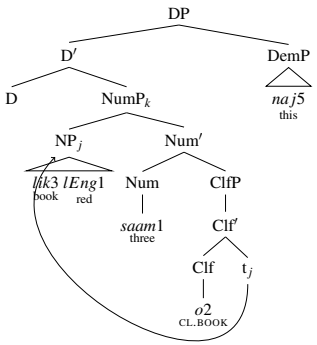
WAROTAMASIKKHADIT, UDOM. 1972. *Thai syntax: An outline*, vol. 68. Mouton De Gruyter.

WATANABE, AKIRA. 2004. Parametrization of quantificational determiners and head-internal relatives. *Language and Linguistics* 5.59–97.

WILLIAMSON, JANIS. 1987. An indefiniteness restriction for relative clauses in Lakhota. *The Representation of (In) definiteness*, 168–190.

A Details of analysis

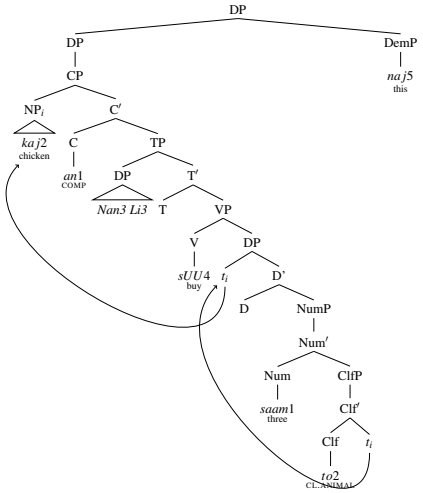
Proposed Shan DP



A.1 Relative clauses

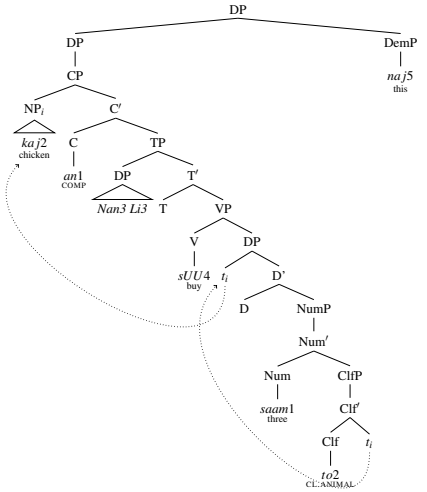
- For post-head relatives, the NP moves into the SpecCP of the relative clause, as in (23).

(23) **Post-head RC with internal quantifier**



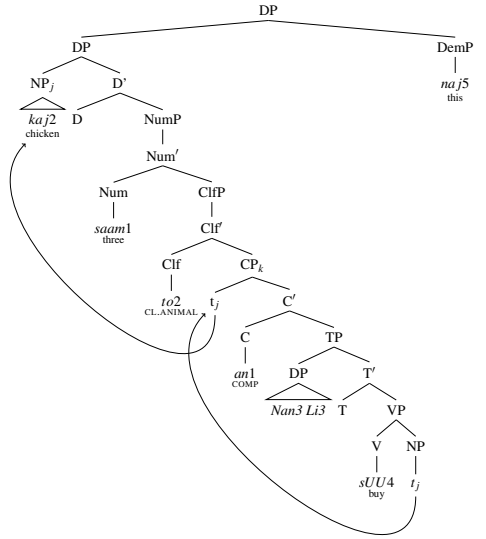
- For IHRCs, the head still moves to SpecCP, but this happens at LF, as in (24).

(24) **IH RC with internal quantifier: LF movement**



- When there is quantificational material modifying the head outside of the clause, it could either be analyzed as having the quantifier generated outside the relative clause, as in (25), or having the quantifier generated inside the clause and move to SpecCP along with the head NP.

(25) **Post-head RC with external quantifier: Base generate Q**

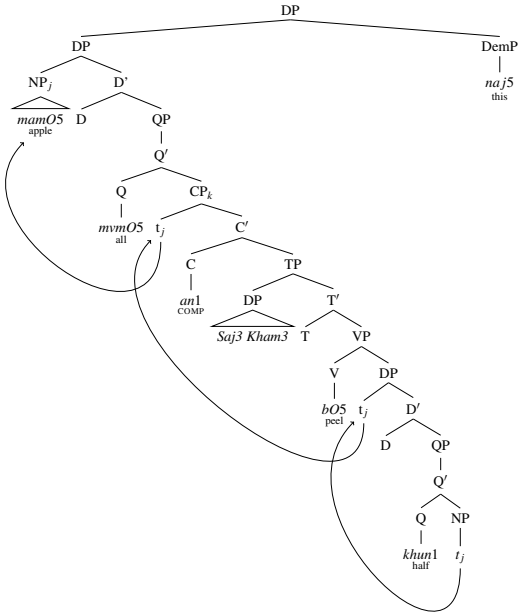


- Given that it is possible for there to be one quantifier inside the clause and one outside the clause, as in (26-27), the surface base generation analysis will work best.

(26) [Saj3 Kham3 bO5 **mamO5** **khun1** naj5] Nan3 Li3 cin1 pen3 mvmO5.
Saj Kham peel apple half this Nan Li eat ASP all
'Nan Li ate all of the half of the apples that Saj Kham peeled.'
(**Shan IH**; 2017.24.01.MM.25a)

(27) Nan3 Li3 cin1 **mamO5** **mvmO5** [an1 Saj3 Kham3 bO5 **khun1** naj5].
Nan Li eat apple all COMP Saj Kham peel half this
'Nan Li ate all of the half of the apples that Saj Kham peeled.'
(**Shan Post**; 2017.24.01.MM.25b)

(28) **Base generated external quantifier**



B Quantifier Scope

- Shimoyama (1999) makes the claim that in Japanese quantifiers are interpreted within the clause they appear.
- Hastings (2004) claims that in Cuzco Quechua one class of quantifiers ('strong quantifiers') is interpreted within the matrix clause and another class ('weak quantifiers') is interpreted within the relative clause regardless of where the quantifier appears.
- Shan quantifiers appear to take scope within the clause they appear.

(29) Nan3 Li3 cin1 pen3 [**mamO5** **saam1** hwi4 [$_{CP}$ an1 Saj3 Kham3 bO5] naj5].
Nan Li eat ASP apple three CL.ROUND COMP Saj Kham peel this
'Nan Li ate three apples that Saj Kham peeled.'
(**Shan Post**; 2017.02.28.MM.14)

- Number of apples Saj Kham peeled: at least 3
- Apples Nan Li ate: 3 of the peeled apples

(30) Nan3 Li3 cin1 pen3 [**mamO5** [$_{CP}$ an1 Saj3 Kham3 bO5 **saam1** hwi4] naj5].
Nan Li eat ASP apple COMP Saj Kham peel three CL.ROUND this
'Nan Li ate apples that Saj Kham peeled which are three in number.'
(**Shan Post**; 2017.02.28.MM.17)

- Number of apples Saj Kham peeled: 3
- Apples Nan Li ate: some number of the peeled apples

(31) Nan3 Li3 cin1 pen3 [[$_{CP}$ an1 Saj3 Kham3 bO5 **mamO5** **saam1** hwi4] naj5].
Nan Li eat ASP COMP Saj Kham peel apple three CL.ROUND this
'Nan Li ate apples that Saj Kham peeled which are three in number.'
(**Shan IH**; 2017.02.28.MM.13)

- Number of apples Saj Kham peeled: 3
- Apples Nan Li ate: some number of the peeled apples