Interpretation of internally-headed relative clauses in Shan

MARY MORONEY
Cornell University*

1 Introduction

Shan is a Southwestern Tai language related to Thai that is spoken in parts of Myanmar, Thailand, and other countries (Glick & Moeng 1991). In this paper, I work through evidence showing the existence of internally headed relative clauses (IHRCs) in Yün Shan, a dialect of Shan, and highlight an indefinite interpretation of IHRCs in this language that does not fit well with previous IHRC typologies. The basic IHRC and Post-Head constructions as shown in 1. In 1a is an IHRC. The head of the clause, lik ‘book’, appears in the same place it would in independent clause SVO structure. In 1b, on the other hand, the head is in the typical Post-Head position, appearing before the whole clause, in front of òàn, which I analyze as a complementizer.

(1) Relative clauses in Shan

   a. [ʔǎn háw hǎn lik nǎj] mán lēŋ.
      COMP 1.SG see book this 3.SG red
      ‘The/A book that I see is red.’ (IHRC)

      book COMP 1.SG see this 3.SG red
      ‘The/A book that I see is red.’ (Post-Head)

*Thanks to Aye Twei Soe who provided the Shan data. Thanks also to Molly Diesing, Carol-Rose Little, Sarah Murray, and John Whitman and the audiences of the Chulalongkorn International Student Symposium on Southeast Asian Linguistics and BLS 44 for all their feedback. Any errors are my own.

¹The Shan data comes from my fieldwork with a Shan speaker in Ithaca, NY from January 2016 to September 2017. My consultant is from Meiwai village, near Papun in Kayin (Karen) State in Myanmar. She speaks the Yûn Shan dialect, which is very different from the Taunggyi dialect. She also speaks Karen, Burmese, and English. She has been in the United States for 5 years. Data was collected using a variety of elicitation methods: telling short stories, grammaticality judgments, and felicity judgments.

Mary Moroney

The availability of IHRCs in Shan is somewhat typologically unusual because SVO languages are less likely to have internally headed relative clauses: Cole (1987) had proposed that IHRCs were only found in OV languages. Recently, IHRCs have been identified in verb initial languages like Seediq and Tagalog (Aldridge 2004) and SVO like Buli (Hiraiwa 2003). WALS counts 58/580 OV languages and 5/608 VO languages as having IHRCs (Dryer & Haspelmath 2013).

Section 2 gives several examples of internally headed relative clauses in Shan. Section 3 goes over the previously described typology of internally headed relative clauses, focusing on the issues of maximality and island sensitivity. Section 4 demonstrates that Yunn Shan IHRCs are island sensitive and non-maximal which complicates the typology of IHRCs. Section 5 describes the IHRC data of Navajo, which seems to have relative clauses of the same type as Shan. Section 6 sketches a head raising analysis and concludes.

2 Yunn Shan relative clauses

Example 2 shows the strict SVO word order of Shan. Lik ‘book’ in 2 could be definite or indefinite because Shan lacks overt articles. As 3 shows, it is also a classifier language.

(2) háw hǎn lik. 1.SG see book ‘I see the/a book.’
(3) háw hǎn máa sǎam tó. 1.SG see dog three CL.ANIM ‘I see three dogs.’

In order to distinguish between internally and externally headed relatives, I use: (i) the complementizer, Pǎn, to identify the left edge of the clause, and (ii) wánñâj ‘today’, to identify the right edge of the clause. In 4a, the head kǎj ‘chicken’ appears in the usual position for objects. It is clearly inside of the clause because wánñâj ‘today’, which is also inside the relative clause, appears to the right of the head. In 4b, the head is clearly outside of the relative clause since it appears to the left of the complementizer Pǎn.

(4) Extracted object as matrix subject
   a. [Pǎn Nan Li sú kǎj wánñâj] máan pěn ści khǎaw.
      COMP Nan Li bought chicken today 3.SG be color white
      ‘The/A chicken Nan Li bought today was white.’ (IHRC)
Examples 4a-b show that the head of the relative clause can function as the matrix subject, and 5a and 5b show examples where the head is functioning as the matrix object. From these examples it is possible to see that the Shan IHRC is not simply a topicalization structure.

(5) Extracted object as matrix object

a. Saj Kham ꞌayı [ꞌӓn Nan Li stī máa ꞌaj wānnāj].
   Saj Kham take COMP Ms. Li buy come chicken today
   ‘Saj Kham took the/a chicken Nan Li bought today.’ (IHRC)

b. Saj Kham ꞌayı [ꞌaj, ꞌӓn Nan Li stī máa ꞌaj wānnāj]).
   Saj Kham take chicken COMP Ms. Li buy come today
   ‘Saj Kham took the/a chicken Nan Li bought today.’ (Post-Head)

Many of my examples were elicited and tested in a constructed context, but 6a is an example from a fable my consultant told. Here the head of the relative clause is svkho ꞌhàw ‘clothes’, and it appears in typical object position rather than in front of the relative in a post-head construction.

(6) Extracted object: Example from story

a. Luk ꞌō kāʔ ꞌążhankó mán stī ꞌaw [ꞌӓn mán ꞌeax svkho ꞌhàw].
   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,’ (IHRC)

It is also possible to extract the subject of the relative clause, as in 7a-b, or the object of a ditransitive as in 8a-b.

(7) Extracted Subject

a. [ꞌӓn ꞌaj cǐn ꞌhaw jù nāj] mán pēn sī ꞌhàaw.
   COMP chicken eat rice IMPF this 3.SG be color white
   ‘The/A chicken eating rice is white.’ (IHRC)
b. [kàj, [ʔân t, čǐn kʰaw jù] nāj] mán pên sǐ khāaw.
   chicken COMP eat rice IMPF this 3.SG be color white
   ‘The/A chicken eating rice is white.’

3 Island sensitivity and maximality: IHRC typology

Some analyses of IHRCs have noted an apparent correlation between maximal interpretations and wh-island constraints (Grosu 2002, Watanabe 2004), leading them to categorize IHRCs as one of two types: (i) the Lakhota type: restrictive/non-maximalizing, island insensitive; and (ii) the Quechua and Japanese type: maximalizing, island sensitive.

3.1 Islands: CNPC

Lack of island sensitivity has typically been used as evidence that syntactic A’-movement is not taking place in these constructions and that instead binding is responsible for the available interpretation (Bonneau 1990, Grosu 2000, Watanabe 2004). Languages like Lakhota are not sensitive to islands, as 9 shows. Here, the head, wowapi ‘paper’, is being extracted out of a relative clause that has a different head (i.e. wichota ‘many people’). Languages like Japanese have IHRCs that are sensitive to islands, as shown in 10. In this example, the head, subarashii ronbun ‘excellent paper’, is being extracted out of a relative clause whose head is hito ‘person’. This extraction out of another relative clause is not possible in Japanese, as it violates the Complex Noun Phrase Constraint (CNPC)
Mary Moroney

identified by Ross (1967).

(9) 

many-people paper a read PL ind with we-speak PL the that L.A. Times be
‘The newspaper that we talk to many people who read (it) is the L.A. Times.’

(Lakhota: Williamson 1987, cited in Watanabe 2004: (8))

(10) *[John-ga [subarashii ronbun-o kaita hito]-o homete-ita no]-ga
John-NOM excellent paper-ACC wrote person-ACC praised-had COMP-NOM
shuppan-sareta.
publish-PASS
‘An excellent paper which John had praised the person who wrote (it) was pub-
lished.’

(Japanese; Watanabe 2004: (4b))

Since languages fall neatly into these patterns, researchers have explained the difference in
sensitivity to the CNPC using the differences in on how the relative clause is formed: often,
movmenet versus binding. This paper focuses on languages that display island sensitivity.

3.2 Maximality

An internally headed relative clause can be interpreted maximally or non-maximally. The
distinction between a maximal and non-maximal interpretation corresponds approximately
to having a definite determiner ‘the’ modifying the head of the relative in English, as
shown in 11-12. 11 has a non-maximal interpretation because it is not the case that Bill
must have eaten all of the apples Susan bought. 12 has a maximal interpretation: it would
be infelicitous unless Bill ate all of the apples Susan bought.

(11) Bill ate apples that Susan bought.

(12) Bill ate the apples that Susan bought.

The figure below summarizes these interpretations. The maximal interpretation is obliga-
tory when the set described by the relative clause (e.g., the apples that Susan bought) has to
be co-extensive with the set described by the matrix clause (the apples eaten by Bill). The
maximal interpretation is the case where all of the elements in the relative clause must be in the set described by the matrix clause.

<table>
<thead>
<tr>
<th>Non-maximal</th>
<th>Maximal</th>
</tr>
</thead>
<tbody>
<tr>
<td>set described by matrix clause</td>
<td>set described by matrix clause</td>
</tr>
<tr>
<td>set described by RC</td>
<td>set described by RC</td>
</tr>
</tbody>
</table>

Analyses derive the maximal interpretation of IHRCs in various ways: an e-type analysis (Shimoyama 1999), feature checking with a null D of a wh-phrase (Watanabe 2004), quantificational disclosure (Grosu & Landman 2012), and trace conversion (Erlewine & Gould 2016). Grosu and Landman (1998) noted that this difference in maximal/non-maximal IHRCs correlates to the presence (Lakhota) or absence (Quechua/Japanese) of an overt D, but this has been challenged by Gur languages, which have definite articles and maximalizing IHRCs (Hiraiwa et al. 2017). A more robust correlation has been claimed to be the correlation between island sensitivity (sensitivity to CNPC) and the maximal interpretation of the internally headed relative clause (Grosu 2002, Watanabe 2004, Hiraiwa et al. 2017).

Lakhota, the language whose internally headed relative clauses are not island sensitive, also has non-maximalizing internally headed relative clauses. In (13), the interpretation of this internally headed relative clause is not maximal. The set described by the matrix clause, the set of apples that the speaker wants, does not have to contain all the elements of the set described by the relative clause, the set of well-washed apples. This sentence is felicitous even if the speaker only wants one of the well-washed apples.\(^2\)

(13) \([\text{Thasp} \text{a} \text{wa} \text{ži} \text{t} \text{a} \text{yuza} \text{ža} \text{pi} \text{cha}] \text{wach}i\)

apple a-IRR well wash PL SM I-want

‘I want an apple (nonspecific) that is well washed.’

(\text{Lakhota}; Williamson 1987, cited in Grosu & Landman 1998: (92a))

\(^2\)It is not clear from the paper whether a maximal interpretation is allowed in this sentence, so the Maximal interpretation in the figure for 13 is marked as ‘?’. 
Quechua is a language whose internally headed relative clauses are island sensitive. The interpretation of its relative clauses has been claimed to be obligatorily maximal, as shown in 14. The set described by the matrix clause, which is the set of horses that are good horses, must have all the same elements as are in the set described by the relative clause, which is the set of two horses that the man bought. It is not possible to use this sentence to say that only some of the horses that the man bought were good horses.

\[(14) \text{[Nuna ishkay bestya-ta ranti-shqa-n] alli bestya-m ka-rqo-n}\]
\[\text{man two horse-ACC buy-PERF-3 good horse-VAL be-PAST-3}\]
\[\text{‘The two horses that the man bought were good horses.’}\]

Unavailable interpretation: ‘Two horses that the man bought were good horses.’

(Quechua; Dayal 1991, cited in Grosu & Landman 1998: (93a))

Japanese is another language with island sensitive internally headed relative clauses. It has also been said to have obligatorily maximal internally headed relative clauses.\(^3\) 15 illustrates this. Here the set described by the matrix clause is the set of cookies brought to the party by Taro. This set must be coextensive with the set of cookies Yoko put in the fridge.

\(^3\)However, there are some, such as Kubota and Smith (2007) who disagree.
Taro-TOP Yoko-NOM refrigerator-LOC cookie-ACC most put-AUX-NM-ACC paatii-ni motteitta.
party-to brought
‘Yoko put most cookies in the refrigerator and Taro brought them, *some to the party.’

(Japanese; Shimoyama 1999, cited in Grosu 2002: (28b))

Thus, Japanese and Quechua seem to have maximalizing, island sensitive IHRCs, and Lakhota seems to have non-maximalizing, island insensitive IHRCs.

<table>
<thead>
<tr>
<th>Summary</th>
<th>Maximalizing</th>
<th>Island sensitive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lakhota</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>Quechua, Japanese</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

4 The case of Yûn Shan

4.1 Sensitivity to islands

Yûn Shan internally headed relative clauses do appear to be island sensitive. The Shan equivalent of the Lakhota example in 9 is not acceptable, as shown in 16.
4.2 Maximality

Given that Shan lacks overt determiners and is sensitive to island constraints, we might expect its IHRCs to be maximalizing. However, its IHRCs appear to allow non-maximal interpretations, giving a counterexample to the correlation between island sensitivity and maximality noted by (Grosu 2002, Watanabe 2004, Hiraiwa et al. 2017).

In 17, the set described by the matrix clause is the set of apples that Nan Li wants to eat, and the set described by the relative clause is the set of apples that Saj Kham will wash. If this were a maximalizing internally headed relative clause, we would expect the follow up sentence, which restricts the number of apples that Nan Li wants to eat to one, to be infelicitous. A maximal interpretation would require Nan Li to want to eat all the apples that Saj Kham will wash.

Similarly, 18 shows that a maximal interpretation is not obligatory. When there is a number modifying the internal head of the relative clause, it restricts the number of elements described by the relative clause but not the number of elements described by the matrix clause. The set described by the matrix clause is the set of apples that Nan Li ate, and the set described by the relative clause is the set of apples that Saj Kham peeled. The
interpretation of this sentence is that Saj Kham peeled three apples and Nan Li ate some of the peeled apples. However, Nan Li does not have to have eaten all three peeled apples.

(18) Nan Li cǐn ṃĕn [ʔăn Saj Kham p`ık màmò săam hwí nâj].
Nan Li eat up COMP Saj Kham peel apple 3 CL.RND this
‘Nan Li ate apples that Saj Kham peeled which are three in number.’

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

<table>
<thead>
<tr>
<th>18</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Non-maximal</strong> ✓</td>
</tr>
<tr>
<td>apples N.L. ate</td>
</tr>
<tr>
<td>☐ ☐ ☐</td>
</tr>
<tr>
<td>apples that S.K. peeled</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>18</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Maximal</strong> ✓</td>
</tr>
<tr>
<td>apples N.L. ate</td>
</tr>
<tr>
<td>☐ ☐ ☐</td>
</tr>
<tr>
<td>apples that S.K. peeled</td>
</tr>
</tbody>
</table>

Finally, it is possible to describe elements in a set of things using an internally headed relative clause and then describe elements of the same set of things with contradictory information without it being infelicitous. This is similar to the Consistency test used by Dayal (2004) to identify the true definite determiner of a language. In 19, the internally headed relative clause describes the set of beans on the white book. The matrix clause is the set of black beans. If this were interpreted maximally, we would expect all the beans on the book to be black. However, it is possible to follow up that sentence with one that says that some of the beans are not black. This suggests that a non-maximal interpretation is possible. Importantly, in this context, there are no beans anywhere except on the white book, so it is not possible for ‘some beans’ to be describing other beans in the context. Also, this sentence is not being interpreted generically in the context. The interpretation is that there are black beans and non-black beans on the white book.
Yūn Shan does not fit with the typology that connects island sensitivity of IHRCs with a maximal interpretation. It has internally headed relative clauses that are sensitive to island constraints, but it also allows non-maximal interpretations of its internally headed relative clauses.

<table>
<thead>
<tr>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximalizing Island sensitive</td>
</tr>
<tr>
<td>Lakhota</td>
</tr>
<tr>
<td>Quechua, Japanese</td>
</tr>
<tr>
<td>Shan</td>
</tr>
</tbody>
</table>

5 The case of Navajo

Navajo seems like a good candidate to fit with the Shan IHRC type, with one caveat. Grosu (2012) amends his typology by introducing data found in Navajo. Navajo, like Shan, has IHRCs that are non-maximal but sensitive to island constraints, as in 20-21. The English translation of 20 uses an indefinite to modify the head, which is not compatible with a maximal interpretation. 21 has the head ˈtɛɛchqɪ́ ‘dog’ being extracted out of another
relative clause with the head hastiin ‘man’. This is not possible in Navajo, so it is clearly sensitive to the CNPC.

(20) [Bilasana hazho’o tanasgis-ígíí] nisin. apple carefully washed-REL 1-want ‘I want an apple that is well washed.’ (Navajo; Grosu 2012: (4))

(21) *[Hastiin ́éécháaí bishxash-é] be’eldooh néidiitá-(n)-é naha’ín. man dog bit-REL gun pick-up-REL bark ‘The dog that the man who was bitten by (it) picked up the gun is barking.’ (Navajo; Grosu 2012: (49b))

Strong quantifiers in Navajo seem to obligatorily be interpreted with matrix clause scope, as in 22. If altso ‘all’ were interpreted inside of the clause, it would have the interpretation that John bought all the cars (and motorcycles) from Bill.

(22) [John Bill chidí t’áá altso (dóó dzi’izí dilchxoshí t’áá altso) yaa John Bill car 3 all and motor cycle 3 all from nayisnii’cé] t’éiyá nizhónígo nidaajeeh. 3.3.buyP.REL only well da.3.run.1 ‘All the cars (and all the motorcycles) that John bought from Bill —and only those —run well.’ (Navajo; Grosu 2012: (48))

Grosu analyzes this internally headed relative clause as being a case of ‘cyclic re-merger’ until the head is outside of the relative clause and claims that the source of island sensitivity is ‘traceable to whatever factors require matrix scope for IHs [internal heads]’ (Grosu 2012: 25). This suggests that there is a connection between the matrix scope-taking quality of some Navajo quantifiers with the formation of internally headed relative clauses. Grosu notes that this idea fits with Hastings’s (2004) claim that Cuzco Quechua IHRCs are restrictive when strongly quantifying and maximalizing otherwise. Hastings (2004) had claimed that strongly quantified heads in Cuzco Quechua internally headed relative clauses obligatorily take matrix scope.

While it would be nice to be able to connect these two observations, it cannot be the whole story: in Yűn Shan, quantifiers seem to take surface scope obligatorily. This can be seen in the contrast between 23 and 24. In 23 is an externally headed relative clause with a
Mary Moroney

strong quantifier *mv³ût ‘all’ modifying the head in the matrix clause. The interpretation of this sentence is that Saj Kham peeled an unspecified number of apples, and Nan Li ate all of those peeled apples. The quantifier takes scope in the matrix clause where it appears. In 24, on the other hand, the quantifier appears inside the relative clause, and the head is either inside or outside the clause. The interpretation when the quantifier is inside the clause is that Saj Kham peeled all the apples (12 in this context), and Nan Li ate some unspecified number of peeled apples. The quantifier takes scope inside the relative clause where it appears. There is a clear difference in meaning that corresponds to the position of the quantifier, not the position of the head.

(23) Nan Li cí [mámmó mv³út [ʔān Saj Kham p’ik] ᵃn nâj].
    Nan Li eat apple  all COMP Saj Kham peel CL this
    ‘Nan Li ate all the apples that Saj Kham peeled.’  (Shan; Post-Head)

    ● Number of apples Saj Kham peeled: some number of apples
    ● Apples Nan Li ate: all the apples peeled by Saj Kham

(24) Nan Li cí p’en [(mámmō) [ʔān Saj Kham p’ik (mámmō) mv³út] nâj].
    Nan Li eat up apple  COMP Saj Kham peel apple all this
    ‘Nan Li ate apples that Saj Kham peeled all of.’  (Shan; Post-Head/IHRC)
    (Context: there are 12 apples)

    ● Number of apples Saj Kham peeled: 12
    ● Apples Nan Li ate: some number of the peeled apples

This data is motivation for an analysis of internally headed relative clauses found in Shan and Navajo where the interpretation is non-maximal but the syntax is island sensitive. While Grosu (2012) has suggested that the analysis could be connected to the matrix scope taking qualities of some quantifiers in Navajo and Cuzco Quechua, that cannot be the whole story since Shan quantifiers obligatorily take scope where they appear. Further, this type of internally headed relative clause is only available in Cuzco Quechua when the head is strongly quantified, but Navajo internally headed relative clauses are all non-maximal/restrictive and island sensitive. It is not clear how an analysis that relies on strongly quantified heads can account for the examples that lack strong quantification.
Mary Moroney

This data from Yùn Shan does not fit previous analyses for IHRCs that (1) require a maximal interpretation (Shimoyama 1999, Grosu & Landman 2012, Erlewine & Gould 2016); (2) employ un-selective binding to explain IHRCs that are not subject to wh-island constraints (Watanabe 2004); or (3) rely on strongly quantified heads to explain the availability of non-maximal IHRCs (Grosu 2012).

6 Sketch of Analysis: Head raising

The island sensitivity of IHRCs in Shan, Japanese, and Navajo suggests that there is movement involved, so I will propose LF movement of the head to SpecCP. I will assume a head raising analysis, following the theory proposed Kayne (1994) and Bianchi (1999), and revised by De Vries and colleagues (2002). The head being in SpecCP can explain the island sensitivity of Shan IHRCs. Since it is not necessary to account for maximality, no special analysis is required to account for the interpretation here.

(25) \[[CP \text{ mám}s_i \text{ [CP } \text{ ?án [IP Saj Kham] [VP [V' pýk [CIHp t_i [CIHp sám [CIh' hwí [NP t_i ]]]]]]]]]

An interesting thing to consider is how to integrate the analysis for Yùn Shan IHRCs with the analyses of other island sensitive IHRCs, as are found in Japanese and Navajo, other than simply having a null definite operator appearing where necessary in Japanese and Navajo. The analysis in Grosu & Landman 2012 could be used to account for this data. Grosu and Landman (2012) use event semantics and operator movement instead of covert head movement. The definiteness of Japanese IHRCs is accounted for using a null definiteness operator that is above -no in the syntax (Grosu & Landman 2012: 179). Shan could simply lack this null definite operator. Still, finding a more principled explanation for the appearance of a null definite operator is important for future research.

REFERENCES

Mary Moroney


Mary Moroney


