INTERNALLY-HEADED RELATIVE CLAUSES IN HAKHA CHIN

July, 2019

Stefon Flego

1Indiana University Department of Linguistics

ABSTRACT

Hakha Chin, an underdocumented Tibeto-Burman language, is reported to have internally-headed relative clauses (IHRCs), a typologically rare syntactic structure in which the head noun phrase surfaces within the relative clause itself. The current study provides new data and novel observations which bear on several outstanding issues regarding IHRCs in this language: 1) Relativization of locative and instrumental adjuncts in IHRCs is avoided. 2) Conflicting stem allomorph requirements of negation and relativization of non-subjects give rise to optionality in stem choice when the two are brought together in an IHRC. 3) To relativize an indirect object, an IHRC is either avoided altogether, or the noun phrase is fronted to the absolute left-most position in the embedded clause. 4) Relativization of NPs with a human referent in an IHRC exhibit relativizer gender agreement, which has not been previously reported for this clause type in Hakha Chin.

Key Terms— internally-headed relative clause, relative clause, Hakha Chin, Laiholh, Tibeto-Burman

1. INTRODUCTION

Internally-headed relative clauses (IHRCs) are typologically rare; for example, of the 824 languages for which the order of head noun and relative clause (RC) is established in Dryer (2013), only 24 (3%) use an IHRC as their dominant relative clause type. Furthermore, as many of the languages which are reported to have IHRCs are themselves understudied (e.g. indigenous languages of North America, Tibeto-Burman languages, signed languages), the body of literature informing our understanding of this clause type is still relatively small. Languages which have been reported to have IHRCs include Navajo (Platero 1974), Lakhota (Williamson 1987), Diegueño (Gorbet 1976), Mojave (Munro 1976), Japanese (Hiraiwai 2017), and a variety of signed languages (Wilbur 2017). IHRCs are also quite prevalent in Tibeto-Burman languages (Dryer 2008, Subbarao 2008), such as Tenyidie (Subbarao and Kevichüsa 1999), Sema (Subbarao and Kevichüsa 2005), Qiang (Huang 2008), and Hakha Chin (Kathol and VanBik 1999). Hakha Chin is an underdocumented Central Kuki-Chin language spoken by approximately 140,000 people in Burma/Myanmar (Simons and Fennig 2018), and by many of the over 25,000 Burmese refugees living in Indiana (“College Enrollment Rate,” 2018). The purpose of this study is to collect and present data on IHRCs in Hakha Chin, with the dual aim of contributing to the literature on this clause type and of providing further documentation of this language.

In contrast to an externally-headed relative clause (EHRC), an IHRC is one in which the relativized noun phrase (RNP) remains structurally inside of the embedded clause instead of occurring
outside of it, in the matrix clause (MC). The examples from Japanese given in (1a) and (1b) illustrate both constructions. An EHRC is shown in (1a), in which the RNP, ringo ‘apple,’ shown in blue bolded text, occurs structurally outside of the RC (outside of the innermost set of brackets). The corresponding IHRC is shown in (1b), in which the RNP surfaces within the relative clause. Both sentences have the same meaning.

(1) Externally and Internally-Headed RCs in Japanese

   Naomi-TOP Ken-NOM buy-come.BEN-PAST apple-ACC eat-PAST

   Naomi-TOP Ken-NOM apple-ACC buy-come.BEN-PAST c-ACC eat-PAST

   Naomi ate the apple that Ken bought for her. (Hiraiwa 2017:2041)

One issue that is cross-linguistically common to IHRCs is that they are potentially ambiguous with respect to which noun phrase in the RC is to be interpreted as its head (i.e. which is the RNP), and many languages that have IHRCs utilize one or more strategies for disambiguation (Basilico 1996, Hiraiwa 2017). The sentences from Diegueño given in (2) serve to illustrate: because multiple NPs (here ‘rock’ and ‘dog’) occur inside the embedded clause in (2a), it is not clear which is to be interpreted as the RNP (i.e. whether it is the rock or the dog that is black). To avoid ambiguity, the NP which is to be interpreted as the head of the RC, here ‘rock’, can be fronted in Diegueño, leaving a pronoun, n’im ‘that,’ occupying its base position, as seen in (2b).

(2) Ambiguous Relativization in Diegueño

a. [xatkocok-∅ wi:m tuc]-pu-c n’iLy
   dog-OBJ rock.COMIT I.hit-DEM-SUBJ black

   The rock I hit the dog with was black.
   The dog I hit with the rock was black.

b. [‘wil’ ‘xa-t’ n’i-m ‘tu;]-pu-c n’iL\-cis
   rock dog-OBJ that-COMIT I.hit-DEM-SUBJ black.indeed

   The rock I hit the dog with was black. (Gorbet 1976:51-53)

Both sentences in (2) contain IHRCs and can express the same meaning, but the second is considered unambiguous. As will be shown in the following section, similar strategies for disambiguation of IHRCs have been reported for Tibeto-Burman languages.

Although the existence of IHRCs in Hakha Chin has been established (Kathol and VanBik 1999 & 2002, Peterson 1998 & 2017), many questions about them—and about disambiguation strategies in the language—remain open. For example, although it is well-documented that case marking and verb stem allomorphy affect the interpretation of IHRCs, it is not known whether fronting or avoidance of an IHRC altogether are ever preferred (or, in the case of fronting, possible) methods for disambiguation. In addition, it has been reported that verbal negation and relativizing a non-subject NP make conflicting demands on verb stem allomorphy (Kathol 2001, Kathol and VanBik 1999, Lehman 1996, Peterson 2017), but it is still unclear which verb stem should surface in a negated IHRC in which a non-subject is to be interpreted as the RNP. Kathol and VanBik (1999 & 2002) and Peterson (2017) have shown that
verbal arguments can be relativized in an IHRC, but it is not known how and if instrumental and locative NPs can also be internally relativized. To address these gaps, relative clause data were elicited from a native speaker of Hakha Chin.

The rest of this paper is as follows. Section III summarizes what has been reported for IHRCs in Hakha Chin and other Tibeto-Burman languages. Section IV outlines a methodology for eliciting further data using a preexisting fieldwork questionnaire targeting RCs. In Section V, I present selected data to address the research questions posed above, and section VI concludes the paper. In addition, the full set of elicited data is included at the end of this paper in the form of an appendix.

The main findings are briefly summarized here. It was found that fronting can be employed to relativize an indirect object in an IHRC, but that relativization of instrumental and locative adjuncts in IHRCs is avoided. In addition, when relativizing the non-subject of a negated verb in an IHRC, either of the two possible verb stems is acceptable. Because so little data on this clause type is available for Hakha Chin, much of the elicitation carried out in the current study was necessarily exploratory, and several unexpected aspects of the behavior of IHRCs were encountered. For example, relativization of noun phrases with human referents often requires gender agreement on the relativizer, which has not yet been reported for this clause type in Hakha Chin. And finally, internal relativization of RNPs with one specific combination of syntactic roles in the RC and MC seems to be avoided.

2. IHRCS IN HAKHA CHIN AND OTHER TIBETO-BURMAN LANGUAGES

Although IHRCs are thought to be cross-linguistically rare, they are relatively common in Tibeto-Burman languages (Dryer 2008, Subbarao 2008). Sema, Tenyidie, and Hakha Chin are all Tibeto-Burman languages of the Kuki-Chin group (classification according to Dryer 2008) reported to have this type of clause structure. All three have OV word order, as is typical for languages which have IHRCs (Cole 1987), and all three allow for EHRCs in addition to IHRCs. In Sema, spoken in the state of Nagaland, fronting of the RNP can affect the interpretation of IHRCs (Subbarao and Kevichüsa 2005). With ditransitive verbs, ambiguity arises as to whether the indirect object (IO) or the direct object (DO) is the RNP when they surface in canonical word order in an IHRC, as shown in (3). This sentence contains an IHRC, as all NPs which have a syntactic role in the RC occur within the RC itself (i.e. to the left of the nominalizer -ke-, marked in red).

(3) Ambiguous Relativization of Non-Subjects of Ditransitive Verb in Sema

nO - nO timi yesI (pewo) tsl- ke- u- ye iOnO khuSuwo
you [+TR] person letter ACC give NOZ DEF [-TR] very tall/long

The person you gave a letter to is very tall.
The letter you gave to the person is very long. (Subbarao and Kevichüsa 2005)

However, linear precedence does play a role in interpreting an IHRC: if a non-subject argument is moved leftward of its canonical position within the clause (as in Diegueñño, (2b) above), it is interpreted as the RNP. This is shown in (4a) and (4b), where the DO and the IO are relativized through leftward movement, respectively.

1 Several examples have been found which challenge this generalization (see Hiraiwa 2017 for spoken languages, and Wilbur 2017 for signed languages).
(4) Unambiguous Relativization of Non-Subjects of Ditransitive Verb in Sema

a. nO - nO yesI timi (pewo) tsl- ke- u- ye ilOnO khuSuwo
   you [+TR] letter person ACC give NOZ DEF [-TR] very tall/long
   The letter you gave to the person is very tall.

b. timi nO - nO yesI (pewo) tsl- ke- u- ye ilOnO khuSuwo
   person you [+TR] letter ACC give NOZ DEF [-TR] very tall/long
   The person you gave a letter to is very tall.

(Subbarao and Kevichüsa 2005)

In Tenyidie, a dialect of Angami Naga, both word order and avoidance of IHRCs are used to circumvent possible ambiguity (Subbarao and Kevichüsa 1999). For example, the second constituent in an RC in Tenyidie is usually interpreted as the RNP. So although IO-DO is considered the unmarked word order in this language, as in other South Asian languages (Subbarao 2008), the DO can precede the IO if it is to be interpreted as the RNP, similar to the example from Sema in (4a). Additionally, IHRCs are avoided when the NP in a time expression is to be relativized; only an EHRC is possible in this situation in Tenyidie.

So far it has been shown that languages with IHRCs typically have OV word order and often employ strategies such a fronting to avoid the ambiguity that often arises in this clause type. Attention is now turned to Hakha Chin, the language under investigation in this paper. This language has (at least) three relativizers: 1) mi, the most often encountered, which can relativize any argument of the verb in the RC, 2) tu, which relativizes only ergative arguments of the verb in the RC, and 3) nak, which relativizes locative or instrumental adjuncts in the RC (Kathol and Van Bik 1999, Peterson 1998:119). Both EHRCs and IHRCs occur in Hakha Chin, as shown in (5). EHRCs are pre-nominal (i.e. they precede the RNP) as is typical for Tibeto-Burman languages (Dryer 2008), shown in (5a). However, the RNP can also surface within the RC, in which case it occurs to the left of the relativizer (here mi, marked in red) and the verbal complex of the embedded clause, as shown in (5b). Both sentences have the same meaning.

(5) Externally and Internally-Headed RCs in Hakha Chin

a. inn=i a=it mi lothlo-pa ka=hmuh
   house=LOC 3S=sleep.I REL farmer-REL 1S=see.II

b. inn=i lothlo-pa a=it mi ka=hmuh
   house=LOC farmer-REL 3S=sleep.I REL 1S=see.II

I saw the farmer who slept in the house. (Peterson 2017)

Verb stem allomorphy and case marking have also been shown to play important roles in the interpretation of IHRCs in Hakha Chin. Most verbs have two forms which appear in complementary syntactic contexts, typically referred to in the literature on Hakha Chin as stem I and stem II (Hyman and VanBik 2002, Lehman 1996, Melnik 1997, Patent 1997, Peterson 1998). The phonetic difference

---

2 The following glossing conventions and abbreviations are used in this paper: relative clause (RC), externally-headed relative clause (EHRC), internally-headed relative clause (IHRC), matrix clause (MC), noun phrase (NP), relativized noun phrase (RNP), 1st person (I), 2nd person (II), 3rd person (III), singular (SG), plural (PL), nominative (NOM), accusative (ACC), ergative (ERG), locative (LOC), instrumental (INSTR), relativizer (REL), stem I (I), stem II (II), reflexive (REFL), subject (SUBJ), direct object (DO), indirect object (IO), transitive (TR), intransitive (INTR), and ditransitive (DITR).
between these two forms is one of idiosyncratic coda allomorphy, with stem II very often ending in a glottal stop or glottalized sonorant. The choice of stem is usually closely tied to the valency of the verb, but these associations can be overridden in an RC depending on whether a subject or non-subject is to be interpreted as the RNP (Kathol 2001, Kathol and VanBik 1999 & 2002, King 2009, Peterson 2017). Stem I, normally associated with intransitivity, is used with transitive verbs in IHRCs when the subject is relativized (an example of what King (2009) calls ‘subject focus’), as with bat ‘hang.I’ in (6).

(6) Relativization of Subject of Transitive Verb in Hakha Chin

\[
\begin{align*}
nikum=i & \quad lothlo-pa & \quad thil & \quad a=bat & \quad mi & \quad ka=hmuh \\
last.year=LOC & \quad farmer-M & \quad thing & \quad 3s=hang.I & \quad REL & \quad 1s=see.II
\end{align*}
\]

I saw the farmer who hung up the clothes last year. \hspace{1cm} \text{(Peterson 2017)}

In addition, if the subject of an IHRC is to be interpreted as the RNP, the normally cliticized ergative marker nih does not appear, as in (6). However, if the ergative marker nih is present, a non-subject must be interpreted as the RNP, as in (7). This sentence also illustrates that stem II of the verb is required when a non-subject is be interpreted as the head (an example of King’s (2009) ‘object focus’), as with bah ‘hang.II’.

(7) Relativization of Direct Object of Transitive Verb in Hakha Chin

\[
\begin{align*}
nikum=i & \quad lothlo-pa=nih & \quad thil & \quad a=bah & \quad mi & \quad ka=hmuh \\
last.year=LOC & \quad farmer-M=ERG & \quad thing & \quad 3s=hang.II & \quad REL & \quad 1s=see.II
\end{align*}
\]

I saw the clothes (that) the farmer hung up last year. \hspace{1cm} \text{(Peterson 2017)}

Apart from RCs, there are also other syntactic contexts in which one or the other verb stem is required. For example, stem I is used before the negative particle lo (Lehman 1996). As no examples of negated IHRCs are given in the previous literature on RCs in Hakha Chin, one wonders how the conflicting requirements of negation (stem I) and for relativizing non-subjects (stem II) interact, and if this could have an effect on which constituent is interpreted as the RNP, or whether IHRCs of this type are avoided altogether.

As in Sema ((3) above), it is similarly reported that ambiguity arises in Chin IHRCs when there are multiple non-subject NPs, as is often the case with ditransitive verbs. In (8), it can be seen that either the IO or DO can be interpreted as the RNP (i.e. it is ambiguous as to whether it is the pig or the food that is being seen).

(8) Ambiguous Relativization of Non-Subjects of Ditransitive Verb in Hakha Chin

\[
\begin{align*}
lothlo-pa & \quad vok & \quad rawl & \quad a=pek & \quad mi & \quad ka=hmuh \\
farmer-M & \quad pig & \quad food & \quad 3s=give.II & \quad REL & \quad 1s=see.II
\end{align*}
\]

I saw the pig which the farmer gave food to.

I saw the food which the farmer gave to the pig. \hspace{1cm} \text{(Kathol and VanBik 1999)}

Fronting (leftward movement) of constituents has been reported to have the ability to disambiguate such IHRCs in Diegueño, Sema, and Tenyidie, yet this type of strategy has not been described for Hakha Chin. It is also not clear how often potentially ambiguous IHRCs, such as in (8), are simply avoided in favor of unambiguous EHRCs.
Although previous work on IHRCs in Hakha Chin has provided description of relativization of a variety of verbal arguments within the embedded clause, the question of whether instrumental and locative adjuncts can be relativized within an IHRC remains open. The handful of examples of internally-relativized verbal arguments available in the previous literature also show limited variety in the syntactic role of the RNP in the matrix clause (usually the object of the verb ‘see,’ as in examples (5)-(8) above). Given how little we know about this construction, it would seem desirable to have examples of IHRCs in which the RNP has a variety of syntactic roles in both the RC and the MC. By utilizing the methodology outlined in the next section, I collected data as a first step towards addressing the issues framed in this section.

3. METHODS

Sentences from an RC questionnaire developed for Bantu languages by Downing et al. (2010) served as initial prompts for elicitation. These were often adapted to utilize vocabulary more familiar to the language consultant and to better suit the specific research questions outlined above. The questionnaire includes example sentences in which the RNP has a variety of grammatical functions in the embedded clause, for example, the subject, DO, IO, possessor, instrumental, locative, and temporal adjunct. The verbs in the embedded clause in the example sentences also exhibit a wide variety of syntactic features (e.g. possible reflexive / reciprocal marking, negation, etc.).

Zai Sung, a college-age female native speaker of Hakha Chin, served as a language consultant. The data were collected during 10 1-hour-long elicitation sessions, conducted over the course of 12 weeks. The language consultant was asked to provide Hakha Chin translations of English sentences containing target relative clauses. When she produced sentences containing EHRCs, these were noted, but the given sentences were also manipulated by the author to form an IHRC by placing the RNP structurally inside of the RC, and the consultant was then asked for acceptability judgements on the resulting structure. The consultant was also asked for judgements on other permutations of verb stem allomorphy, word order, and case marking. In this way, from a relatively small elicitation list, a larger set of potential IHRCs was generated, which in turn was culled and vetted by the language consultant, yielding a set of grammatical and ungrammatical sentences containing IHRCs. 67 sentences were collected in total, 29 of which contained EHRCs, 32 contained grammatical IHRCs, and 6 contained ungrammatical IHRCs. The full data set is given in the appendix, where sentences are grouped by their syntactic role of the RNP in the RC, and then sub-grouped by their syntactic role in the MC.

4. DATA & DISCUSSION

The main findings of the investigation are reported in this section. Selected examples from the data set are individually discussed. Each example is given along with its numbering from the appendix, based on the syntactic role of the RNP in both the RC and MC. The following two sets of examples illustrate that IHRCs are avoided when the RNP is a locative (9) or instrumental (10) adjunct. The EHRCs in (9a) and (10a) are grammatical, while the IHRCs in (9b) and (10b), in which the RNP occurs to the left of the relativizer nak, are ungrammatical.
(9) Relativization of Locative Adjunct (G.ii, LOC-SUBJ.REFL)\(^3\)

a. ka=tta-pa=nih \(\text{a=hmu}h\) nak biakinn \(\text{kha a=dawh tuk}\) 
1S=brother-M=ERG doctor-M 3S=see.II REL church TOP REL=pretty INT

\[\text{The church where my brother saw the doctor is pretty.}\]

b. *ka=tta-pa=nih sibawi-pa biakinn \(\text{a=hmu}h\) nak \(\text{kha a=dawh tuk}\) 
1S=brother-M=ERG doctor-M church 3S=see.II REL TOP REL=pretty INT

(10) Relativization of Instrumental Adjunct (H.ii, INSTR-SUBJ.REFL)

a. hngakchia-pa=nih thlalangawng \(\text{a=hrawh} nak\) bawlung \(\text{kha a=dawh tuk}\) 
child-M=ERG window 3S=break.II REL ball TOP REL=pretty INT

\[\text{The ball with which the boy broke the window is pretty.}\]

b. *hngakchia-pa=nih thlalangawng bawlung \(\text{a=hrawh} nak\) \(\text{kha a=dawh tuk}\) 
child-M=ERG window ball 3S=break.II REL TOP REL=pretty INT

(11) Relativization of Direct Object of Ditransitive Verb (E.i, DO.DITR-SUBJ.INTR)

a. sayamah-nu=nih hngakchia-nu \(\text{a=pek} mi cauk\) \(\text{kha a=tla}\) 
teacher-F=ERG child-F 3S=give.II REL book TOP 3S=fall.I

\[\text{The book that the teacher gave to the girl fell.}\]

b. sayamah-nu=nih hngakchia-nu cauk \(\text{a=pek} mi\) \(\text{kha a=tla}\) 
teacher-F=ERG child-F book 3S=give.II REL TOP 3S=fall.I

Although the intuitions of the language consultant in this study differ with previous accounts on this point, many similarities can be observed: for the non-subject ‘book’ to be interpreted as the RNP, stem II of the verb in the embedded clause was required, and the subject was cliticized with the ergative marker \(\text{nih}\).

One of the research questions in this study was on the status of IHRCs in which conflicting demands on verb stem allomorphy would be made, as in the case of relativizing a non-subject and negating the verb. The IHRC example in (12a), similar to (11b) above, shows that stem II of the verb, \(\text{pek}\), is required when relativizing the non-subject ‘book.’ (12b) gives a simplex version of the syntactic

\(^3\) Kha, marked here as a topicalizer, could modify the preceding noun (in the current example, \(\text{biakinn}\) ‘church’), or the entire preceding RC. No claim will be made here as to which of these analyses is more appropriate, since the behavior of \(\text{kha}\) does not seem to bear on the phenomena discussed in this paper. For further discussion, see Wamsley (2019), this volume, where \(\text{kha}\) is treated as a familiarity marker.
material in the embedded clause in (12a), but this time with the verb cliticized with the negative marker lo. It can be seen the stem I of the verb, pe, does indeed surface, as has been previously reported for verbal negation with lo. However, when the negated sentence was used in an IHRC, as shown in (12c), optionality arose between stem I, pe, and stem II, pek.

(12) Relativization of Direct Object with Negation (E.ii, DO.DITR-SUBJ.REFL)

a. sayamah-nu=nih hngakchia-nu cauk a=pek mi kha aa=dawh tuk teacher-F=ERG child-F book 3S=give.II REL TOP REFL=pretty INT
   The book that the teacher gave to the girl is pretty.

b. sayamah-nu=nih hngakchia-nu cauk a=pe=lo teacher-F=ERG child-F book 3S=give.I=NEG
   The teacher didn’t give a book to the girl.

c. sayamah-nu=nih hngakchia-nu cauk a=pe/k=lo mi te kha aa=dawh tuk teacher-F=ERG child-F book 3S=give.I/II=NEG REL TOP REFL=pretty INT
   The book that the teacher didn’t give to the girl is pretty.

For the verb in the negated RC, the language consultant found either stem allomorph acceptable. It should be noted that only stem II was acceptable in the positive sentence in (12a), and only stem I in the simplex version in (12b). However, when conflicting requirements were at work within the sentence, either stem allomorph was accepted. Optionality between thla ‘drop.I’ and thlak ‘drop.II’ also arose in the negated IHRC (given in C.ii in the appendix).

So far, the RNPs discussed in this section have had non-human referents. However, IHRCs in which the RNP has a human referent unexpectedly entail relativizer gender agreement for the language consultant. To my knowledge, this phenomenon has not been reported before for IHRCs in Hakha Chin. Relativizer agreement can be seen in (13b), (14b), and (15b) below (equivalent EHRCs are given in (13a), (14a), and (15a) for comparison). In all of these examples, the gender marker which agrees with the RNP (pa, masculine, nu, feminine) surfaces directly after the relativizer. Relativizer agreement is in most cases completely optional, as in (13b), but sometimes preferred, as in (14b), or even obligatory, (15b).

(13) Relativizer Gender Agreement Optional (A.iv, SUBJ.REFL-DO.TR)

a. inn hmai=ah lente aa=celh mi hngakchia-pa kha uico=nih a=seh house outside=LOC play REFL=play REL child-M TOP dog=ERG 3S=bite
   The dog bit the boy who was playing outside.

b. hngakchia-pa inn hmai=ah lente aa=celh mi (pa) kha uico=nih a=seh child-M house outside=LOC play REFL=play REL (M) TOP dog=ERG 3S=bite
   The notation % (nu) in (14b) signifies that the meaning of the sentence is clear to the listener when gender agreement is omitted, but its presence is strongly preferred. It can also be seen in (14b) that in order to relativize the subject of ‘give,’ stem I of the verb surfaces and the ergative marker nih is not cliticized to ‘teacher,’ which is consistent with previous descriptions of disambiguation of IHRCs in Hakha Chin.
(14) Relativizer Gender Agreement Preferred (D.iv, SUBJ.DITR-DO.TR)

a. hngakchia-pa cauk a=pe mi sayamah-nu kha uico=nih a=seh
   child-M book 3s=give.I REL teacher-F TOP dog=ERG 3s=bite

b. sayamah-nu hngakchia-pa cauk a=pe mi % (nu) kha uico=nih a=seh
   teacher-F child-M book 3s=give.I REL F TOP dog=ERG 3s=bite

The dog bit the teacher who gave the book to the boy.

The notation *(pa) in (15b) signifies that the gender agreement marker cannot be omitted; without it the sentence is considered ungrammatical by the language consultant.

(15) Relativizer Gender Agreement Obligatory (B.i, SUBJ.TR-SUBJ.INTR)\(^4\)

a. thlalangawng a=hrawh tu hngakchia-pa kha a=ttap
   window 3s=break.II REL child-M TOP 3s=cry

b. hngakchia-pa thlalangawng a=khuai tu *(pa) kha a=ttap
   child-M window 3s=break REL M TOP 3s=cry

The boy who broke the window is crying.

It is not entirely clear what factors condition the optionality of relativizer agreement, and this issue will be left to future research. I now turn to relativization of the IO in an IHRC. As already mentioned in the explanation surrounding example (11), the language consultant did not find IHRCs with stem II of ditransitive verbs ambiguous as to whether the IO or the DO is to be interpreted as the RNP, as has been previously reported. In these cases, it is unambiguously the DO which is relativized, as in (11b). The IO can of course be relativized without issue in an EHRC, shown in (16a), but cannot be relativized in the IHRC in its canonical position within the clause. However, the RNP can be fronted to the left-most position in the RC, with optional relativizer gender agreement, as shown in (16b).

(16) Relativization of Indirect Object of Ditransitive Verb (F.iv, IO.DITR-DO.TR)

a. sayamah-nu=nih cauk a=pek mi hngakchia-nu kha uico=nih a=seh
   teacher-F=ERG book 3s=give.II REL child-F TOP dog=ERG 3s=bite

b. hngakchia-nu | sayamah-nu=nih cauk a=pek mi (nu) kha uico=nih a=seh
   child-F teacher-F=ERG book 3s=give.II REL (F) TOP dog=ERG 3s=bite

The dog bit the girl to whom the teacher gave the book.

The language consultant also said that she would take a slight pause after the fronted RNP (represented here with a prosodic boundary | ). It is left to future research to determine whether the RNP in these cases is indeed still internal to the RC, or if it has been fronted externally. The EHRC in (16a) was in any case felt to be the most natural to the consultant, so avoidance of IHRCs when relativizing IOs may be the preferred strategy.

Although the syntactic role of the RNP in the matrix clause did not seem to bear on any of the phenomena discussed so far, there was one unexpected interaction between the syntactic role of the RNP

\(^4\) The language consultant preferred the relativizer \(tu\) in these examples, although she said that \(mi\) was also possible. In the case of (15b), relativizer agreement would still obligatory even after \(mi\).
in the RC and MC: when the RNP was the DO of a transitive verb in the RC and the subject of a transitive verb in the MC, an IHRC was avoided, as shown in (17b). The equivalent EHRC given in (17a) was grammatical, however.

(17) Ungrammatical Interaction of Syntactic Roles in the RC and MC (C.iii, DO.TR-SUBJ.TR)
a. uico=nih a=seh mi hngakchia-pa=nih thlalangawng a=khuai
dog=ERG 3S=bite REL child-M=ERG window 3S=break

b. *uico=nih hngakchia-pa a=seh mi (pa) nih thlalangawng a=khuai
dog=ERG child-M 3S=bite REL (M) ERG window 3S=break

The boy who the dog bit broke the window.

(17b) is ungrammatical regardless of the presence or absence of relativizer gender agreement (pa). It should be noted that ungrammaticality arose in all elicited IHRCs with this specific combination of roles in the RC and MC (see all examples for C.iii in the appendix), even though all other IHRCs in which the RNP was the DO of a transitive verb in the RC (C.i, C.ii, and C.iv), or the subject of a transitive verb in the MC (A.iii, B.iii, D.iii, E.iii, and F.iii) were grammatical.

5. CONCLUSION

Hakha Chin, an underdocumented Tibeto-Burman language of the Central Kuki-Chin branch, is reported to have internally-headed relative clauses, a typologically rare syntactic structure in which the RNP surfaces inside of the relative clause. This study has provided new data and novel observations about IHRCs in Hakha Chin. The results of this study bear on several outstanding questions about RCs in this language. It has been shown that relativization of locative and instrumental adjuncts in IHRCs is avoided, and that these types of RNPs can only surface external to the RC, in the matrix clause. Furthermore, the conflicting stem allomorph requirements of negation and relativization of non-subjects give rise to optionality when brought together in an IHRC, with either stem allomorph accepted.

In addition, the intuitions of the language consultant differed in several respects from those previously reported for Hakha Chin. First, ambiguity does not arise as to whether the DO or IO is to be interpreted as the RNP when ergative marking is present and stem II of the verb is used in an IHRC. In these cases, the DO is interpreted as the RNP. To relativize the IO, an IHRC is either avoided altogether, or the IO can be fronted to the absolute left-most position in the embedded clause. Finally, relativization of NPs with a human referent in an IHRC exhibit relativizer gender agreement, which to my knowledge has not been reported before for this clause type in Hakha Chin.

There still remain many questions about IHRCs in Hakha Chin that were either not addressed in or engendered by this investigation. First, it is not known whether temporal adjuncts can serve as the RNP in an IHRC, or if this relativization strategy is simply avoided along with internal relativization of locative and instrumental adjuncts. Second, the exact nature and scope of relativizer gender agreement in Hakha Chin is uncertain, and it is not known what factors may contribute to making relativizer agreement obligatory in certain cases but optional in others. Third, it is not clear why internal relativization of NPs which serve as the direct object of a transitive verb in the RC and the subject of a transitive verb in the MC is avoided, and whether this restriction extends to any other interactions of syntactic role between the two clauses that were not investigated in this study. These questions are open to future research.
ACKNOWLEDGMENTS

I would like to thank my language consultant Zai Sung, without whose expertise, intuitions, and time this study would not have been possible. I would also like to thank Kelly Berkson, Samson Lotven, James C. Wamsley, and Steven Franks for their insightful comments on previous versions of this work.

REFERENCES


APPENDIX

The full set of elicited data is given in the following sections, organized first by the syntactic role of the relativized noun phrase in the relative clause (A-H), and secondarily by its syntactic role in the matrix clause (i-iv). In example sentences, () indicates that the presence of relativizer agreement is optional, %() indicates that the presence of relativizer agreement is preferred, and *() indicates that the presence of relativizer agreement is obligatory. No marking around relativizer agreement indicates that its optionality was not noted during elicitation.

A. Subject of Reflexive Verb in the Relative Clause

i. Subject of Intransitive Verb in the Matrix Clause

\[\text{inn hmai ah lente aa celh mi hngakchia pa kha a ttap} \] \hspace{1cm} (EHRC)
\[\text{hngakchia pa inn hmai ah lente aa celh mi (pa) kha a ttap} \] \hspace{1cm} (IHRC)

The boy who was playing outside is crying.

ii. Subject of Reflexive Verb in the Matrix Clause

\[\text{inn hmai ah lente aa celh mi hngakchia pa kha aa dawh tuk} \] \hspace{1cm} (EHRC)
\[\text{hngakchia pa inn hmai ah lente aa celh mi (pa) kha aa dawh tuk} \] \hspace{1cm} (IHRC)

The boy who was playing outside is attractive.

iii. Subject of Transitive Verb in the Matrix Clause

\[\text{hngakchia pa inn hmai ah lente aa celh mi (pa) nih khan thlalangawng a hrawh} \] \hspace{1cm} (IHRC)

The boy who was playing outside broke the window.
The boy who fell asleep early didn’t get to see his favorite movie.

The dog bit the boy who was playing outside.

The boy who broke the window is crying.

The boy who broke the window was playing.

I know the boy who threw the ball saw my brother at the pharmacy.
C. Direct Object of Transitive Verb in the Relative Clause

i. Subject of Intransitive Verb in the Matrix Clause

ka tta pa nih sidawr ah a hmu mi sibawi nu kha a ttap
sibawi nu ka tta pa nih sidawr ah a hmu mi (nu) kha a ttap
*ka tta pa nih hngakchia pa kha na dawh tuk
The doctor that my brother saw at the pharmacy is crying.

ii. Subject of Reflexive Verb in the Matrix Clause

ka tta pa nih sidawr ah a hmu mi nu kha aa dawh tuk
The woman that my brother saw at the pharmacy is attractive.

ka tta pa nih sibawi nu sidawr ah a hmu mi (nu) kha aa dawh tuk
The doctor that my brother saw at the pharmacy is attractive.

hngakchia pa nih khedan a thlak mi kha aa dawh tuk
The pencil that the boy dropped is pretty.

hngakchia pa nih khedan a thla/k lo mi kha aa dawh tuk
The pencil that the boy didn’t drop is pretty.

iii. Subject of Transitive Verb in the Matrix Clause

ka tta pa nih sidawr ah a hmu mi pa nih khan thlalangawng a hrawh
*ka tta pa nih hngakchia pa sidawr ah a hmu mi (pa) nih khan thlalangawng a hrawh
The boy who my brother saw at the pharmacy broke the window.

uico nih a seh mi hngakchia pa nih thlalangawng a khuai
*uico nih hngakchia pa a seh mi (pa) nih thlalangawng a khuai
The boy who the dog bit broke the window.

uico nih a seh mi hngakchia pa nih ka tta pa sidawr ah a hmuh
*uico nih hngakchia pa a seh mi (pa) nih ka tta pa sidawr ah a hmuh
The boy who the dog bit saw my brother at the pharmacy.

iv. Direct Object of Transitive Verb in the Matrix Clause

ka tta pa nih sidawr ah a hmu mi hngakchia pa kha uico nih a seh
ka tta pa nih hngakchia pa sidawr ah a hmu mi (pa) kha uico nih a seh
The dog bit the boy who my brother saw at the pharmacy.
D. Subject of Ditransitive Verb in the Relative Clause

i. Subject of Intransitive Verb in the Matrix Clause

\[ \text{hngakchia nu cauk a pe mi nu sayamah nu kha a ttap} \]
\[ \text{The teacher who gave the book to the girl is crying.} \] (EHRC)

\[ \text{sayamah nu hngakchia pa cauk a pe mi %} (\text{nu}) \text{ kha a ttap} \]
\[ \text{The teacher who gave the book to the boy is crying.} \] (IHRC)

\[ \text{sayamah nu hngakchia pa cauk a pe mi %} (\text{nu}) \text{ kha a thau tak} \]
\[ \text{The teacher who gave the book to the boy is fat.} \] (IHRC)

ii. Subject of Reflexive Verb in the Matrix Clause

\[ \text{hngakchia nu cauk a pe mi (nu) sayamah nu kha aa dawh tuk} \]
\[ \text{The teacher who gave the girl a book is pretty.} \] (EHRC)

\[ \text{sayamah nu hngakchia nu cauk a pe mi %} (\text{nu}) \text{ kha aa dawh tuk} \]
\[ \text{The teacher who gave the girl a book is pretty.} \] (IHRC)

iii. Subject of Transitive Verb in the Matrix Clause

\[ \text{hngakchia nu cauk a pe mi sayamah nu nih khan ka tta pa sidawr ah a hmuh} \]
\[ \text{The teacher who gave a book to the girl saw my brother at the pharmacy.} \] (EHRC)

\[ \text{sayamah nu hngakchia nu cauk a pe mi (nu) nih khan ka tta pa sidawr ah a hmuh} \]
\[ \text{The teacher who gave a book to the girl saw my brother at the pharmacy.} \] (IHRC)

iv. Direct Object of Transitive Verb in the Matrix Clause

\[ \text{hngakchia pa cauk a pe mi sayamah nu kha uico nih a seh} \]
\[ \text{The dog bit the teacher who gave the book to the boy.} \] (EHRC)

\[ \text{sayamah nu hngakchia pa cauk a pe mi %} (\text{nu}) \text{ kha uico nih a seh} \]
\[ \text{The dog bit the teacher who gave the book to the boy.} \] (IHRC)

E. Direct Object of Ditransitive Verb in the Relative Clause

i. Subject of Intransitive Verb in the Matrix Clause

\[ \text{sayamah nu nih hngakchia nu a pek mi cauk kha a tla} \]
\[ \text{The book that the teacher gave to the girl fell.} \] (EHRC)

\[ \text{sayamah nu nih hngakchia nu cauk a pek mi kha a tla} \]
\[ \text{The book that the teacher gave to the girl fell.} \] (IHRC)

ii. Subject of Reflexive Verb in the Matrix Clause

\[ \text{sayamah nu nih hngakchia nu cauk a pek mi kha aa dawh tuk} \]
\[ \text{The book that the teacher gave to the girl is pretty.} \] (IHRC)
The book that the teacher didn't give to the girl is pretty.

The cat that the teacher gave to the girl bit my brother.

The dog ate the book that the teacher gave to the girl.

The dog bit the cat that the teacher gave to the girl.

The girl who the teacher gave the book to is crying.

The girl who the teacher gave a book to is pretty.

The girl who the teacher gave the book to broke the window.

The dog bit the girl that the teacher gave the book to.
sayamah nu nih cauk a pek mi (nu) hngakchia nu kha ka hmuh
hngakchia nu | sayamah nu nih cauk a pek mi %(nu) kha ka hmuh
I saw the girl who the teacher gave the book to.

G. Locative Adjunct in the Relative Clause

ii. Subject of Reflexive Verb in the Matrix Clause

ka tta pa nih sibawi pa a hmuh nak biakinn kha a dawh tuk
*ka tta pa nih sibawi pa biakinn a hmuh nak kha a dawh tuk
The church where my brother saw the doctor is pretty.

H. Instrumental Adjunct in the Relative Clause

ii. Subject of Reflexive Verb in the Matrix Clause

hngakchia pa nih thlalangawng a hrawh nak bawlung kha aa dawh tuk
*hngakchia pa nih thlalangawng bawlung a hrawh nak kha aa dawh tuk
The ball that the boy broke the window with is attractive.

iv. Direct Object of Transitive Verb in Matrix Clause

*hngakchia pa nih hai namte a ah nak kha ka hmuh
I saw the knife that the boy cut the mango with.