The Tibetan correlative and its counterparts in Hindi and English

Floris van Vugt
florisvanvugt@ucla.edu

December 10, 2009

Tibetan[1] employs a construction that has been called the *correlative*. Cable (2009) defines the construction as having the following two properties (p.3):

1. an adjunct CP containing a (wh– or relative) operator, and

2. a pronoun or demonstrative phrase[2], occupying an argument position and 'associated with' the aforementioned adjunct CP.

In what follows I will refer to the adjunct CP as the *correlative* CP and the demonstrative phrase as the *correlative* DP. The former surfaces sentence–initially and the latter generally remains in the matrix clause in the default location considering its grammatical function[3]. One of the many examples is given below (C.ex.1)[4].

(1) [$_{CP}$ Khyodra–s gyag gare nyos yod na] nga–s [$_{DP}$ de] bsad pa yin.
you–erg yak what buy aux if I–erg that kill perf aux.
'I killed whatever yak you bought'

[1]The dialect under consideration here is Lhasa Tibetan, the *lingua franca* in the Tibetan Autonomous region. Unfortunately I have no access to a speaker of this dialect, which is why my analysis is entirely based on data from Cable (2009). All transcriptions given are in Wylie transliterationWylie (1959) to reflect ancient pronunciation.

[2]For the sake of simplicity, I will simply consider them to be DPs in the present analysis.

[3]Though pre–verbal word order in Tibetan is essentially free.

[4]Since the Tibetan examples come from Cable (2009) I will simply mark them as C.ex. followed by the example number in that paper. I will also freely add brackets to indicate constituents wherever I think that clarifies the presentation, although they are not always present in the original paper.

An obvious difference with other languages that employ relatives is that 'na' *if*, is also used in conditional constructions, such as (2) (C.ex.24) and which we will not classify as correlatives since no wh–operator occurs in the embedded clause.

(2) [$_{CP}$ Kyodrang Lhasa la 'gro na] nga [$_{DP}$ ∅] 'gro gi yin.
you Lhasa dat go if I PRO go non.past aux.
'If you go to Lhasa, I will go there.'

The example also shows a phenomenon common in Tibetan, which is that pronouns can be null, as shown independently by the following example.

(3) Nga Norbu la dgagi yod. ∅ Gyag bsad pa red.
I Norbu dat good aux. (he) yak kill perf aux.
'I like Norbu. He killed a yak.'

1 Hypothesis

One might ask if we can analyse the Tibetan correlatives as English–like relative clauses(Schachter, 1973; Kayne, 1994) with the correlative CP starting out inside the correlative DP. Under this analysis, the English and Tibetan would have the same underlying form in (4). In English, the head noun moves to the specifier of CP, whereas in Tibetan it would remain *in situ*, with the correlative CP in its turn moving to a sentence–initial position.

(4) (*UF*) I killed [$_{DP}$ the [$_{CP}$ [$_{C'}$ REL you bought yak]]]
(*Tibetan*) [$_{CP_i}$ [$_{C'}$ REL you bought yak]] I killed [$_{DP}$ the t_i]
(*English*) I killed [$_{DP}$ the [$_{CP}$ yak$_i$ [$_{C'}$ REL you bought t_i]]]

2 Hindi correlatives

In order to gain a better understanding of this problem we turn to Hindi, language in which correlatives have been studied in much more depth. An example construction is (M.ex.9)[5].

(5) [$_{CP}$ jo a:dmi: si:ta:–ko pasand he] mujhe [$_{DP}$ vo] acc^{h}a: nahĩ:
REL man Sita–dat like be.pres I–dat DEM nice not
lagtaa.
seem.imp.
'I do not like the man who Sita likes.'

Under Mahajan (2000)'s analysis the underlying structure for this sentence is (after M.ex.29):

(6) mujhe [$_{DemP}$ vo [$_{CP}$ [$_{IP}$ si:ta:–ko jo a:dmi: acc^{h}a: lagtaa he]]]
I–dat DEM Sita–dat REL man nice seem.imp be.pres
pasand nahĩ: he
like not be.pres

Subsequently, the following operations take place:

1. The relative together with the head–N scramble to a IP–specifier position. At this stage the phrase corresponds to a grammatical sentence. In (7) I show only the relative clause for clarity.

   (7) [$_{CP}$ [$_{IP}$ [jo a:dmi:]$_i$ si:ta:–ko t$_i$ acc^{h}a: lagtaa he]]
   REL man Sita–dat nice seem.imp be.pres

2. The head–N optionally moves further to the CP–specifier. Again the sentence is surface–valid.

   (8) [$_{CP}$ a:dmi:$_j$ [$_{IP}$ [jo t$_j$]$_i$ si:ta:–ko t$_i$ acc^{h}a: lagtaa he]]
   man REL Sita–dat nice seem.imp be.pres

3. The DemP is copied to the left of the matrix clause subject.

[5] I will mark examples from Mahajan (2000) as M.ex. followed by the example number in that paper. I will add brackets to indicate certain constituents under the analysis proposed in that paper. Finally the phonological transcription here is simplified.

4. The CP in the copy left behind is deleted. The surface result is (9), assuming that the head–N did not move to the specifier of CP in the earlier steps.

(9) [$_{\text{DemP}}$ vo [$_{\text{CP}}$ [$_{\text{IP}}$ [jo a:dmi:]$_i$ si:ta:–ko t$_i$ acc$^{\text{h}}$a: lagtaa he]]]
DEM REL man Sita–dat nice seem.imp be.pres
mujhe [$_{\text{DemP}}$ vo [$_{\text{CP}}$ [$_{\text{IP}}$ [~~jo~~ ~~a:dmi:~~]$_i$ ~~si:ta:–ko~~ t$_i$ ~~acc$^{\text{h}}$a:~~ ~~lagtaa~~ ~~he~~
I–dat DEM REL man Sita–dat nice seem.imp be.pres
]]] pasand nahĩ: he
like not be.pres

5. To yield the expected output in (5) we further need to allow the initial DEM 'vo' to delete.

Although step 4 speaks of deletion of the CP it should be noted that there is one item that can escape deletion (or be deleted, optionally). Hindi allows the head noun to be present in both the relative clause and the main clause if the former is preposed (i.e. if it has not remained adjacent to the determiner as it is underlyingly)[6]:

(10) [$_{\text{DemP}}$ vo [$_{\text{CP}}$ a:dmi:$_j$ [$_{\text{IP}}$ [jo t$_j$]$_i$ si:ta:–ko t$_i$ acc$^{\text{h}}$a: lagtaa he]]]
DEM man REL Sita–dat nice seem.imp be.pres
mujhe [$_{\text{DemP}}$ vo [$_{\text{CP}}$ a:dmi:]] pasand nahĩ: he
I–dat DEM man like not be.pres

This means it was not the entire CP that was deleted.

Mahajan (2000) cites convincing evidence for each of the intermediate states of the posited movement, showing either directly that they are either grammatical sentences, or that the processes that generate them occur widely. Furthermore, on semantic grounds the account is satisfying since there is a direct link between what we have been calling the correlative CP and DP. More precisely, the CP is underlyingly a sister to the demonstrative head of the correlative DP.

[6]Traces in the embedded clause are omitted.

2.1 How to regulate what is deleted and what is not

But one question remains open. Under the copying theory of movement we have doubled the DemP by moving it up front and subsequently delete most, but not all, of what is left. Clearly the copying theory of movement accomodates deletion of the entire remainder after a making a copy, for allegedly this is what makes a trace unpronounced in e.g.

(11) 'What did you buy ~~what~~?'

But if we broaden our theory to allow partial deletions, should there be elements that we want to force to be deleted? In other words, one then wonders why it is not possible to leave the entire original correlative CP in to start with, to yield the following unacceptable sentence:

(12) *[$_{\text{DemP}}$ [$_{\text{CP}}$ [$_{\text{IP}}$ jo si:ta:–ko acc$^{\text{h}}$a: lagtaa he]]] mujhe [$_{\text{DemP}}$ vo [$_{\text{CP}}$ [$_{\text{IP}}$ [jo a:dmi:] si:ta:–ko acc$^{\text{h}}$a: lagtaa he]]] pasand nahĩ: he

Mahajan (2000)'s response is two–fold.

To prevent too much material from being deleted he assumes there exists a *constraint of identity* (p.216): one instance of a copied item can be deleted only if it is not deleted elsewhere. This is a reasonable assumption governed by some principle that the essential material must be recoverable.

To on the other hand prevent too much material from making it into the phonological representation there can be a sort of *c–command constraint*[7] which would posit that "two copies, x_i and y_i can be spelled–out simultaneously in a representation only if neither c–commands the other"(p.221 of Mahajan (2000) and inspired by Wilder (1995)). This principle would already be present in syntactic theory, he argues, to rule out pronunciation of traces as in (11).

[7]The term is of my making for convenience of presentation.

These two principles account for a number of otherwise puzzling facts in Hindi, such as that the head of the relative head noun cannot be present both in the relative and matrix clause if the former is not preposed. The reason is that in this case the head–N in its raised form c–commands the lower "original" and hence both cannot be pronounced at the same time (if the clause is preposed as illustrated before, this problem does not arise), e.g.[8]

(13) *mujhe [$_{\text{DemP}}$ vo [$_{\text{CP}}$ a:dmi:$_j$ [$_{\text{IP}}$ [jo a:dmi:$_j$]$_i$ si:ta:–ko t$_i$ acc$^{\text{h}}$a: lagtaa
I–dat DEM man REL man Sita–dat nice seem.imp
he]]] pasand nahĩ: he
be.pres like not be.pres

However, I will argue that the c–command constraint by itself is not sufficient. The reason is that we have not excluded sentences such as (12). My consultant assured me that any sentence with the REL 'jo' in both the matrix and relative clause is ungrammatical. Mahajan (2000) does not provide any principle on the basis of which these examples can be excluded. There does not seem to be a structural criterion on the basis of which we can formulate a principle that would prohibit out the REL 'jo' from appearing in both clauses, but not the head–N.

2.2 IP–only deletion hypothesis

An ad–hoc solution based on the data so far could be to consider that it is not the embedded CP left behind after copying that deletes, but rather the IP. In this way, only the head–N could escape deletion since it is the only element that can move into the specifier of CP. The reason why this solution remains unsatisfying is that it raises the question why it would be the IP that deletes and not another arbitrary constituent.

It does yield one empirical prediction, however. If we delete the IP rather than the CP and if it is this that enables the head N to escape deletion, then we predict that the head–N can appear in the main clause only if it occurs *before* REL in the preposed clause,

[8]This grammaticality judgement was provided by Hindi consultant Anoop Mahajan.

since the preposed clause is a copy of the DemP in the main clause[9].

In other words, we would expect that (14-a),(14-b),(14-d) below are grammatical but crucially (14-c) is not, since we can see in the preposed clause that the head–N has not raised out of the IP and hence should have been deleted with the remainder of the IP in the main–clause DemP.

(14) a. [vo a:dmi: jo si:ta:–ko acc^{h}a: lagtaa he] mujhe vo a:dmi: pasand nahi$\tilde{}$: he

b. [vo a:dmi: jo si:ta:–ko acc^{h}a: lagtaa he] mujhe vo pasand nahi$\tilde{}$: he

c. [vo jo a:dmi: si:ta:–ko acc^{h}a: lagtaa he] mujhe vo a:dmi: pasand nahi$\tilde{}$: he

d. [vo jo a:dmi: si:ta:–ko acc^{h}a: lagtaa he] mujhe vo pasand nahi$\tilde{}$: he

However, my Hindi consultant judged (14-c) and actually all sentences as felicitous. Thus we have to reject our IP–only deletion hypothesis as well.

2.3 Summary

In sum, my commentary is that Mahajan (2000)'s account provides a reason why certain information must be deleted (and this principle is employed effectively to rule out the head–N from being in both relative and main clause when the former is not preposed), but doesn't tell us how to force deletion of other material, in particular what part of remaining embedded clause needs to disappear.

3 Hindi analysis for Tibetan

Can we apply the analysis proposed in Mahajan (2000) to the Tibetan correlative?

The underlying form of a sentence such as (1) would contain the wh–operator adjacent to the head–N as well as the 'na' (*if*) marker.

[9]Notice that we are assuming here also that the (optional) raising of the head–N in step 2 happens crucially before copying.

(15) nga–s [$_{DP}$ de [$_{CP}$ Khyodra–s gyag gare nyos yod na]] bsad pa yin.
I–erg that you–erg yak what buy aux if kill perf aux.

In Hindi, step 3, the entire correlative DP would front. In Tibetan, however, there seems to be no reason to assume this. So I propose the correlative CP would move to frontal or quasi–frontal position[10].

(16) [$_{CP}$ Khyodra–s gyag gare nyos yod na]$_i$ nga–s [$_{DP}$ [$_{CP}$ Khyodra–s gyag gare nyos
you–erg yak what buy aux if I–erg you–erg yak what buy
yod na]$_i$ de] bsad pa yin.
aux if that kill perf aux.
'I killed whatever yak you bought.'

Since Tibetan correlatives never appear to exhibit any remainders of the correlative CP material in the correlative DP, we can conclude that the CP–deletion step that Mahajan (2000) posited for Hindi *does* apply without exception in Tibetan.

(17) [$_{CP}$ Khyodra–s gyag gare nyos yod na]$_i$ nga–s [$_{DP}$ [$_{CP}$ ~~Khyodra–s~~ ~~gyag~~ ~~gare~~ ~~nyos~~
you–erg yak what buy aux if I–erg you–erg yak what buy
~~yod~~ ~~na~~]$_i$ de] bsad pa yin.
aux if that kill perf aux.
'I killed whatever yak you bought.'

3.1 Evidence for movement

An argument for this movement's taking place is that a quantificational DP can bind a (null) pronoun inside the raised correlative CP, as exemplified here (C.ex.40), which is grammatical:

(18) [$_{CP}$ $\emptyset_1$ mogmog gare mthong na]$_2$ [mi tshangma–s]$_1$ [$_{DP}$ de$_2$] njo gi
PRO momo what see if man every–erg that buy non.past
red.
aux.
'Every man buys whatever momos[11] he sees.'

[10] Cable (2009) notes that certain subjects can appear before the correlative CP

[11] Traditional Tibetan dumplings.

This shows that the pronoun $\emptyset_1$ needs to underlyingly be c–commanded by 'mi tshangma–s' *every man*, which is indeed the case for the underlying structure in our analysis.

3.2 Empiricial prediction

If indeed the Tibetan correlative is parallel to English and Hindi and our analysis of it is correct, it would seem reasonable to predict that interrogative wh–words do not generally raise in Tibetan. The reason is that the wh–element remains *in situ* in Tibetan whereas in English the wh–element raises in relatives clauses, i.e.

(19) The man whom we saw ~~whom~~.

∗The man we saw whom.

On the basis of this it would seem reasonable to postulate that wh–raising is not required in Tibetan. Indeed standard interrogative phrases confirm this (Tournadre (1996) p.158):

(20) Khyedrang gi rkanggaril ga–par bzhag yod
you gen bike where put aux
'Where did you put your bike?'

It seems furthermore reasonable to assume that in Tibetan the head–N either cannot occur in both the main and relative clause. The reason is that it does not seem to front to the relative clause like in Hindi (or in English for that matter). The reason that Hindi head–N in the main clauses could remain overt was that they escape deletion in some meaningful way that REL does not, the most likely structural reason being its raising to the specifier of CP[12]. If then, as it seems from the data here, Tibetan does not raise the subordinate head–N and REL, it would follow in our current framework that they cannot escape deletion. So in particular sentences like the following (which are possible in Hindi) are predicted to

[12]As it was remarked before, we otherwise have no explanation for why the REL cannot escape deletion in the same way.

be ungrammatical.

(21) [$_{CP}$ Khyodra–s gyag gare nyos yod na] nga–s [$_{DP}$ de gyag] bsad pa yin.
you–erg yak what buy aux if I–erg that yak kill perf aux.
'I killed whatever yak you bought.'

4 Conclusion

I have presented the Tibetan correlative construction, which seems a somewhat restricted counterpart to the Hindi correlative. For the latter, we have investigated Mahajan (2000)'s account that assumes that the correlative CP starts out as a complement of the head determiner of the correlative DP. This analysis succesfully predicts a number of ungrammaticalities pertaining to the case in which the correlative DP does not prepose to the beginning of the sentence, but it does not allow us to formulate a condition that rules out incorrectly spelling out part of the original DP–internal CP. A simplified version of this account can be straight–forwardly applied to the Tibetan correlatives, with the difference that the relative and head noun do not seem to raise to a relative clause–initial position and furthermore instead of the entire DP only the CP appears to then be preposed.

References

Cable, Seth. 2009. The syntax of the tibetan correlative. Pages 195–222 of: Liptak, Aniko (ed), Correlatives cross-linguistically. Amsterdam: John Benjamins.

Kayne, Richard S. 1994. The antisymmetry of syntax. MIT Press.

Mahajan, Anoop. 2000. Relative asymmetries and hindi correlatives. Pages 201–230 of: Artemis Alexiadou, Paul Law, Andre Meinunger, & Wilder, Chris (eds), The syntax of relative clauses. Linguistics Today.

Schachter, Paul. 1973. Focus and relativization. Language, **49**(1), 19–46.

Tournadre, Nicolas. 1996. L'ergativit en tibtain. Bibliothque de l'Information Grammaticale. Louvain – Paris: ditions Peeters.

Wilder, Chris. 1995. Rightward movement as leftward deletion. Pages 273–309 of: Lutz, U., & Pafel, J. (eds), Extraction and extraposition in german. Amsterdam: John Benjamins.

Wylie, Turell. 1959. A standard system of tibetan transcription. Harvard journal of asiatic studies, **22**(December), 261–267.