Focalization and Topicalization in Nepali*

Renato Lacerda, University of Connecticut

renato.lacerda@uconn.edu

Abstract: This paper provides novel data to characterize focalization and topicalization in Nepali, an SOV Indo-Aryan language. Focalization is argued to not trigger a specific syntactic operation (such as focus movement). It is proposed that an element may be focalized in any position it may independently scramble to, and it is shown that no adjacency is required between the focus and the verb (unlike e.g. Hindi). Topicalization, conversely, must involve dislocation, which is taken to indicate that an operator-variable configuration must be created in the syntax. Broadly, the Nepali data suggest that focalization and topicalization should not be treated equally in the mapping from syntax to Information Structure.

Key words: Focalization, Topicalization, Syntax, Information Structure, Nepali.

1.   Introduction

Nepali can be considered a "free word order" language, in that it allows for multiple scrambling possibilities. For instance, all the sentences in the (non-exhaustive) list in (1) are acceptable.

(1)   a.   *raam-le   hijo   pasal-maa   kitaab   kin-yo.*
           raam-ERG   yesterday   store-LOC   book   buy-3SG.MASC
           'Raam bought a book at the store yesterday.'
      b.   *raam-le   hijo   kitaab   pasal-maa   kin-yo.*
           raam-ERG   yesterday   book   store-LOC   buy-3SG.MASC
      c.   *raam-le   hijo   pasal-maa   kin-yo   kitaab.*
           raam-ERG   yesterday   store-LOC   buy-3SG.MASC   book
      d.   *hijo   raam-le   kitaab   pasal-maa   kin-yo.*
           yesterday   raam-ERG   book   store-LOC   buy-3SG.MASC
      e.   *kitaab   raam-le   hijo   pasal-maa   kin-yo.*
           book   raam-ERG   yesterday   store-LOC   buy-3SG.MASC

The fact that the grammar of Nepali allows for a vast array of word order possibilities does not mean that related sentences like the ones in (1) are freely interchangeable. In particular, it has been argued that in a number of "free word order" languages, variations of this sort are often linked to different information structure configurations (see e.g. Butt 2014 for Hindi; Kučerová 2012 for Czech, Russian, and Serbo-Croatian; Şener 2010 for Turkish).[1] In a neutral scenario, as in out-of-the-blue utterances or in response to a question of the sort 'What happened?', the all-new, broad-focus sentence in Nepali clearly reveals an SOV canonical order. For instance, in (2), the answer in B1 is preferred over that in B2.

(2)　　A:　*ke　bha-yo?*
　　　　　　what　happen-3SG.MASC
　　　　　　'What happened?'
　　　　B1:　*raam-le　hijo　　pasal-maa　　kitaab　kin-yo.*
　　　　　　raam-ERG　yesterday　store-LOC　　book　buy-3SG.MASC
　　　　　　'Raam bought a book at the store yesterday.'
　　　　B2:　*#kitaab raam-le　hijo　　　pasal-maa　　kin-yo.*
　　　　　　book　　raam-ERG　yesterday　store-LOC　　buy-3SG.MASC
　　　　　　'The book, Raam bought it at the store yesterday.'

The goal of this paper is to provide an initial characterization of how the information structure processes of focalization and topicalization influence word order variation in Nepali. In particular, I will argue that focalization does not trigger a specific syntactic operation; an element may be focalized in any position it may independently scramble to, and often in situ. Topicalization, on the other hand, must involve dislocation, which is taken as an indication that an operator-variable configuration must be created in the syntax. Broadly, the observation of the Nepali data suggests that focalization and topicalization should not be treated equally in the mapping from syntax to Information Structure.

---

[1] By no means do I intend to claim that information structure is the only factor regulating word order variation in free word order languages, even though factors other than information structure in Nepali will be left aside in this paper.

2. Focalization

I will start this section by discussing (narrow) contrastive focalization. Contrastive focalization is easily observed in contexts of correction, as in (3).[2] The first salient property of corrective sentences is that they make use of a different form of the verb than the one used in the broad-focus counterpart.[3] Given its interpretation, one could hypothesize that (3)B involves some sort of cleft construction. However, that analysis is ruled out by the fact that there is no promotion of the focalized constituent. While in (3)B the direct object is in the immediately pre-verbal position, that does not have to be the case for corrective subjects, as shown in (4)B.

(3)　　A:　　*raam-le　　pasal-maa　　*jacket　*kin-yo.*
　　　　　　　raam-ERG　store-LOC　　　jacket　buy-3SG.MASC
　　　　　　　'Raam bought a jacket at the store.'
　　　　B:　　*ho-ina,　　us-le　　kitaab　**kin-ek-o**　　　　ho.*
　　　　　　　AUX-NEG　3SG-ERG　book　　buy-EK-3SG.MASC　AUX
　　　　　　　'No, he bought *a book*.'

(4)　　A:　　*raam-le　　pasal-maa　　*jacket　*kin-yo.*
　　　　　　　raam-ERG　store-LOC　　　jacket　buy-3SG.MASC
　　　　　　　'Raam bought a jacket at the store.'
　　　　B:　　*ho-ina,　　**hari-le**　　*jacket　*kin-ek-o　　　　ho.*
　　　　　　　AUX-NEG　hari-ERG　jacket　buy-EK-3SG.MASC　AUX
　　　　　　　'No, *Hari* bought a jacket.'

Nepali exhibits a preference for a corrective sentence to be parallel in word order to its corresponding corrected sentence (and accordingly, question-answer pairs), as a change in the word order of the corrective sentence (or answer) could reveal additional informational moves (I will descriptively refer to this as the *corrected-corrective parallelism*). The object being in pre-verbal position in the corrective sentence (3)B thus naturally follows from the corrected sentence

---

[2] Sentences were elicited with the aid of slides on a computer screen, where all participants of the event (and potential alternatives where necessary) were accessible to the consultant.

[3] Acharya (1990) analyses *-eko* as a 'perfective aspect marker'. Although that gloss is compatible with the examples presented here, I will not discuss aspectual properties of corrective sentences and therefore I will remain agnostic about the proper analysis of *-ek* (note that the form *-ek-i* may be used for feminine subjects).

(3)A being in the canonical SOV order. Note that the corrective subject in (4)B also stays in its usual position, showing that, unlike other SOV languages like Hindi (Butt & King 1996, Kidwai 2000) and Turkish (Şener 2010), Nepali does not require adjacency between the focus and the verb. Note that the focalized subject (or any other element, for that matter) can independently end up in the immediately pre-verbal position if the potentially intervening elements all scramble to the its left, as in (5)B.

(5) A: *raam-le   pasal-maa   jacket   kin-yo.*
         raam-ERG   store-LOC   jacket   buy-3SG.MASC
         'Raam bought a jacket at the store.'
     B: *ho-ina,   jacket **hari-le   kin-ek-o   ho.*
         AUX-NEG   jacket hari-ERG   buy-EK-3SG.MASC   AUX
         'No, the jacket, *Hari* bought it.'

The contrastively focalized constituent is thus allowed to remain in situ, and often that is the preferred option. In (6)B, the corrective direct object cannot appear higher than the middle-field locative adverb *pasal-maa* 'at the store'. If it does, the sentence is reported to have an interpretation where it is the locative that is being corrected (presumably with the locative staying in situ and the non-focus direct object scrambling past it). By the same token, if the focalized argument is fronted to the left periphery, as in (7)B, it also leads to an odd result. Contrastively focalized constituents thus seem to be preferably realized in a position parallel to the corrected sentence (although there is also such preference for non-focus elements, it may not be as strong).

(6) A: *raam-le   pasal-maa   jacket   kin-yo.*
         raam-ERG   store-LOC   jacket   buy-3SG.MASC
         'Raam bought a jacket at the store.'
     B: *ho-ina,   us-le   {#kitaab} pasal-maa   {kitaab} kin-ek-o   ho.*
         AUX-NEG   3SG-ERG {#book}   store-LOC   {book}   buy-EK-3SG.MASC AUX
         'No, he bought *a book* at the store.'

(7) A: *raam-le* *{juu-maa}* *baagh* *{juu-maa}* *Dekh-yo.*
raam-ERG {zoo-LOC} tiger {zoo-LOC} see-3SG.MASC
'Raam saw a tiger at the zoo.'

B: *#ho-ina,* **kukkur(-laai)** *us-le* *Dekh-ek-o* *ho.*
AUX-NEG dog(-DAT) 3SG-ERG see-EK-3SG.MASC AUX
'No, he saw *a dog*.'

The data above suggest that Nepali does not make use of a specific syntactic operation (e.g. movement) to mark contrastive focus in *-ek* sentences. Rather, contrastive focus interpretation may often be assigned to an element in its canonical position, in opposition to the informational status of the other elements of the sentence (topicalization and discourse givenness will be discussed in the next section). In order to reinforce the existence of alternatives to the corrected focus (especially if they are salient in the context), the particle *chãi* may be used, as in (8)B.[4]

(8) *Context: Raam and Hari went to the store and bought different things.*
A: *raam-le* *pasal-maa* jacket *kin-yo.*
raam-ERG store-LOC jacket buy-3SG.MASC
'Raam bought a jacket at the store.'

B: *ho-ina,* **hari-le-chãi** jacket *kin-ek-o* *ho.*
AUX-NEG hari-ERG-ALTSHIFT jacket buy-EK-3SG.MASC AUX
'No, *Hari* (instead) bought the jacket.'

One possible exception to the corrected-corrective parallelism preference for contrastive foci is the postverbal position, as in (9)B.[5] Butt (2014) proposed that the postverbal position in Hindi "functions as a secondary focus position" (p.170), used when the verb receives primary focus. Without further discussion I will tentatively assume that something similar might be true for Nepali. With Butt (2014) I assume that movement to the right of the verb is not triggered by focus, but rather that the postverbal position is a position where NPs can independently scramble to. Crucially, the focus interpretation in that position may be possible because it does not create

---

[4] The particle *chãi* is analyzed as an alternative shifter (glossed ALTSHIFT), rather than a (contrastive) focus particle, due to the fact that it can also be used with (contrastive) topics (see section 3).
[5] The status of the postverbal position in the (verb-final) Indo-Aryan languages is controversial and is the subject of a number of recent works (see e.g. Bhatt and Dayal 2007, Manetta 2012, and Simpson and Choudhury 2015).

a clash with the informational status of the other elements of the sentence. With the focus at the end of the sentence (and possible focus on the verb also), all given/topical elements appear to the left of the verb, which is independently possible in Nepali.[6]

(9) A: *raam-le   pasal-maa   jacket kin-yo.*
        raam-ERG store-LOC   jacket buy-3SG.MASC
        'Raam bought a jacket at the store.'
    B: *ho-ina,   us-le   kin-ek-o   ho   kitaab.*
        AUX-NEG 3SG-ERG buy-EK-3SG.MASC AUX book
        'No, he bought *a book*.'

I propose that an XP can be focalized in any position it can normally scramble to, as long as it does not create a clash with the informational status of other elements. New information focus can be used to further illustrate the latter point. For instance, the direct object *kitaab* 'book' in (10), which stands for the *wh*-word in the question, may appear on either side of the locative *pasal-maa* 'at the store'. Crucially, the locative is also new information, although non-prominent. Butt & King (1996) label information that is new but not prominent as "completive information". In Hindi, they argue, completive information has to appear between the topic(s) and the focus (cf. the focus-verb adjacency requirement in Hindi). In Nepali, on the other hand, the completive information may appear before the focalized direct object, as the focus may stay in situ, or appear after the direct object, as the direct object may scramble to its left. Since both relevant elements are new information, no informational clash is obtained and therefore both orders are allowed in (10). These facts naturally follow if focalization in Nepali neither is a syntactic operation such as movement nor involves a fixed position (relative or absolute).

---

[6] I should note that the postverbal position is not the preferred position for foci for my consultant. Often times, upon accepting a postverbal focus, she would describe some additional effects or even backtrack and assign a given/topical interpretation to the postverbal element. Note that Butt (2014:162) claims that *wh*-elements in Hindi, when appearing in positions other than their default ones, cause additional pragmatic effects, a matter that still awaits a comprehensive analysis. As mentioned in footnote 5, the status of postverbal elements in Indo-Aryan languages is the topic of ongoing debate, and unfortunately at the moment I will cast more shadow, rather than light, to the matter.

(10)   A: *raam-le*   **ke**   *kin-yo?*
              raam-ERG   what   buy-3SG.MASC
              'What did Raam buy?'
       B1:   *us-le*   *pasal-maa*   **kitaab**   *kin-yo.*
              3SG-ERG   store-LOC   book   buy-3SG.MASC
              'He bought a book at the store.'
       B2:   *us-le*   **kitaab**   *pasal-maa*   *kin-yo.*
              3SG-ERG   book   store-LOC   buy-3SG.MASC
              'He bought a book at the store.'

The dialogue in (11) illustrates a case where a scrambled informational focus creates a clash with the informational status of another element, thereby leading to pragmatic infelicity. Note that *sihã* 'lion' appears in the question, so it is interpreted as a discourse-given element. Since the new-information status of *galaaichaa-maa* 'on the carpet' is anchored solely on the discourse givenness of *sihã* 'lion', the language obeys the given-before-new preference, as in (11)B1, and disfavors scrambling of the focalized element, as in (11)B2. (Right-dislocation of the discourse-given element is also an option, as in (11)B3, which is discussed in the next section.)

(11)   A: *sihã kahã*   *chh-a?*
              lion   where   AUX.PRES-3SG.MASC
              'Where is the lion?'
       B1:   *sihã galaaichaa-maa*   *chh-a.*
              lion   carpet-LOC   AUX.PRES-3SG.MASC
              'The lion is on the carpet.'
       B2:   *#galaaichaa-maa sihã*   *chh-a.*
              carpet-LOC   lion   AUX.PRES-3SG.MASC
              'The lion is on the carpet.'
       B3:   *galaaichaa-maa*   *chh-a*   *sihã.*
              carpet-LOC   AUX.PRES-3SG.MASC   lion
              'The lion is on the carpet.'

Informational focus in postverbal position does not seem to be possible, as in (12)B and (13)B2. I will leave an answer for why that is the case for future research (pending data collection). At the moment I will speculate that one or both of the two following possibilities might be on the right track. First, if it is true that in Nepali the postverbal position may be used for secondary focus but not for primary focus (as Butt 2014 proposed for Hindi), the pragmatic

clash is expected: the narrow answer to a *wh*-question has to be the primary focus of the sentence. Second, if right-dislocation is a marked movement in Nepali, informational focalization may be too trivial an operation to license it. Note that I am proposing that foci can appear in any position an XP can independently scramble to, which amounts to saying that foci and neutral elements have basically the same distribution (modulo potential informational clashes).

(12) A: *raam-le* **ke** *kin-yo?*
          raam-ERG what buy-3SG.MASC
          'What did Raam buy?'
      B: *#us-le pasal-maa kin-yo **kitaab**.*
          3SG-ERG store-LOC buy-3SG.MASC book
          'He bought a book.'

(13) A: *hari-le hijo katiotaa kitaab kin-yo?*
          hari-ERG yesterday how.many book buy-3SG.MASC
          'How many books did Hari buy yesterday?'
      B1: *Tinotaa (kin-yo).*
          three buy-3SG.MASC
          'He bought three.'
      B2: *#kin-yo Tinotaa.*
          buy-3SG.MASC three

To conclude this section, I would like to reiterate my claim that focalization in Nepali is not a syntactic operation, in that it does not trigger movement/dislocation. Also, it does not make use of fixed positions (relative or absolute), as opposed to other SOV languages where a focus-verb adjacency requirement is observed. Topicalization, on the other hand, must involve ex situ positions, as we will see next.

3.   Topicalization

Nepali has a tendency to mark discourse-given elements via dislocation. Given that foci often stay in situ and that Nepali has a given-before-new preference, the identification of the focus is facilitated if given elements and topics are dislocated past the focus. Elements that are discourse-

given (but not necessarily topicalized) tend to be dislocated, but that is not obligatory (especially if their in situ position already favors identification of the focus; see (11) above). However, if a constituent is unambiguously topicalized (indicated by the topic marker *Ta*), it must necessarily involve dislocation, either to the left periphery, as in (14)B1, or to the postverbal periphery, as in (14)B2. Crucially, *Ta* cannot be used if an element stays in situ, as in (14)B3. Note that the subject in (14)B3 is a contrastive focus that stays in its canonical position (cf. corrected-corrective parallelism), which shows that the direct object does not reach the left periphery.[7] That amounts to saying that either a topic-marked constituent in Nepali must be licensed via movement or that a topic-marked constituent must be base-generated in a (non-argumental) topic position. Either way, contrary to foci, *Ta*-marked topics cannot stay in situ.[8]

(14) A: *hari-le   tai kin-yo.*
          hari-ERG tie buy-3SG.MASC
          'Hari bought a tie.'
      B1: *ho-ina,   **tai(-Ta)**, raam-le   kin-ek-o   ho.*
          AUX-NEG   tie(-TOP) raam-ERG buy-EK-3SG.MASC AUX
          'No, the tie, *Raam* bought it.'
      B2: *ho-ina,   raam-le   kin-ek-o   ho   **tai(-Ta)**.*
          AUX-NEG   raam-ERG buy-EK-3SG.MASC AUX tie-TOP
          'No, *Raam* bought it, the tie.'
      B3: *ho-ina,   raam-le   **tai(#-Ta)** kin-ek-o   ho.*
          AUX-NEG   raam-ERG tie(#-TOP) buy-EK-3SG.MASC AUX
          'No, *Raam* bought the tie.'

It must be pointed out that the topic particle *Ta* is not restricted to bare nouns (in fact, none of the observations above changes if the relevant nouns are modified). For instance, *Ta* can also appear on oblique objects, on top of the dative marker *laai*, as in (15)B, as well as on locatives,

---

[7] Here I assume that string-vacuous scrambling of the subject is ruled out (see e.g. Hoji 1985; see also footnote 10 below); thus, the direct object does not reach the left periphery of the sentence. I additionally assume that vacuous topicalization to the middle field is banned. In Lacerda (2015:25) I argued that is also the case in Brazilian Portuguese, another language with rich use of the middle field for informational purposes.

[8] Although compatible with dedicated topic projections, as in the cartographic approach to information structure (e.g. Rizzi 1997, Belletti 2004), topic movement/dislocation does not entail the presence of a TopicP in the spine of the sentence. Conversely, the lack of focus movement does entail the lack of a FocusP in the spine of the sentence.

on top of the locative postposition *maa*, as in (16)B. (16) additionally shows that *Ta* cannot appear in questions or on focalized elements.[9]

(15) A: *gitaa-le   hari-laai   aa'p   D-ii.*
          gitaa-ERG   hari-DAT   mango   give-3SG.FEM
          'Gitaa gave Hari a mango.'
      B: *ho-ina,   **hari-laai-Ta**   us-le   siau   D-ek-i   ho.*
          AUX-NEG   hari-DAT-TOP   3SG-ERG apple   give-EK-3SG.FEM   AUX
          'No, to Hari she gave *an apple*.'

(16) A: *sofa-maa(#-Ta)   ke/ko   chh-a?*
          couch-LOC(#-TOP)   what/who AUX.PRES-3SG.MASC
          'What/Who is on the couch?'
      B: *sofa-maa(-Ta)   kukkur(#-Ta)   chh-a.*
          couch-LOC(-TOP) dog(#-TOP)   AUX.PRES-3SG.MASC
          'A dog is on the couch.'

Somewhat surprisingly, the particle *Ta* is also odd in contrastive topicalization; the particle *chãi* is preferred. Recall from the previous section (cf. (8)) that contrastive foci too can be marked with *chãi*. Given that, I analyze this particle as an alternative shifter, rather than a focus or topic marker. Observe (17). The elements *sofa* 'couch', *galaichaa* 'carpet', and *table* in (17) were not previously mentioned in discourse, but were given as being part of the set under discussion (i.e., the set of pieces of furniture) and were given as alternatives to *khaat* 'bed'. Qualifying as contrastive topics (see Büring 2003), these elements may be marked with *chãi*.

---

[9] I first hypothesized that the ban on *Ta* in questions could follow from the lack of discourse anaphoricity, if the question is the first utterance in a dialogue. However, my consultant also rejected *Ta* in follow-up questions, in contexts where the topic-marked element had been previously mentioned in discourse. On a rather speculative note, this might suggest that Force and Topic are not two independent heads in the Nepali CP system, that is, question and topicalization are encoded by the same functional head, thereby only one of either is possible at a time. I suspect there might be here a correlation between Nepali having fewer functional heads in the CP domain and its being an article-less language, in Bošković's (2008) sense. If the noun-clause parallelism holds, languages without a DP layer on the noun may also lack functional projections in the high clausal domain.

(17) Context: There are several pieces of furniture, on each of which there is an animal.

A: *ma-laai Thaaha chh-a*   *khaat-maa ke   chh-a,*
1SG-DAT know   AUX.PRES-3SG.MASC bed-LOC   what AUX.PRES-3SG.MASC
Tara aru   chij-maa   ke   chh-a?
but   other thing-LOC what AUX.PRES-3SG.MASC
'I know what is on the bed, but what is on the other things?'

B: **sofa-maa(-chãi)**   *kukkur chh-a,*
couch-LOC(-ALTSHIFT)   dog   AUX.PRES-3SG.MASC
'On the couch, there is *a dog*,'
galaichaa-maa(-chãi) *sihã*   *chh-a,*
carpet-LOC(-ALTSHIFT)   lion   AUX.PRES-3SG.MASC
'on the carpet, there is *a lion*,'
table-maa(-chãi)   *musaa chh-a.*
table-LOC(ALTSHIFT)   mouse AUX.PRES-3SG.MASC
'on the table, there is *a mouse*.'

That *chãi* requires salient alternatives in order to be licensed is corroborated by (18). Here, *chãi* is odd in the first sentence of B. Note that there is no salient set of alternatives by the moment Raam is introduced (the set of people who went to the zoo is introduced by speaker B). Therefore, Raam cannot be marked by *chãi*. However, when Hari and Sitaa are introduced, a salient set of people is already established; these can thus be marked with the particle.

(18) A: *hijo   juu-maa ke   bha-yo?*
yesterday zoo-LOC what happen-3SG.MASC
'What happened at the zoo yesterday?'

B: *hijo   juu-maa **raam-le(#-chãi)**   kukkur Dekh-yo,*
yesterday zoo-LOC raam-ERG(#-ALTSHIFT) dog   see-3SG.MASC
'Yesterday at the zoo Raam saw a dog,'
hari-le(-chãi)   *baagh   Dekh-yo,*
hari-ERG(-ALTSHIFT)   tiger   see-3SG.MASC
'Hari saw a tiger,'
ra **sitaa-le(-chãi)**   *bããDar   Dekh-ii.*
and sitaa-ERG(-ALTSHIFT) monkey see-3SG.FEM
'and Sitaa saw a monkey.'

Even if an alternative to the *chãi*-marked element is not previously mentioned, as in (19), *chãi* can be used to introduce a contrast and give rise to contrastive topicalization, as long as there is a salient set of alternatives. If there are no alternatives, as in (20), *chãi* is ruled out.

(19)   Context: Raam, Sitaa, and Hari went to the zoo.
   A: *hijo*   *juu-maa*   *ke*   *Dekh-ee?*
        yesterday zoo-LOC   what   see-3PL
        'What did they see at the zoo yesterday?'
   B: (Not knowing about Sitaa and Hari)
        hijo   *juu-maa*   **raam-le(-chãi)**   *kukkur*   *Dekh-yo.*
        yesterday   zoo-LOC   raam-ERG(-ALTSHIFT)   dog   see-3SG.MASC
        'Yesterday at the zoo Raam saw a dog.'

(20)   Context: Only Raam went to the zoo.
   A: *hijo*   *juu-maa* *raam-le*   *ke*   *Dekh-yo?*
        yesterday   zoo-LOC   raam-ERG   what see-3SG.MASC
        'What did Raam see at the zoo yesterday?'
   B: *hijo*   *juu-maa*   **raam-le-(#-chãi/✓-Ta)**   *kukkur*   *Dekh-yo.*
        yesterday   zoo-LOC   raam-ERG(#-ALTSHIFT/✓-TOP)   dog   see-3SG.MASC
        'Yesterday at the zoo Raam saw a dog.'

It is important to note that the use of topicalization and alternative shifting strategies is optional. That is, the use of *Ta* and *chãi* depends on the discursive move by the speaker (granted the dislocation requirement on *Ta*). In (21), once it is common ground that three people went to the zoo, speaker A can use *chãi* to introduce a contrastive topicalization strategy (as in (19) above). The other speaker may correct A's statement as in (21)B1 — here, since Raam is both a previously established topic and an alternative in the set of people, either *Ta* or *chãi* is possible. Interestingly, as seen in (21)B2, *Ta* and *chãi* can co-occur in a sentence: here, while *Ta* marks a topic (dislocated), *chãi* marks a contrastive focus (in situ).[10] Note that *chãi*, not being a topic marker, does not require dislocation.

---

[10] Although in (21)B2 the relative position of the subject does not distinguish between the canonical or a dislocated position, I maintain that the subject is dislocated here, given the analysis presented in the paper (although in general I remain uncommitted as to whether topic dislocation in Nepali means movement or base-generation in a non-argumental topic position). Given that Nepali is a *pro*-drop language, an analysis of (21)B2 where the topic binds a *pro* in subject position is plausible, especially if vacuous subject topicalization via movement is banned (see Bošković 2016 for an account; see also footnote 7 above).

(21)   Raam, Hari, and Sitaa went to the zoo.
          Speaker A doesn't know what Hari and Sitaa saw.

          A: *raam-le-chãi*   *baagh* *Dekh-yo.*
               raam-ERG-ALTSHIFT tiger   see-3SG.MASC
               'Raam saw a tiger.'

          B1:   *ho-ina,*   **raam-le(-Ta/-chãi)**   *kukkur* *Dekh-ek-o*   *ho.*
               AUX-NEG raam-ERG(-TOP/-ALTSHIFT) dog   see-EK-3SG.MASC AUX
               'No, Raam saw *a dog*.'

          B2:   *ho-ina,*   **raam-le-Ta**   **kukkur-chãi** *Dekh-ek-o*   *ho.*
               AUX-NEG raam-ERG-TOP dog-ALTSHIFT see-EK-3SG.MASC AUX
               'No, Raam saw *a dog*.'

Ta, being a topic marker, requires dislocation. As we saw above, the *Ta*-marked topic can be dislocated either to the left or to the right periphery of the sentence. Those, however, are not the only possibilities. Scrambling to the middle field of the sentence suffices (as long as it is not vacuous; see footnote 7). In (22)B1, *Thapa-ko-Ta* 'by Thapa' is topicalized and is linked to a position within the NP headed by *kitaab* 'book'. Since *Thapa-ko-Ta* surfaces in a position higher than the locative adverb (but still lower than the temporal adverb and the subject), a dislocation configuration is created which licenses the *Ta*-marked topic. If the topic marker is dropped, as in (22)B2, the sentence is reported to have an interpretation where *Thapa-ko* is associated with *pasal* 'store' (note that this interpretation is infelicitous in the context of (22)). This reading is unavailable in (22)B1: if *Ta* marks a dislocation, the association of *Thapa-ko* with *pasal* would require that either *Thapa-ko-Ta* be topicalized in situ or be dislocated vacuously, but both of these possibilities are ruled out.

(22)   Context: Hari bought some books by different authors at the store yesterday.

          A: *hari-le* *hijo*   *pasal-maa* *Thapa-ko* *katiotaa* *kitaab* *kin-yo?*
               hari-ERG yesterday store-LOC   Thapa-GEN how.many book   buy-3SG.MASC
               'How many books by Thapa did Hari buy at the store yesterday?'

          B1:   *us-le*   *hijo*   **Thapa-ko-Ta**   *pasal-maa*   *duiotaa* *kitaab* *kin-yo.*
               3SG-ERG yesterday Thapa-GEN-TOP store-LOC   two   book buy-3SG.MASC
               'He bought two books by Thapa at the store yesterday.'

          B2:   #*us-le*   *hijo*   **Thapa-ko** *pasal-maa*   *duiotaa* *kitaab* *kin-yo.*
               3SG-ERG yesterday Thapa-GEN store-LOC   two   book buy-3SG.MASC
               'He bought two books at Thapa's store yesterday.'

Topicalization in Nepali is thus shown to necessarily involve overt, non-vacuous dislocation. In other words, the licensing of *Ta*-marked topics must occur via a specific syntactic configuration (contrary to focalization, which is not constrained as such). *Ta*-marked elements in Nepali are true, unambiguous topics, while *chãi*-marked elements are potentially ambiguous in being either given or new information; *chãi* therefore does not require dislocation (as both discourse-given and focalized elements may stay in situ). In the final remarks, I will briefly discuss some consequences of these observations for a theory of Information Structure in Nepali (and in general).

4.   Final remarks

In analyzing word order and information structure in Turkish (which like Nepali is SOV), Şener (2010) proposed that all movements in that language are triggered by the syntactic counterpart of a discourse/pragmatic function (more precisely, an operator feature on the moving element). In that state of affairs, there is no room for pragmatically vacuous scrambling. The caveat is the following: all topics and discourse anaphoric elements must move, but foci must stay in situ (which is what derives the focus-verb adjacency requirement, as all other elements move to higher positions). The reason, he argues, is that in Turkish topicalization and discourse anaphoricity are encoded in the syntax by an operator-variable relation, which is not the case with focalization.

Given the data presented in this paper, we can seamlessly transpose Şener's (2010) analysis to focalization and topicalization in Nepali: while topics are read off syntax via an operator-variable relation, focalization is not encoded as such. Recall that foci in Nepali may move, but I argue that this movement is not triggered by a focus feature or the necessity to create a specific structural configuration in the syntax. Discourse-given elements also do not require dislocation in Nepali (unlike Turkish), which suggests that discourse anaphoricity is not marked in the syntax of the language (although it is subject to the usual PF effects, such as deaccenting).

The observation that discourse-given and neutral elements alike may either appear in their canonical positions or scramble (and perhaps also the observation that the positions of foci

and completive information may be freely interchangeable, cf. (10)) leads us to the conclusion that Nepali leaves room for pragmatically vacuous scrambling. That creates the possibility that constituents be scrambled to a certain position, and be additionally — and independently — focalized in that position. As we saw, focus movement is not only generally not required, but it is also dispreferred when an element is focalized with respect to a constituent that was in situ in the (antecedent) corrected sentence or question.[11]

The discussion above boils down to the question of what aspects of Information Structure, if any, are encoded in the syntax of natural languages. In Nepali, at least, it seems that topicalization must be achieved in syntax, unlike focalization. That focalization is not part of the narrow syntax was explicitly proposed by Horvath (2010), who argues that so-called "focus movement" in Hungarian is in fact driven by a semantic exhaustive identification operator. She however takes a strong view and proposes that no discourse operations at all exist in the syntax.

While focalization in Nepali seems to go along with Horvath's general claim, the same cannot be said about topicalization. If anything, the novel data presented in this paper bring additional evidence for a body of work that indicate that grammars tend to (or solely) mark in the narrow syntax information that is old and discourse-given/topical, rather than information that is new/focal.[12] For instance, Kučerová (2012) claimed that in Czech, Russian, and Serbo-Croatian, discourse-given elements must always be grammatically marked, by moving to (or being base-generated in) positions that linearly precede new elements (which she captures by proposing a *G-operator* which marks elements in its scope as given). Nepali, as we saw, is not as strict with respect to discourse givenness, as given and new elements may intertwine (for instance in the correction of the sentential subject). Although it does not have a G-operator in its syntax, Nepali is a givenness-oriented language, as topics but not foci require that an operator-variable configuration be established in the syntax.

---

[11] We should bear in mind the complications of the postverbal position (see footnotes 5 and 6). If we adopt an analysis similar to that of Bhatt & Dayal (2007) for Hindi, where rightward movement is restricted to verbal projections, it might be the case that the postverbal focalized direct object in e.g. (9)B above is actually in its base position, inside a VP that is further right-dislocated (whose head previously moved to a higher projection). I will however leave the matter open here.

[12] This view can be summarized by Du Bois's (1987) claim that "grammars code best what speakers do most."

Finally, the fact that the syntax of topics and the syntax of foci are clearly asymmetric in many languages (like the ones mentioned in this paper[13]) challenges the notion that topicalization and focalization are two sides of the same coin when it comes to the interface between syntax and Information Structure. That comparison thus presents itself as a very fruitful domain to tease apart accounts of information structure, such as the cartographic approach (e.g. Rizzi 1997, 2013; Belletti 2004) and the mapping approach (e.g. Neeleman & van de Koot 2008). The data presented here show that in Nepali topicalization may or may not be cartographic (as dislocation may or may not involve movement to or base-generation in dedicated topic projections), while focalization clearly is not (as no focus-driven dislocation is involved).[14] That the syntax of topics and the syntax of foci are so frequently distinct within languages may be a suggestion that the interface mechanism (that is, how discourse functions are read off syntax by the pragmatic component) may also be distinct for topics and foci, in the end calling for a hybrid approach to information structure.

5.   References

Acharya, Jayaraj. 1990. A descriptive grammar of Nepali and an analyzed corpus. PhD Dissertation, Georgetown University.

Belletti, Adriana. 2004. Aspects of the low IP area. In *The Structure of CP and IP: The Cartography of Syntactic Structures, Volume 2*, ed. by Luigi Rizzi, 16–51. New York: Oxford University Press.

Bhatt, Rajesh, and Veneeta Dayal. 2007. Rightward scrambling as rightward remnant movement. *Linguistic Inquiry* 38: 287–301.

Bošković, Željko. 2008. What will you have, DP or NP? In *Proceedings of the North East Linguistic Society 37*, ed. by Emily Elfner and Martin Walkow, 101–114. Amherst GLSA, University of Massachusetts.

Bošković, Željko. 2016. On the timing of labeling: Deducing Comp-trace effects, the Subject Condition, the Adjunct Condition, and tucking in from labeling. *The Linguistic Review* 33: 17–66.

---

[13] In work in preparation, I argue that Brazilian Portuguese must be added to that list. While topicalization involves dislocation and dedicated projections (TopicP) in the spine of the sentence, focalization does not.

[14] If the speculations in footnote 9 are plausible, it might also be plausible that article-less (i.e. DP-less) languages lack TopicP and FocusP altogether.

Butt, Miriam. 2014. Questions and information structure in Urdu/Hindi. In *Proceedings of the LFG14 Conference*, ed. by Miriam Butt and Tracy Holloway King, 158–178. Stanford: CSLI Publications.

Butt, Miriam, and Tracy Holloway King. 1996. Structural topic and focus without movement. In *Proceedings of the First LFG Conference*, ed. by Miriam Butt and Tracy Holloway King. Stanford: CSLI Publications.

Du Bois, John W. 1987. The discourse basis of ergativity. *Language* 63: 805–855.

Hoji, Hajime. 1985. Logical Form constraints and configurational structures in Japanese. PhD Dissertation, University of Washington.

Horvath, Julia. 2010. "Discourse features", syntactic displacement and the status of contrast. *Lingua* 120: 1346–1369.

Kidwai, Ayesha. 2000. *XP-Adjunction in Universal Grammar: Scrambling and Binding in Hindi-Urdu*. Oxford: Oxford University Press.

Kučerová, Ivona. 2012. Grammatical marking of givenness. *Natural Language Semantics* 20: 1–30.

Lacerda, Renato. 2015. Asymmetries between high and low topics in Brazilian Portuguese. General examination paper, University of Connecticut.

Manetta, Emily. 2012. Reconsidering rightward scrambling: Postverbal constituents in Hindi-Urdu. *Linguistic Inquiry* 43: 43–74.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In *Elements of Grammar*, ed. by Liliane Haegeman, 260–318. Dordrecht: Kluwer.

Rizzi, Luigi. 2013. Notes on cartography and further explanation. *Probus* 25: 197–226.

Şener, Serkan. 2010. (Non-)peripheral matters in Turkish syntax. PhD Dissertation, University of Connecticut.

Simpson, Andrew, and Arunima Choudhury. 2015. The nonuniform syntax of postverbal elements in SOV languages: Hindi, Bangla, and the rightward scrambling debate. *Linguistic Inquiry* 46: 533–551.

Renato Lacerda

UConn Department of Linguistics

renato.lacerda@uconn.edu