JSAL

Volume 4, Issue 1 December 2011

Journal of South Asian Linguistics

CSLI PUBLICATIONS

Contents

Editorial Statement

We are happy to present to you, our readers, the fourth volume of the Journal for South Asian Linguistics. We owe you an apology as this issue has been much delayed. We will spare you a full listing of our failings but you should know that we are very much aware of them. We promise to do better and to help us make the Journal be a more responsive operation, we have asked Tracy King to join us as co-editor of *JSAL*. Tracy has not only worked extensively on Slavic syntax, clitics, and various theoretical and computational issues within LFG, she also has published several papers on South Asian languages. She currently works in industry, where her renowned organizational prowess is much appreciated and which we are excited to benefit from along with her significant theoretical and computational expertise. Tracy has accepted our invitation and so we ask you to join us in welcoming her. She will begin tenure as a co-editor in January 2012.

In terms of this current issue, we are again happy to present what we see as a perfect illustration of the mix of theoretical, psycholinguistic, and descriptive concerns that we want *JSAL* to cover. The first paper analyzes the syntax of numerical classifiers in Bangla. The eastern Indo-Aryan languages, to which Bangla belongs to, are special within the Indo-Aryan languages in requiring classifiers with numerals. In this paper, Dustin Alfonso Chacón gives a new treatment of classifier phenomena in Bangla which argues that there is DP-internal movement of the NP to [Spec, DP] to mark definiteness and that the Numeral does not form a constituent with the classifier. He also expands the empirical database by bringing to bear upon the debate data from the 'quantificationally approximate' construction where the classifier precedes the numeral.

The second paper, by Sudha Arunachalam and Anubha Kothari, is an excellent illustration of the application of psycholinguistic methodologies to semantic questions. There has been a lively debate in the semantics of aspect about the links between telicity and perfectivity. For English, it is often claimed that to successfully 'eat an apple', one has to finish said apple. This claim has been challenged for a number of languages, including Hindi/Urdu. The authors use experimental methodologies to explore this question and conclude that the issue of completion is not completely represented in the semantics; instead a variety of factors including contextual ones are at play. They also use the results of their investigation into Hindi/Urdu to suggest a re-examination of the data assumed for English.

The final paper, by Gwendolyn Hyslop, provides a detailed discussion of mirativity in Kurtöp, a Tibeto-Burman language of Bhutan. Mirativity is the explicit grammatical marking of 'surprise'; it is closely related to the phenomenon of evidentiality, which has been the subject of much recent discussion in the semantics and acquisition literature. Hyslop shows how mirativity interacts with aspect and is manifested throughout the verbal paradigm of Kurtöp.

Kurtöp is an under-documented language and we are happy to contribute to its documentation. In this context, we are also happy to note that this issue advances by two the number of languages that we have had papers on. One of the goals of *JSAL* is to provide a venue for work on the large number of languages spoken in South Asia, and in doing so, to go beyond the natural focus on the dominant languages and language-families of this area. The language count from preceding issues goes as follows: Hindi/Urdu (3), Telugu (2), Brahui, Malto, Persian, Sanskrit, and Tamil (1 each). This volume has a paper on Hindi/Urdu, one on Bangla, and one on Kurtöp. In the fullness of time, we would like to have a much wider coverage; the current issue is an excellent step in that direction.

Finally, we would like to thank Annette Hautli, who helped significantly with the necessary LaTeX type-setting and to Sebastian Sulger for maintaining the journal website. Their labor was partly funded by the University of Konstanz as part of an on-going CSLI-Konstanz cooperation and we would like to thank the University of Konstanz for supporting our publication effort. As ever we also thank Dikran Karagueuzian of CSLI Publications for his role as a continually outstanding and supportive publisher.

Rajesh Bhatt, University of Massachusetts, Amherst Miriam Butt, University of Konstanz

Head Movement in the Bangla DP

Dustin Alfonso Chacón, *University of Maryland, College Park*

Received March 2010; Revised March 2011

Abstract

Bengali/Bangla is unusual among South Asian languages in that it uses numerical classifiers. In this paper, I propose a new analysis of the DP structure in Bangla motivated by data previously unaccounted for and typological concerns. Specifically, I propose that Bangla has DP-internal NP movement to Spec,DP to mark definiteness, that the numeral and classifier form separate heads in the syntax, and that there is noun to classifier movement when there is no overt classifier. I propose a feature for each of these phenomena, and attempt to explain the ungrammatical examples using principled reasons derived from this structure. Also, I give an analysis for the quantificationally approximate construction, in which the classifier appears on the left of the numeral. I claim that the model presented in this paper can account for these constructions, and that the differences found between "classifier-compatible" nouns and "classifier-less" nouns with regard to the quantificationally approximate structures follows naturally from my analysis.

1 Introduction

One goal of linguistic theory is to identify properties present in all languages and to discover how languages may differ and why. Identifying universal properties helps us understand better the underlying architecture of grammar and its constraints. Similarly, identifying how languages vary illustrates the flexibility of the language faculty. The internal structuring of noun phrases is one area in which languages vary widely. However, since many well studied languages have a fairly rigid nominal structure, thus far it has been difficult to make generalizations about the universal and language-specific properties of noun phrases.

Bangla is a language with a fairly flexible nominal structure. It is interesting from a typological standpoint, since it is a classifier language in the Indo-Aryan family, comprised predominantly of languages with no numerical classifiers. Previous analyses of the Bangla DP (representatively, Bhattacharya 1999b) have accounted for DP-internal alternations of the position of the NP and the distinction between nouns that combine with classifiers versus nouns that do not. The model pursued by Bhattacharya (1999a,b, 2001) requires that the numeral and classifier form a fused head to derive the other properties of Bangla DPs. However, this approach does not account for a construction that is used to express quantificational approximateness, a "vague" reading on the quantity of a nominal expression. In these constructions, the numeral and classifier invert positions. This phenomenon is difficult to account for in a theory in which the numeral and classifier morphemes form a composite head. Additionally, an approach that relies on a fused head consisting of the numeral and classifier does not generalize to other classifier languages (Simpson 2005). To make stronger generalizations about the structure of DPs in classifier languages and in languages universally, it is important to try

JSAL Volume 4, Issue 1, December 2011.

to reconcile data between disparate languages that show similar phenomena, such as Bangla and the classifier languages of Southeast Asia, in order to construct a stronger theory of universal grammar. Additionally, a theory of UG in which two classifier languages differ in minute and opaque ways presents a potentially unnecesssarily complicated learning problem for the child.

This paper introduces a model that accounts for the quantificational approximate constructions and expands structural similarities between Bangla and other classifier languages. I will posit that there is DP-internal NP movement to Spec,DP to mark definiteness. In order to explain the differences between nouns that generally are found with classifiers and those that are not, I will posit that there is N^0-to-Cl^0 movement for the nouns that do not appear with classifiers. I will formalize this by claiming that some nominal roots do not merge with a nominalizer n, and thus move to adjoin to a null Cl^0 for interface reasons. I will explain the ungrammaticality of NP movement marking definiteness in cases where no overt classifier appears by suggesting that n is required to license NP movement to Spec,DP. Establishing this, I will propose a feature [VAGUE] on D^0 that enters into an AGREE relation with Cl^0, triggering left-adjoining head movement to generate the quantificationally approximate DP structure. I will also show that the resulting word orders provide independent evidence for an analysis in which Num^0 and Cl^0 are separate heads, as has been claimed to be the case in other classifier languages.

2 Structures

In this section, I briefly describe the forms of the morphemes that constitute the Bangla classifier system. Then, I introduce the syntactic phenomena under discussion that hinge on the distribution of the classifier. Afterwards, I describe the phenomena for the two classes of nouns that exist in Bangla: those that are commonly seen in a DP that contains a classifier ("classifier-compatible") and those that are frequently seen in a DP that does not contain a classifier ("classifier-less").

2.1 Classifiers in Bangla

Bangla utilizes a relatively small set of classifiers.[1] The most common classifier is *(go)ʈa*, which surfaces as the clitic *=ʈa* in certain configurations. The cliticized form additionally has the diminutive *=ʈi* and allomorphs *=ʈo* and *=ʈe* in some dialects. The allomorph *=ʈo* surfaces when *(go)ʈa* is cliticized onto *du* 'two', and *=ʈe* surfaces after *tin* 'three' and *car* 'four'. Additionally, *(go)ʈa* has a plural variant *gulo* and a plural diminutive *guli*. The forms of *(go)ʈa* are shown below.[2]

(1) a. *kɔek=ʈa bɔsta*
few=CL sack
'a few sacks'

b. *goʈa=kɔek bɔsta*
CL=few sack
'a few sacks'[3]

c. *kɔtok=gulo bɔsta*
few=CL sack
'some sacks'

d. *du=ʈo kɔlom*
two=CL pen
'two pens'

[1] Much of this description is dependent on Dasgupta (1983).

[2] Contrasting with closely-related languages, it seems that the forms of *(go)ʈa* need to be cliticized onto something — either a Num or NP on its left, or material on its right for the quantificationally approximate constructions. Thus, forms like **goʈa kham* "CL envelope" are not attested. Sahoo (1999) provides data that, to me, seems to suggest that this constraint does not apply in Oriya, however.

[3] In Dasgupta (1983), this is translated as 'a couple of sacks'.

e. *tin=ʈe kɔlom*
three=Cl pen
'three pens'

Additionally, Bangla uses *khan(a)*, for flat square objects, and *jɔn(a)*, for human-denoting nouns. The classifier *khana* surfaces as *khana* after a numeral, and as *khan* before a numeral. Similarly, *jɔna* appears as *jɔn* after a numeral, and *jɔna* before a numeral. These are shown in (2).

(2) a. *pãc=jɔn kormi*
five=Cl employee
'five employees'

b. *jɔna=pãc=ek kormi*
Cl=five=Ek employee
'five employees or so'

c. *kɔek=khana biskuʈ*
several=Cl cookie
'several cookies'

d. *khan=kɔek biskuʈ*
Cl=several cookie
'several cookies'[4]

2.2 Classifier-Compatible Nouns

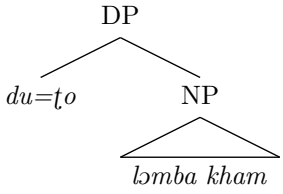
In the next two sections, I sketch the syntactic patterns under discussion. First, I introduce the DP structures that contain an overt classifier. These are shown in (3), with partial structures in (4).[5]

(3) a. *du=ʈo (lɔmba) kham*
two=Cl long envelope
'two (long) envelopes'

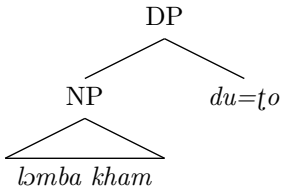
b. *(lɔmba) kham du=ʈo*
long envelope two=Cl
'the two (long) envelopes'

c. *goʈa du=ek (lɔmba) kham*
Cl two=Ek (long) envelope
'two (long) envelopes or so'

(4) a. *du=ʈo lɔmba kham*

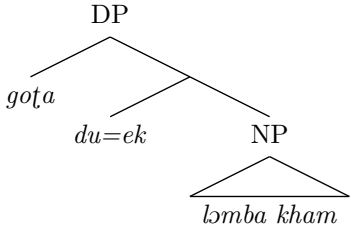


b. *lɔmba kham du=ʈo*



[4]Examples adapted from Dasgupta (1997).
[5]In these structures, the numeral and classifier are represented on one node. I reject this later.

c. *goʈa du=ek lɔmba kham*



In (3a), the DP not marked for definiteness or quantificational approximateness has the word order NUM-CL-(ADJ)-N, numeral-classifier-adjective-noun. For the moment, we assume that the adjective and noun form some intermediate structure, the NP. For DPs marked for definiteness and not marked for quantificational approximateness, the word order is NP-NUM-CL, shown in (3b). Once more, the potential to place an adjective *lɔmba* 'long' left of the noun suggests that an entire NP appears to the left of the numeral and classifier. Finally, for DPs marked for quantificationally approximateness and unmarked for definiteness, the order of the elements is CL-NUM-NP with a morpheme *ek* cliticized to the numeral, shown in (3c). Definite, quantificationally approximate structures are discussed in section 7.[6]

2.3 Classifier-less Nouns

Now, I direct my attention to nouns that do not normally occur in nominals that contain a classifier. First, it is important to establish what comprises this category. In Dasgupta (1983), the author mentions that "measure words" do not *normally* take classifiers, except to achieve certain referential or contextual effects.

(5) a. i. *John amar sathe du din thaklo*
John my with two day stayed
'John stayed two days with me.'

ii. *tin mas*
three month
'three months'

b. i. *du=ʈo bɔchor kharap gælo — 1966 ar 1976.*
two=CL year bad went — 1966 and 1976
'Two bad years occurred — 1966 and 1976'[7]

ii. *du=ʈo mas-er nam-er ʃeṣ-e "ari" – januari ar phebruari.*
two=CL month-GEN name-GEN end-LOC "ari" – January and February
'Two months have names ending in *ari* – January and February'

The forms with the unmarked interpretations are are shown in (5a). These refer to spans of time and not entities. Examples of the marked referential usages occur in (5b). Here, the noun phrases do refer to entities. The DPs with the marked interpretation for these sets of nouns will behave as other classifier-compatible nouns.

Additionally, Bhattacharya (1999a) gives the forms of nominals without a span-of-time denoting noun in (6). None of these surface with an overt classifier.

(6) a. *du deʃ-er moitri*
two country-GEN friendship
'friendship between two countries'

[6] Examples (3) and (7) adapted from Bhattacharya (1999a) and Dasgupta (1983). Parentheses mark optionality in the gloss and in the Bangla forms.

[7] This is slightly reworded from the original paper.

b. *tin caka-r gaɽi*
three wheel-Gen vehicle
'three-wheeled vehicle'

c. *car paʃ*
four side
'four sides'

d. *tin dik*
three direction
'three directions'

e. *ora car bon tin bhai*
they four sister three brother
'They are four sisters and three brothers.'

Dasgupta's examples show that for time span denoting nouns there is a referential distinction associated with the classifier — reference to entities is associated with a classifier, and reference to spans of time is associated with the absence of a classifier. However, the examples in (6) show that for other nouns, classifiers may be absent as well. In other words, these are not all "measure words", per the description in Dasgupta (1983). It is unclear whether there is a semantic property all the nouns in (6) share that could motivate this behavior or whether it is an arbitrary lexically specified feature. The absence or presence of a classifier in DPs with a time span denoting noun are accounted for in the model put forth in this paper. However, the cases in (6) still remain a problem. Interestingly, Simpson (2005) finds similar classes of bare nouns in other classifier languages. A unified analysis from a cross-linguistic or diachronic perspective could potentially explain this phenomena better than any speculation I could offer in this paper. I leave this for further research.

The word orders given in (3) are not the same for classifier-less nouns. The paradigm for classifier-less nouns is given in (7).

(7) a. *tin sɔptaho*
three week
'three weeks'[8]

b. * *sɔptaho tin*
week three
intended: 'the three weeks'

c. *sɔptaho tin(=ek)*
week three(=Ek)
'three weeks or so'

For the indefinite quantificationally unmarked DPs shown in (7a), the obvious difference from the classifier-compatible analogs is the absence of a classifier. Additionally, for definite DPs, there is no analog to (3b) for classifier-less nouns — the example in (7b) is ungrammatical. Finally, for quantificationally approximate indefinite interpretations, the classifier-less DPs surface as N-Num, with the optional morpheme *ek* after Num, shown in (7c). This is quite different from the order Cl-Num-N seen with the classifier-compatible analogs in (3c).

Interestingly, the semantic asymmetry triggered by the presence or absence of an overt classifier and the word order in the quantificationally approximate structures, exemplified in (5) and (7c) respectively, do not occur as such for nouns like *deʃ* 'country' or *caka* 'wheel'. That is, for nouns that are not "measure words" or that denote spans of time, the paradigm diverges. This is demonstrated in (8) for *caka* 'wheel'. The question of whether this means we should see the data as crossing across

[8]This construction may be allowed to have a definite reading in addition to the indefinite reading represented by the gloss. This is discussed briefly later.

three paradigms or two are not be considered here. The model pursued in this paper is only able to account for the nouns like *din* or *mas*, that do have a CL-NUM-EK-type of quantificationally approximate structures.[9]

(8) a. *tin caka*
three wheel
'three wheels'

b. (?) *tin-ʈe caka*
three=CL wheel
'three wheels'

c. * *caka tin=ek*
wheel three=EK
intended: 'three wheels or so'

d. *goʈa tin=ek caka*
CL three=EK wheel
'three wheels or so'

These patterns are summarized in the table in (9), with +DEF representing a specification for definiteness, and +QA representing a specification for quantificational approximateness. In this table, +CL represents nouns that normally appear with a classifier, and −CL represents nouns that generally do not.[10]

(9)

	−DEF, −QA	+DEF, −QA	−DEF, +QA
+CL	NUM-CL-NP	NP-NUM-CL	CL-NUM-EK-N
−CL	NUM-NP	*NP-NUM	N-NUM-EK

The data in (3) and (7) drives the analysis presented in the next sections.

3 Previous Analyses

The most articulated model of Bangla DPs within the generative tradition has been advanced by Tanmoy Bhattacharya, primarily in Bhattacharya (1999a, 1999b, 2001), though referenced in other work.[11]

3.1 Bhattacharya's (1999) Analysis

An adapted tree representing the analysis of Bangla DPs in Bhattacharya (1999b) is given in (10).[12]

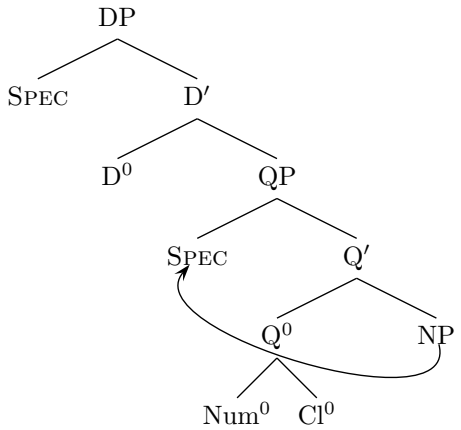
[9] Rajesh Bhatt (p.c.) suggested that this class of nouns may only seem to appear without a classifier on the surface, and that there may be a surface-deletion rule in the context of these kinds of nouns. This is an interesting proposal that I leave to future research. This kind of analysis is wholly consistent with the analysis offered in this paper.

[10] These are not syntactic features that play a role in the model pursued here. For the moment, they are just convenient abbreviations.

[11] For alternative analyses and dissenting views within the generative tradition, see Bhattacharja (2008), Dirdal (2004) and Ghosh (2006). In other theoretical models, see Bhattacharja (2007).

[12] It is also worth noting that Bhattacharya posits further structure within the NP, including a nP between the QP and NP shells. He uses these to account for further phenomena not under analysis here. I do not comment on this aspect of his theory; I only discuss the QP and the structures that it is directly involved with in this section. I continue to refer to all items below the QP as NP, and later I offer an analysis in which the NP is constructed out of a nominalizer n and $\sqrt{}$, for different empirical reasons than Bhattacharya's (1999) nP.

(10)



In this model, there is only one additional head between the higher DP layer and the lower NP layer, Q^0 for quantifier. The Q^0 is a complex head consisting of Num^0, the syntactic position for numerals and quantifiers, and Cl^0, the position for classifiers.

To account for the alternation in (3a) and (3b), repeated in (11), Bhattacharya (1999b) proposes an optionally strong feature [Specific] on N^0. This feature causes the NP to move to Spec,QP.

(11) a. *du=ʈo (lɔmba) kham*
two=Cl long envelope
'(the) two (long) envelopes'

b. [*(lɔmba) kham*]$_i$ *du=ʈo* t_i
long envelope two=Cl
'the two (long) envelopes'

To explain the ungrammaticality of (7b), repeated in (12), Bhattacharya proposes that the classifier is required in Q^0 to license the [Specific] feature. In other words, NP movement for specificity optionally occurs, but only when a classifier has merged in Q^0. Since there is no classifier in Q^0 in example (12), NP movement is illicit.

(12) * *sɔptaho tin*
week three
intended: 'the three weeks'

The presence of the classifier does seem to be the difference between the forms in (3b) and (7b), now in (11b) and (12). However, although the classifier is the catalyst for NP movement, the NP's target is to the left of the numeral, not to the left of the classifier. A fused Q^0 captures the correct word order. Having a separate Num^0 and Cl^0 would require having [Specific] checked by the Num^0 in order to explain the NP's position, contrary to the observation that the presence of the classifier is required for NP movement. Otherwise, an analysis with separate Num^0 and Cl^0 heads would require an additional movement to have the NP checking [Specific] surface to the left of the numeral, assuming Bhattacharya's mechanisms. The fused head ensures that the classifier element in Q^0 motivates NP movement, without positing a separate step to get the correct word order. This is a strong advantage of Bhattacharya's model. Later, we will try to capture this intuition by tying the licensing of the classifier and of the movement for definiteness to a distinct, third property of the DP, namely, whether a nominalizer *n* merges.

3.2 The Classifier as a Marker of Specificity

Bhattacharya (1999b) hypothesizes that the feature motivating NP movement to Spec,QP is an optionally strong [Specific] feature. This is consistent with Dasgupta (1983) and Junghare (1983), who refer to the classifier as marking specificity.

Why analyze the presence of the classifier or NP movement as marking specificity? There does not seem to be a clear argument one way or another for a definiteness or a specificity-marking analysis, nor is it clear what syntactic reflexes are contributing to this reading. As the idiomatic translations for the examples with a moved NP in (3b) and (7b) would have the reader believe, NP movement seems to mark definiteness pretheoretically. If we were to analyze the classifier as a marker of "just" specificity, we can explain why we can get a classifier in indefinite configurations, shown earlier in (3), repeated in (13), and (14). Here, the classifier is compatible with an indefinite interpretation if it appears before the noun in (13). Additionally, the classifier can cliticize to *ek* with both the indefinite 'one boy' and 'a boy' readings in (14), meaning we can not suggest that the appearance of the classifier itself is what triggers a definiteness reading.

(13) a. *du=ʈo (lɔmba) kham*
two=CL long envelope
'(the) two (long) envelopes'

b. *(lɔmba) kham du=ʈo*
long envelope two=CL
'the two (long) envelopes'

c. *goʈa du(=ek) kham*
CL two(=EK) envelope
'two envelopes or so'

(14) a. *chele*
boy
'(a) boy'

b. *ek(=ʈi) chele*
one=(CL) boy
'one boy' or 'a boy'

An analysis in which the classifier marks definiteness alone, separate from NP movement, would be too strong in light of these data.

One potential challenge to the definiteness analysis of NP movement is the appearance of the classifier in forms expressing "anti-definiteness."[13] An example of the anti-definiteness phenomenon is given in (15), taken from Dasgupta (1997). In this example, there are two DPs: *lebu=ʈa* 'lemon=CL' and *lɔŋka=ʈa* 'chili=CL'. These two DPs do not refer to either lemon or chili. Instead, they refer to a superset containing both lemon and chili, glossed as 'vegetables.'[14]

(15) *poɽʃir kache lebu=ʈa lɔŋka=ʈa cee newa*
neighbor from lemon=CL chili=CL wanting take
'Borrowing (some vegetables) from a neighbor.'

In this section, I presented the current perspective on the data. In the next sections, I critically evaluate these standing claims. Using the original arguments as a springboard, I explore a new analysis of these data. I argue that this new analysis accounts for the quantificationally approximate readings, left unaccounted for in previous models. The implications of the definiteness/specificity distinction regarding the structural configurations of Bangla DPs is explored in the next section. I ultimately make explicit an analysis in which NP movement does indeed mark definiteness.

4 NP-Movement as a Marker of Definiteness

In this section, I specifically assess the claim that specificity is the motivation for NP movement in Bangla. I first show that there are separate motivations for the appearance of the classifier and for NP

[13] Originally in Dasgupta (1997), *antidéfinitude*.
[14] The fact that lemon is a fruit notwithstanding.

movement, and that the NP movement facts are incompatible with an analysis relying uniquely on specificity. I then argue that anti-definiteness is not evidence against an analysis in which definiteness motivates NP movement. Finally, I propose that definiteness, formalized as [Def], motivates NP movement to Spec,DP.

4.1 Definiteness and NP Movement

In section 3.2, I presented arguments that the classifier marks specificity. However, is specificity an appropriate motivation for NP movement? One thing is certain — we cannot have both the NP-movement and the appearance of the classifier triggered by specificity strictly. If specificity were the feature triggering NP movement as well as the feature triggering appearance of a classifier, then there should be no meaning difference between a moved NP and an *in-situ* NP appearing with a classifier. For any DP with a specific interpretation, a classifier would appear to signal specificity, and the NP would optionally move to the left of the classifier to redundantly check [Specific], or optionally remain *in-situ*. If the DP were non-specific, there would be neither a classifier nor motivation for NP movement. In other words, if the appearance of the classifier and NP movement share a motivation, moving the NP would only redundantly mark specificity, and there should be no change in interpretation. This is contrary to the observations presented thus far. Therefore, the two phenomena require distinct triggers. I propose definiteness, formalized as a [Def] feature on D^0, motivates NP movement. If this is correct, we can maintain the observation that NP movement signals the observed semantic effect.

Further evidence against an analysis in which specificity motivates NP movement is shown in (16).[15] If specificity were indeed the relevant property motivating NP movement, and definiteness an irrelevant feature, we should find examples of the string NP-Num-Cl or NP-Cl with a specific indefinite reading. As shown below, however, such a configuration is actually infelicitous with an intended specific indefinite reading. That is, a moved NP must receive a definite interpretation, and specificity does not suffice as a trigger for NP movement.

(16) a. *jɔkhon amar bacca-r bɔes tin chilo, tɔkhon amra ek bondhu-r / (#*
when my child-Gen age three was then we one friend-Gen
bondhu=ʈa)-r baɽi-te chilam, ar se ama=ke boleche je ...
friend=Cl-Gen house-Loc were and he 1Sg=Acc told that ...
'When my child was three years old, we were at a friend's house, and he told me that ...'

b. *ami æk=ʈa kukur kinte cai*
I one=Cl dog to.buy want
'I want to buy a dog (any dog will do).' or 'I want to buy a dog (and I have one in mind).'

c. *# ami kukur=ʈa kinte cai*
I dog=Cl to.buy want
intended: 'I want to buy a dog (any dog will do).' or 'I want to buy a dog (and I have one in mind).'

d. *ami kukur=ʈa kinte cai*
I dog=Cl to.buy want
'I want to buy the dog.'

I do not at this time make any explicit claim about any additional referential effects contributed by the classifier.[16] However, as Rajesh Bhatt (p.c.) suggested, if there is any additional referential

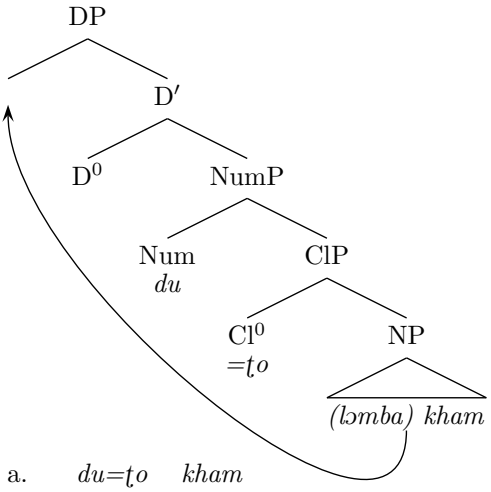
[15] The order of the genitive -r as following the clitic ʈa in (16a) is somewhat surprising and points to the need for further investigation into the interplay between case marking and classifiers in Bangla.

[16] However, for a particularly promising suggestion as to why classifiers appear, see Dasgupta (1985), in which Dasgupta offers an analysis in which classifiers are used to give a value to an otherwise unspecified aggregate feature on the noun.

effects that are contributed by the classifier alone and classifiers are required for counting, we would expect all quantified NPs to have the same referential status, which seems to be wrong.

I assume that if definiteness is indeed the motivation for NP movement, then [DEF] must be checked at Spec,DP. This move seems plausible, since D^0 is the generally accepted position for definiteness markers. If the NP moves to Spec,DP, there is no issue with resulting word orders since the NP still appears to the left of the numeral. This is shown in (17), with the assumption that Num^0 and Cl^0 form separate heads as argued later. However, this leaves us with a problem. As mentioned earlier, there is a link between the presence of a classifier and overt NP movement, exemplified in (18). How do we account for this apparent relation, if we divorce the two phenomena from each other? I address this in section 5.

(17) *lɔmba kham du=ʈo*



(18) a. *du=ʈo kham*
two=CL envelope
'two envelopes'

b. *kham du=ʈo*
envelope two=CL
'the two envelopes'

c. *dui sɔptaho*
two week
'two weeks'

d. * *sɔptaho dui*
week two
intended: 'the two weeks'

One further note. Definiteness is often associated with bare classifiers in classifier languages, per Cheng and Sybesma (2005), Simpson (2005), and Aikhenvald (2000). An analysis in which definiteness sensitive to the presence of a classifier motivates NP movement dovetails nicely with this observation. In other words, the model presented in this paper may generalize to other classifier languages. This is compatible with our secondary goal of proposing a model that would not require the learner to consider a larger class of grammars than is necessary.

4.2 Anti-Definiteness

Before we accept that the NP movement is a reflex of definiteness, we must address whether anti-definiteness à la Dasgupta (1997) is a fatal phenomenon. I do not believe that it is. I entertain two possible explanations that tease out the definiteness from these anti-definiteness constructions without losing Dasgupta's intuitions. The relevant constructions are shown in (15), repeated in (19).

(19) *porʃur kace lebu=ʈa lɔŋka=ʈa cee newa*
neighbor from lemon=Cl chili=Cl wanting take
'Borrowing (some vegetables) from a neighbor.'

Gundel et al. (2001) note cases of definite DPs referring to classes, shown in (20). Here, no particular entity is being referred to. Instead, the speaker refers to entire classes — the class of religious tourists, the class of night life tourists, and so on. One could suggest that these apparently definite examples in (20) are parallel (pragmatically, if not also syntactically) to the forms in (19).

(20) Montreal has something to offer *the religious tourist* and *the night life tourist*, *the art connoisseur*, *the sports fan*, *the intrepid walker*, and *the avid consumer.*

If this is the case, then the syntactic configuration of the DP in (19) does not necessarily preclude an analysis in which NP movement serves to mark definiteness. In fact, since English marks these class references with the definite determiner *the*, the apparent structural definiteness in (19) and (20) may serve as indirect evidence in our favor.

5 N^0-to-Cl^0 Movement

I now turn my attention to classifier-less nouns. In Bhattacharya (1999b), the explanation for the ungrammaticality of forms like **sɔptaho tin* 'week three' "the three weeks" is straightforward — NP movement occurs only when a classifier is present, because the classifier licenses [Specific]. However, following the proposal outlined thus far, NPs do not move to Spec,QP or any position in the domain of the classifier. Thus, we cannot make a similar move. In this section, I posit that Bangla undergoes a N^0-to-Cl^0 movement for classifier-less nouns. This is triggered by the failure to merge a nominalizer element *n* above the root lexical item, triggering the root to move to Cl^0 for interface purposes. I then derive the ungrammaticality of NP movement for classifier-less nouns from this approach.

First, I follow Simpson (2005) in claiming that if a classifier language contains a construction with no overt classifier, the apparent lack of a classifier is due to N^0-to-Cl^0 movement.[17] One piece of evidence in favor of such a theory is that in some languages the same morpheme may appear in both the N^0 and Cl^0 positions. Simpson (2005) suggests that in these cases the lower trace or the lower copy of the N^0 as well as the higher copy are pronounced. Examples of this are shown below: Thai in (21) and Burmese in (22). Presumably, *hoong* 'room' and *cun* 'island' appear in both the N^0 and Cl^0 positions, and hence are pronounced twice.

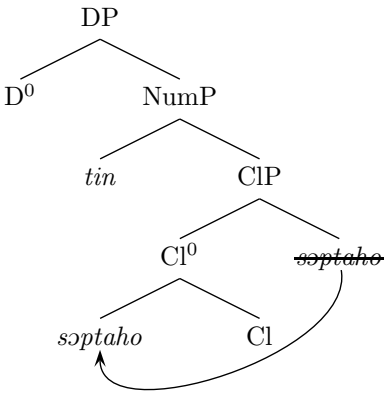
(21) *hoong* $[_{NumP}$ *saam hoong]*
room three room
'three rooms' (Thai)

(22) *cun* $[_{NumP}$ *ta cun]*
island one island
'one island' (Burmese)

Do we have evidence for this movement in Bangla? Note that in *tin sɔptaho* 'three week' "three weeks", N^0-to-Cl^0 movement would not change the word order, since the Cl^0 and N^0 positions are adjacent. The movement would be string-vacuous with these kinds of constructions. This is shown in (23).

[17] Simpson argues that these cases cannot arise due to a lack of a noun with the apparent noun undergoing a lexical operation that changes it into a Cl^0. There would be no predicate nominal for the classifier to individuate if such a lexical rule existed. In other words, a structure with a Cl^0 and no N^0 has an unsaturated Cl^0.

(23) *tin sɔptaho*



I propose that this movement operation does occur, however, and that it is specifically an instance of head movement, as shown above. As we will see shortly, intervening material, including adjectives, block this movement. If this were a phrasal movement, we might not expect intervening heads to block movement.[18]

In (3c) and (7c), repeated in (24) and (25), the classifier-less noun *sɔptaho* 'week' occurs in the same position as the classifier in the nominal headed by the classifier-compatible noun *kham* 'envelope'. A theory in which N^0-to-Cl^0 movement occurs before the operations that derive the quantificational approximateness forms, in (3c) and (7c), predicts these forms. Thus, we have some language-internal evidence for this kind of movement.

(24) *goʈa du=ek kham*
CL two=EK envelope
'two envelopes or so'

(25) *sɔptaho tin=ek*
week three=EK
'three weeks or so'

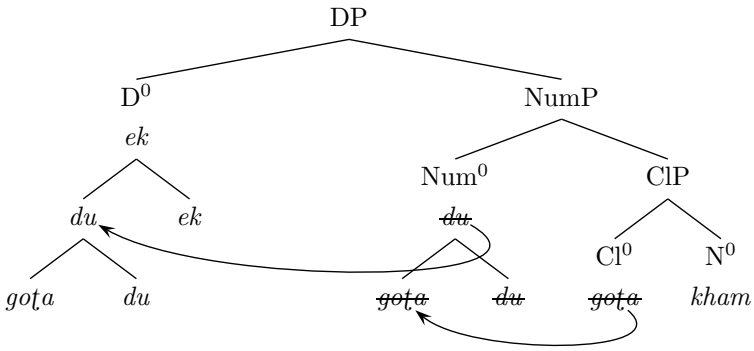
The structures for (24) and (25) are shown in (26). Here, the head movement of Cl^0 to Num^0, then Num^0 to D^0 represents the head movement used to express quantificational approximateness. This is formulated more specifically later. For now, notice that the appropriate word orders are predicted.

What drives N^0-to-Cl^0 movement? For nouns that do not denote spans of time, given in (6) and repeated in (27), I can at the moment only claim that it is a lexically specified feature. That is, some nouns may simple be marked in the lexicon as not combining with an overt classifier, but only a null classifier. Thus, we could claim that these nouns move to a phonetically null Cl^0, which then projects as a ClP.

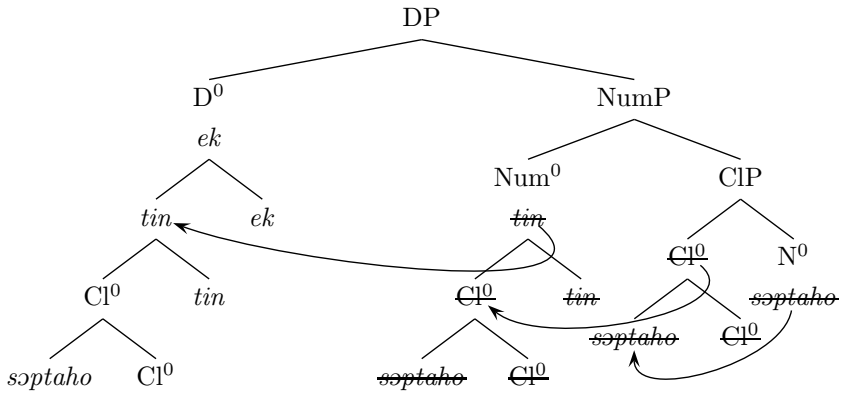
However, an account that relies solely on lexical specification does not capture all the data. Recall the referential distinction that Dasgupta (1983) describes for time-span denoting nouns. Essentially, a DP containing a noun of this category with no classifier refers only to the time span, whereas with a classifier the DP can refer to entities. This distinction is repeated in (28) from the original in (5).

[18] However, we also might not expect head movement to be blocked by an adjective either, if we take adjectives to be adjuncts. The kind of claim that I make explains why the adjunct adjective also blocks this movement, by suggesting that an adjective is only licensed when an optional *n* merges, and that N^0-to-Cl^0 movement only occurs when the nominalizer is not merged. Alternatively, we could also specify that the head-movement operation under discussion is actually one of Local Dislocation, a subspecies of a head-movement-like PF operation that is sensitive to adjuncts and heads alike (Embick and Noyer 2001).

(26) a. *goʈa du=ek kham*



b. *sɔptaho tin=ek*



(27) a. *du deʃ-er moitri*
two country-Gen friendship
'friendship between two countries'

b. *tin caka-r gaɽi*
three wheel-Gen vehicle
'three-wheeled vehicle'

c. *car paʃ*
four side
'four sides'

d. *tin dik*
three direction
'three directions'

e. *ora car bon tin bhai*
they four sister three brother
'They are four sisters and three brothers.'

(28) a. i. *John amar sathe du din thaklo*
John my with two day stayed
'John stayed two days with me.'

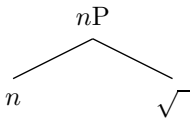
ii. *tin* *maʃ*
three month
'three months'

b. i. *du=ʈo* *bɔchor kharap gælo — 1966 ar 1976.*
two=CL year bad went — 1966 and 1976
'Two bad years occurred — 1966 and 1976.'[19]

ii. *du=ʈo* *maʃ-er* *nam-er* *ʃeʃ-e* *"ari"* *— januari ar phebruari.*
two=CL month-GEN name-GEN end-LOC "ari" — January and February
'Two months have names ending in *ari* — January and February'

How can this be explained? I propose that, following work in Distributed Morphology (Halle and Marantz 1993, Marantz 1997, Siddiqi 2009), lexical items are not inherently specified for a syntactic category. Instead, they gain their syntactic properties from functional heads that select the root. Thus, the structure of a noun is shown in (29), where *n* is the nominalizer, and $\sqrt{}$ represents the lexical root.

(29)



I propose that for classifier-less nouns the lexical root and a null classifier element are selected, and then the lexical root adjoins to Cl^0. I propose this adjunction occurs because lexical roots can only be phonologically realized if assigned a lexical category. Thus, movement to Cl^0 satisfies a PF requirement. Such an analysis stipulates that there is no "real" noun in these kinds of DPs. The implications of this will be clear shortly.

How does this explain the ungrammaticality of NP movement in forms like (30)?

(30) * *sɔptaho tin*
week three
intended: 'the three weeks'

Lexical roots are not inherently definite or indefinite. Only nouns and their projections can be used to refer to entities and have a definite or indefinite value. Thus, the *n*P — and no structure lower than the *n*P — must move to Spec,DP, since this is the lowest layer at which it is meaningful to describe a structure as definite or indefinite. Thus, the ungrammaticality of (30) is due to the fact that the NP cannot move to Spec,DP to check [DEF], as there is no NP to check the feature by virtue of the fact that the root *sɔptaho* has adjoined to the null Cl^0.[20]

Note that this puts two heads in a non-local relation. However, I do not think that there is anything that *a priori* rules out this kind of relation. That is, if we take the Bare Phrase Structure approach seriously, relations between phrases and individual lexical items are necessarily the same ilk. That is, if relating T^0 and D^0 (and thus DP) to check Case in English is an available option, subject to locality constraints, then the D^0-to-Cl^0 AGREE operation that I propose here should also be theoretically viable. However, note that this means we need to state locality constraints on head movement on *movement* as such and not on possible head-to-head relations, or else we will lose the ability to have T^0 and a subject DP enter in the appropriate relationship in English. Additionally, we will need to say something about why particular operations are instantiated by head movement, and why other operations are instantiated by phrasal movement. At this point, I suggest that is

[19] In Dasgupta's paper, the translation was 'There were two bad years — 1966 and 1976.

[20] A more formal way of wording this is that D^0 must utilize the operation AGREE to probe its c-command domain for *n* and check the feature [DEF]. Upon checking [DEF], D^0 checks the EPP feature [*n*] by moving *n*P to Spec,DP. In (30), there is no *n* to satisfy the [DEF] feature.

purely by stipulation, and is encoded in the generalized [Epp] feature that cues movement — that is, just as some Agree operations are instantiated by overt movement and idiosyncratically others are not, so is the distinction between phrasal and head movement.

There are a few ways around this constraint, all with some level of markedness. For one, some speakers allow a coerced reading in which the bare nominal can be construed as definite. Additionally, we have observed that nouns like *din* 'day' in fact can be used to refer to entities with a definite interpretation as long as they appear with a classifier. This is achieved by merging n with the lexical root, giving rise to a structure in which an overt classifier appears by bleeding the movement to a null Cl^0. Afterwards, [Def] on D^0 moves nP to Spec,DP as usual. This is shown in (32). This makes sense, since for a linguistic expression to be definite, the structure must be able to refer to something that can have definite status. Thus, an n is required to form a "real" noun. Alternatively, the grammar may merge a dummy demonstrative in D^0 (or, plausibly, higher in the structure) to check [Def], shown in (33), presumably as a last resort.[21]

(31) *ʃat din*
seven day
'(the) seven days'

(32) *din ʃat=ʈa*
day seven=Cl
'the seven days'

(33) ? *sei ʃat din*
that seven day
'those/the seven days'

The placement of adjectives corroborates this analysis. Let's turn our attention to the data below.

(34) a. *tin din*
three day
'three days'

b. i. * *tin din* [$_{NP}$ *kharap* ~~*din*~~]
three day bad
intended: 'three bad days'

ii. *tin=ʈe kharap din*
three=Cl bad day
'three bad days'

c. i. * *din tin=ek* [$_{NP}$ *kharap* ~~*din*~~]
day three(=Ek) bad
intended: 'three bad days or so'

ii. *goʈa tin=ek kharap din*
Cl three(=Ek) bad day
'three bad days or so'

Consider what an analysis in which N^0 alone moves out of NP would predict. The null Cl^0 selects an NP, whose head N^0 moves and adjoins to Cl^0. If an adjective were to adjoin to NP, we would expect the adjective to be stranded. This does not occur, as shown above. No N^0-to-Cl^0 movement occurs when an adjective appears. This is easy to account for in the model at hand. By hypothesis, adjectives are adjuncts to structures headed by nouns. There must be a phrase headed by a noun for the adjective to adjoin to. Thus, forming a nP is a prerequisite for licensing an adjective. If

[21] This latter strategy seemed marked for at least one informant, so it is shown with a ?.

there is an adjective, then there is a n to form structure to support the adjective. As stated earlier, movement to Cl^0 occurs to give the lexical root phonological support. However, the lexical root is already licensed by n in these structures. Because the grammar has no need to perform a movement to Cl^0, the root remains *in-situ* as the complement of n. This makes the prediction that adjectives with nouns like *sɔptaho* require the appearance of a classifier, shown in (34b-ii).[22]

6 Num^0 and Cl^0 as Separate Heads

Next, I turn my attention to the headedness of the numeral and the classifier. Are they a complex head like the Q^0 proposed by Bhattacharya (1999b), or are they represented as separate heads? Before directly addressing the issue in Bangla, I examine evidence for separate heads cross-linguistically. Note that, as shown in Simpson (2005), the numeral and classifier do not necessarily appear adjacent to one another across languages. For example, in Mandarin Chinese, an adjective may appear between the numeral and the classifier, demonstrated in (35).

(35) *yī xiǎo ben shū*
one small CL book
'one small book' (T'ung and Pollard (1982); Mandarin Chinese)

Likewise, in Nung, the numeral and classifier may appear on opposite sides of the head noun, shown in (36).

(36) *an ahn tahng nuhng ma*
take CL chair one come
'Bring a chair.' (Saul and Wilson (1980); Nung)

Additionally, as Simpson (2005) argues, the numeral and classifier serve different grammatical purposes. The numeral is used to convey the quantity of the noun, whereas the classifier serves to individuate the noun. In other words, the classifier serves a mechanical function in the grammar, and the numeral contributes to the truth conditions of the DP. Furthermore, if Dasgupta (1983) is correct in hypothesizing that classifiers define a value for an aggregate property, another grammatical function is associated with Cl^0, distinct from Num^0. Because of these differences, let's assume that Num^0 and Cl^0 are separate heads, potentially with a (crucially) postsyntactic morphological merger that cliticizes the classifier onto the numeral.

Do these generalizations hold for Bangla? Bhattacharya (1999b) claims that Q^0 is a complex head by citing the constituency tests below, shown in (37). These tests establish that the numeral and classifier cannot be separated. However, the tests only prove that the numeral and classifier are subsumed under a syntactic constituent — they do not provide evidence for the phrasal level. In fact, the tests are even compatible with a model in which the numeral and classifier are separate heads. Suppose the NumP contains ClP and ClP contains nP. If this is the case, then *tin* and *=ʈe* still fall within NumP, and thus form a constituent. We still predict the constituency tests below.

(37) a. *ei [tin-ʈe [$_{NP}$ lal boi]]*
this three=CL red book
b. *ei [[$_{NP}$ lal boi] tin=ʈe]*
this red book three=CL
c. *ʔtin=ʈe ei [$_{NP}$ lal boi]*
three=CL this red book

[22] These facts may be restricted to a few adjectives that merge low enough in the structure to modify the nominal root alone. For adjectives like *gɔto* 'last', as in *gɔto tin sɔptaho* '(the) last three weeks' which appear on the left of the numeral, there is no interference.

d. **tin-ei=ʈe [$_{NP}$ lal boi]*
three-this=Cl red book
'these three red books'

In the next section, I offer an analysis for quantificationally approximate DPs. This account is compatible with an analysis in which the numeral and classifier form separate heads.[23]

7 Quantificationally Approximate Reading

In this section I revisit the quantificationaly approximate DPs from (3c) and (7c), repeated in (38) and (39) respectively.

(38) *goʈa du=ek kham*
Cl two=(Ek) envelope
'two envelopes or so'

(39) *sɔptaho tin=ek*
week three(=Ek)
'three weeks or so'

To account for these data, I propose that D^0 may take a feature [Vague] to express a quantificationally approximate (or "vague") reading. Suppose that [Vague] requires a Cl^0 to adjoin to it in order to check the feature. I suggest that this occurs because Cl^0 is responsible for appropriately dividing the space of referrable entities, à la Borer (2005). That is, I propose that the quantificationally approximate reading is a function of vagueness or approximateness over the individuation and division of the reference of the noun. Thus, D^0 engages in an Agree relation with the Cl^0, and Cl^0 continuously goes through a left-adjoining head-movement operation towards the probe D^0, as formulated in Chomsky (2000).

This approach runs into an empirical difficulty with data like (40), in which we get left dislocation of the classifier, but no *ek*. Instead, we get doubling of the numerals.

(40) a. *gɔʈa du tin kham*
Cl two three envelope
'two envelopes or so'

b. *sɔptaho du tin*
week two three
'two weeks or so'

However, if we assume a Late Insertion-type model, in which phonological shape of a morpheme only takes form after Spellout, we may still be safe. Suppose that the [Vague] D^0 has no phonological information stored in its lexical entry. Instead, it may either occur as the form *ek*, or may parasitically gain its form in relation to the numeral that head-adjoins to it. Upon adjoining the Num^0 head, D^0 spells out as the next adjacent numeral. Thus, the pre-insertion form is given in (41), and the two phonological options are in (42). Of course, this then gives the grammar a non-deterministic flavor, which may not be desirable.

(41) *[$_D$ goʈa tin [Vague]] kham*
Cl three Vague envelope

[23] In pursuing this approach, if we want to have the numeral-classifier complex form a morphological unit, we will need to rely on post-syntactic rules to cliticize the classifier onto the numeral (or the NP with no distinct numeral). Note that we will most likely need to utilize some post-syntactic tinkering with the functional morphemes under discussion either way, as later I suggest that elements that fall under D^0 in quantificationally approximate constructions are not a morphological word. In other words, I am not requiring that morphological words map onto syntactic heads. This disassociation seems to be needed in some approaches to morphology, such as Embick and Noyer (2001).

(42) a. *goʈa tin car kham*
CL three four envelope
'three or four envelopes'

b. *goʈa tin=ek kham*
CL three=EK envelope
'three or four envelopes'

Why is D^0 the position for checking [VAGUE]? Nothing too crucial relies on this part of the analysis. All that matters is that the probe for Cl^0 is higher than Num^0 in order to derive the correct word orders for the quantificationally approximate structures, and I continue to assume that it is D^0 in the discussion for concreteness. However, there are some tentative arguments that D^0 should indeed be this probe, and that we may be able to avoid having to posit an additional head solely to act as the target for Cl^0-movement.

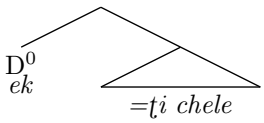
First, we find *ek* appearing with indefinites that are not quantificationally approximate.

(43) a. *chele*
boy
'(a) boy'

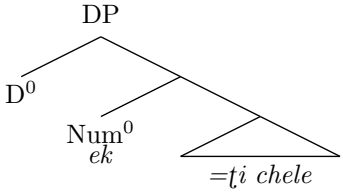
b. *ek(=ʈi) chele*
one=CL boy
'one boy' or 'a boy'

It is likely no accident that this morpheme is (largely) homophonous with the numeral *æk/ek*, making the form in (43b) ambiguous between two interpretations: 'one boy' and 'a boy'. For either interpretation, a classifier *=ʈi* may appear after *ek*. Because of the two interpretations, I posit there are two morphemes *ek*: one being a numeral 'one', the other an indefinite determiner, with structures shown below.[24]

(44) a.



b.



For simplicity, I assume that *ek* in (3c) and (7c), repeated in (45), merge in D^0. Presumably, it does not merge at Num^0, since there is already a numeral. This precludes an analysis in which *ek* merges at Num^0 and moves to D^0, for these structures at least.

(45) a. *goʈa du=ek kham*
CL two=(EK) envelope
'two envelopes or so'

b. *sɔptaho tin=ek*
week three(=EK)
'three weeks or so'

[24] This is obviously a common phenomenon, and can be seen across many languages and language families. Whether all of these languages truly have an ambiguity between 'one' and 'a' structurally is not relevant here.

There is another strategy for expressing quantificational approximateness that seems to be disjoint from the one under discussion, although it does show some similarity to the structures originally shown in (40). This alternate strategy is shown in (46).

(46) *tin car=ʈi kukur*
three four=Cl dog
'three or four dogs'

In this strategy, two adjacent integers appear together, with a classifier on the last numeral. In other words, for a numeral *n*, the expression *n, n*+1=Cl-NP can mean '*n* NPs or so'. One could argue that *ek* in (47) may be one of the two numerals used this way. In other words, this *ek* is *ek* 'one'.

(47) *jɔn car=ek sromik*
Cl four(=Ek) laborer
'four laborers or so' (Dasgupta 1983)

However, this seems unlikely. First, *ek* in (7c) and (47) is a clitic, signalled by the equal sign in the gloss. Secondly, the difference between the two numerals is larger than one, contrary to the pattern in (46). Additionally, the classifier has distinct positions in (46) and (47). Thus, I maintain that *ek* in (47) remains a distinct morpheme from the numeral, by hypothesis, in D^0.[25]

Thus, let's assume that *ek* appears in D^0. Otherwise, if *ek* is interpreted as a numeral, then it merges in Num^0. It is possible that the *ek* in quantificational approximate structures is a separate, third morpheme in a distinct position. However, for parsimony and concreteness, I assume that in quantificationally approximate DPs it is located in D^0.[26]

A curious fact that may be indirect support for the hypothesis that [Vague] merges at D^0 is that given below in the paradigm in (48). If we assume that [Vague] D^0 and [Def] D^0 are separate lexical items that must be in complementary distribution, we can explain the unacceptability of the form in (48b). In order to get NP movement, we need to have [Def] merging at D^0, precluding the D^0 that checks the [Vague]. The reverse case is also true if the quantificationally approximate D^0 merges.

(48) a. *goʈa du=ek kham*
Cl two=Ek envelope
'two envelopes or so'

b. * *kham goʈa du=ek*
envelope Cl two=Ek
intended: 'the two envelopes or so'

[25]Both Dasgupta (1983) and Bhattacharya (2001) note *ek* seems to be cliticized on quantifiers, shown in (1). An obvious exception is *sɔb* 'all'.

(1) a. *ɔnek* 'a lot'
b. *kɔek* 'a few' (*c.f. kɔe* 'few')
c. *khanik* 'a bit'
d. *ɔlpek* 'a little' (*c.f. ɔlpo* 'little')
e. *protteek* 'each one' (*c.f. proti* 'every')
f. *kɔtok* 'a few' (*c.f. kɔto* 'some, few')

Bhattacharya (2001) calls this *ek* 'Vague-one'. As I develop the model in this paper, I make no claim about this morpheme *ek*. Since I am not proposing a position for quantifiers in the DP apart from numerals, I have nothing to say about this morpheme. Though, it is not obvious to me that the *-k* needs to be analyzed as a morpheme at all. Etymology aside, it may just be that quantifiers generally end in *-k* — similar to question words in English beginning with *wh-* for the most part.

[26]If *ek* merges as D^0, then we might expect other numerals to merge. This would predict forms like **ek tin=ʈe chele* 'a three boys', which does not occur. However, we can sidestep this issue by suggesting that the vague *ek* and the indefinite singular *ek* are two separate D^0's.

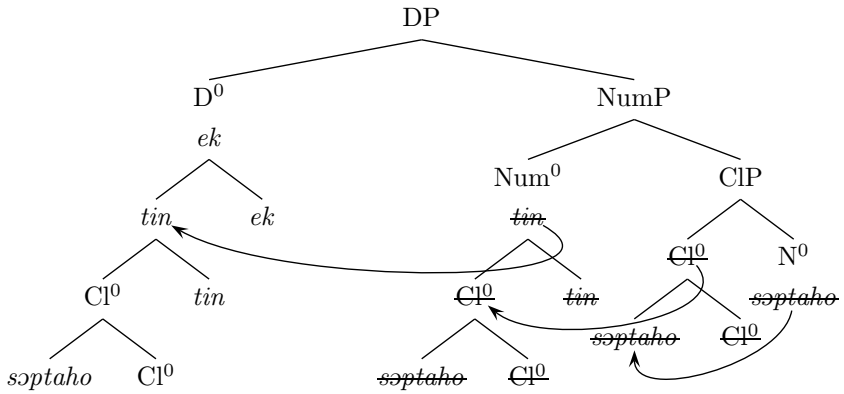
If Num^0 and Cl^0 are separate heads, as I propose, there is an intervening NumP shell between DP and ClP. If this is the case, then Cl^0 must adjoin to Num^0, which then must adjoin to D^0. The roundaboutness of this movement is due to the Head Movement Constraint (HMC), defined in (49). A more updated version of the HMC that also includes XP-level locality constraints is the notion of Relativized Minimality, defined in (50).

(49) **Head Movement Constraint:** Head movement may not skip intermediate heads (Travis 1984)

(50) **Relativized Minimality:** A movement operation cannot involve X and Y over a Z which is relevantly identical to Y in this configuration ...X...Z...Y... if Z c-commands Y. (Rizzi 1990, 2001, Hornstein 2006)

Suppose that this cyclical head movement is left-adjoining.[27] If this is the case, then the word order CL-NUM-EK-N would arise, as in (48a). Let's see how. First, Cl^0 adjoins to Num^0, then Num^0 adjoins to D^0. Additionally, if this movement were to occur after N^0-to-Cl^0 movement, the null hypothesis, then the N^0 would still arise in the Cl^0 position, correctly producing the word order N-NUM-EK for "classifier-less" nouns in the quantificationally approximate construction. These derivations are schematicized in (26), with the derivation of *sɔptaho tin=ek* repeated below in (51).

(51) *sɔptaho tin=ek*



8 Further Work

One problem still remains with this analysis. I have assumed that *ek* can merge in D^0 with structures that contain numerals, since it appears in the quantificationally approximate constructions. However, taking this at face value is incompatible with (52).

(52) *(*ek) car-ʈe boi*
EK four=CL book
'four books'

If *ek* can merge freely at D^0, then we would incorrectly predict it to be possible here as well. One could posit that *ek* has a conflict in a semantic/syntactic number feature, since the corresponding **a four books* in English is presumably ungrammatical for a similar reason. However, if that were so, then we would also predict *ek* to be illicit in quantificationally approximate structures, since these also contain numerals, supposing [VAGUE] merges as D^0. This brings us back to the non-trivial assumption that *ek* 'a' and *=ek* in the quantificational approximate constructions are the same syntactically. One method of dealing with this issue is to suppose that we have two *ek*'s that can merge in D^0 — the indefinite singular and the quantificationally approximate.

[27] See Bobalijk and Brown (1997) for a discussion of head movement and the Extension Requirement.

There is another set of provocative data. In (53), we see that the morpheme *=ek* can appear without triggering head-movement of the classifier. Instead, it seems to appear to the right of the numeral. One could suppose that this is the same operation driven by [Vague], except probing Num^0 instead of Cl^0, and triggering a more local head movement operation. Thus, the structure would look like (54). If this were the case, then it may not be so critical that D^0 probes Cl^0 *per se*, but requires some quantifying head. Sorting out whether there is a subtle semantic distinction between this form and the quantificational approximateness constructions that were the focus of this paper remains to be seen.

(53) *pãc=ek=ʈa (lɔmba) kham*
five=Ek=Cl (long) envelope
'five envelopes or so'

(54) $[_{DP}\ [_{D}\ pãc\ ek]\ [_{NumP}\ \sout{pãc}\ [_{ClP}\ goʈa\ [_{nP}\ kham]]]]$

This might suggest that we need to factor out the two steps in the Cl^0-to-D^0 operation. There could be a separate Cl^0-to-Num^0 movement that occurs, followed by a a Num^0-to-D^0 operation. This would then give a potentially more modular answer to why here we are only seeing Num^0-to-D^0 movement. However, I am not aware of what kind of evidence could tease out these two analyses.[28]

There are still many distributional questions left to address, as well. For instance, it is still unclear whether both numerals and quantifiers can fall under the same head. This does not seem to be the case, given forms like that in (55). Thus, quantifiers must have a distinct structural position. How this position would figure into the analysis presented here is unclear.

(55) *kon ʃat=ʈa boi*
any seven=Cl book
'any seven books'

It remains to be seen how possessors and demonstratives figure into the structure. If NP moves to Spec,DP, then in what position do possessors appear? Demonstratives occuring in D^0 seem to form separate heads in the syntax. This is consistent with the evidence found in the quantificational approximateness configurations and in other classifier languages.

Ultimately, this presents a model in which both the Bangla data and other Southeast Asian languages can fit comfortably. This provides us with a stronger theory of Universal Grammar that does not require the learner to choose between Bangla-type and, say, Chinese-type classifier languages. Instead, the child is pre-equipped with the structural frame that can accomodate a variety of constructions cross-linguistically, with the features and their methods of instantiation varying. This ultimately should be the goal of comparative syntax — reducing superficially different phenomena across a variety of languages to an appropriately flexible base that does not force the child to choose between two options that would only give rise to minimal differences.

Acknowledgments

Thanks to Priyanka Biswas, Syed Saurov, Samira Choudhury, Aniqua Hasan, Nandita Rahman, and Atif Safi for their help in providing data and judgments. Thanks to Jeanette Gundel, Brian Reese, Howard Lasnik, and Colin Phillips for their comments. Special gratitude is due to Hooi Ling Soh for her guidance and supervision. Finally, thanks to Kaitlin Johnson, Andrew Borne, Alissa Hoyme, James Imes, Kevin Myers, and the UCLA–UC Berkeley Joint Conference on the Languages of Southeast Asia for their comments aud suggestions. All remaining errors are my own.

[28] Curiously, the NP seems to be able to be left dislocated in these examples without the definite interpretation: *(lɔmba) kham pãc=ek=ʈa* 'five envelopes or so'.

References

Abney, Steven P. 1987. *The English Noun Phrase in its Sentential Aspect*. Ph.D. thesis, MIT.

Aikhenvald, Alexandra Y. 2000. *Classifiers: A Typology of Noun Categorization Devices*. Oxford: Oxford University Press.

Bernstein, Judy B. 1993. The syntactic role of word markers in null nominal constructions. *Probus* 5(1):25–38.

Bernstein, Judy B. 2001. The DP-hypothesis: identifying clausal properties in the nominal domain. In M. Baltin and C. Collins, eds., *The Handbook of Contemporary Syntactic Theory*, pages 563–561. Oxford: Blackwell.

Bhattacharja, Shishir. 2007. *Word Formation in Bangla: A Whole Word Morphological Description and Its Theoretical Implications*. Lincom Europa.

Bhattacharja, Shishir. 2008. Bangla namomguler antorgoThon. *The Dhaka University Journal of Linguistics* 1(2):67–89.

Bhattacharya, Tanmoy. 1999a. Specificity in the Bangla DP. In R. Singh, ed., *Yearbook of South Asian Languages and Linguistics 2*, pages 71–99. New Delhi/London: Sage Publications.

Bhattacharya, Tanmoy. 1999b. *The Structure of the Bangla DP*. Ph.D. thesis, University College London.

Bhattacharya, Tanmoy. 2001. Numeral/quantifier-classifier as a complex head. In H. van Riemsdijk and N. Corver, eds., *Semi-Lexical Heads*, pages 191–222. Berlin: Mouton de Gruyter.

Bobalijk, Jonathan David and Samuel Brown. 1997. Interarboreal operations: Head movement and extension requirement. *Linguistic Inquiry* 28(2):345–356.

Borer, Hagit. 2005. *In Name Only*, vol. 1 of *Structuring Sense*. Oxford: Oxford University Press.

Cheng, Lisa L.-S. and Rint Sybesma. 1999. *yi-wan tang, yi-ge Tang*: Classifiers and massifiers. *The Tsing Hua Journal of Chinese Studies* 28(3):385–412.

Cheng, Lisa L.-S. and Rint Sybesma. 2005. Classifiers in four varieties of Chinese. In G. Cinque and R. S. Kayne, eds., *The Oxford Handbook of Comparative Syntax*, pages 259–292. Oxford: Oxford University Press.

Chierchia, Gennaro. 1998. Reference to kinds across languages. *Natural Language Semantics* 6:339–405.

Chomsky, Noam. 1995. *The Minimalist Program*. MIT Press.

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In R. Martin, D. Michaels, and J. Uriagereka, eds., *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*, pages 89–155. Cambridge, MA: The MIT Press.

Dasgupta, Probal. 1983. On the Bangla classifier *ta*, its penumbra, and definiteness. *Indian Linguistics* 44(1–4):10–26.

Dasgupta, Probal. 1985. On Bangla Nouns. *Indian Linguistics* pages 37–65.

Dasgupta, Probal. 1997. La définitude en bengali. *Faits de Langues* 5(10):137–144.

Dasgupta, Probal. 2003. Bangla. In G. Cardona and D. Jain, eds., *The Indo-Aryan Languages*, pages 351–390. New York: Routledge.

Dirdal, Hildegunn. 2004. *Second Language acquisition of articles and plural marking by Benagli speakers of English*. Ph.D. thesis, University of Cambridge.

Embick, David and Rolf Noyer. 2001. Movement operations after syntax. *Linguistic Inquiry* 32(4):555–595.

Ghosh, Sanjukta. 2006. Honorificity-marking words of Bangla and Hindi: Classifiers or not? *Bhashachintana* 1:21–27.

Giusti, Giuliana. 1991. The categorial status of quantified nominals. *Linguistische Berichte* 136:438–452.

Gundel, Jeanette K., Nancy Hedberg, and Ron Zacharski. 2001. Definite descriptions and cognitive status in English: why accomodation is unnecessary. *English Language and Linguistics* 5(2):273–295.

Halle, Morris and Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In K. Hale and S. J. Keyser, eds., *The View from Building 20*, pages 111–176. Cambridge, MA: The MIT Press.

Hornstein, Norbert. 2006. Deriving c-command. Unpublished Ms., Universtiy of Maryland, College Park.

Jackendoff, Ray. 1977. *X-bar-Syntax: A Study of Phrase Structure*. No. 2 in Linguistic Inquiry. MIT Press.

Junghare, Indira. 1983. Markers of definiteness in Indo-Aryan. *Indian Linguistics* 44(1–4).

Kornai, András and Geoffrey Pullum. 1990. The X-bar theory of phrase structure. *Language* 66:24–50.

Löbel, Elisabeth. 1989. Q as a functional category. In C. Bhatt, E. Löbel, and C. Schmidt, eds., *Syntactic phrase structure phenomena*, pages 133–57. Amsterdam: John Benjamins.

Longobardi, Giuseppe. 2003. The structure of DPs: Some principles, parameters, and problems. In M. Batlin and C. Collins, eds., *The Handbook of Contemporary Syntactic Theory*, pages 562–603. Oxford: Blackwell.

Marantz, Alec. 1997. No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. In A. Dimitriadis, L. Siegel, C. Surek-Clark, and A. Williams, eds., *Proceedings of the 21st Annual Penn Linguistics Colloquium: Penn Working Papers in Linguistics*, vol. 4, pages 201–225.

Nunes, Jairo. 2001. Sideward movement. *Linguistic Inquiry* 32(2):303–344.

Rizzi, Luigi. 1990. *Relativized Minimality*. Cambridge, MA: The MIT Press.

Rizzi, Luigi. 2001. Relativized minimality effects. In M. Baltin and C. Collins, eds., *The Handbook of Contemporary Syntactic Theory*, pages 89–110. Oxford: Blackwell.

Sahoo, Kalyanamalini. 1999. The syntax of definiteness in Oriya. In R. Singh, ed., *Yearbook of South Asian Languages and Linguistics 2*, pages 101–117. New Delhi: Sage Publications.

Saul, Janice E. and Nancy Freiberger Wilson. 1980. *Nung Grammar*. Summer Institute of Linguistics.

Siddiqi, Daniel. 2009. *Syntax within the Word: Economy, allomorphy, and argument selection in Distributed Morphology*. No. 138 in Linguistics Today. Amsterdam: John Benjamins.

Simpson, Andrew. 2005. Classifiers and DP structure in southeast Asia. In G. Cinque and R. S. Kayne, eds., *The Oxford Handbook of Comparative Syntax*, pages 1–22. Oxford: Oxford University Press.

Soh, Hooi Ling. 2001. The syntax and semantics of phonological phrasing in Shanghai and Hokkien. *Journal of East Asian Linguistics* 10:37–80.

Soh, Hooi Ling and Jenny Yi-Chun Kuo. 2005. Perfective aspect and accomplishment situations in Mandarin Chinese. In H. J. Verkuyl, H. de Swart, and A. van Hout, eds., *Perspectives on Aspect*, pages 199–216. Berlin: Springer.

Travis, Lisa. 1984. *Parameters and Effects of Word Order Variation*. Ph.D. thesis, MIT.

T'ung, Ping-Chen and David E. Pollard. 1982. *Colloquial Chinese*. Oxford: Routledge.

An experimental study of Hindi and English perfective interpretation

Sudha Arunachalam & Anubha Kothari, *Boston University & eBay, Inc.*[1]

Received December 2010; Revised November 2011

Abstract

It has been observed for several languages, including many South Asian languages, that some perfective forms do not entail completion of the events they describe. We explore this phenomenon in the current study, contributing experimental and cross-linguistic perspectives. We compare perfective interpretations in Hindi and in English. Using a variety of predicate types, we test the predictions of a semantic account of perfective interpretation, which predicts clear-cut differences between Hindi and English as well as clear patterns for certain types of predicates. Our results indicate that a semantic account is inadequate to capture speakers' interpretations of these perfectives, and suggest that a variety of factors, including contextual ones, may be at play. In addition to contributing empirical data to our understanding of this puzzling phenomenon, these results also indicate a role for experimental investigation into semantic phenomena.

1 Introduction

In a number of languages, perfective verbs can be used to describe events that do not terminate at their natural endpoints (Japanese: Ikegami 1985; Hindi: Kothari 2008, Singh 1998, Verma 1993; Karachay-Balkar: Tatevosov 2007; Malagasy: Travis 2000; Mandarin: Teng 1972; Punjabi: Raja 2003; Salish languages St'át'imcets and Skwxwú7mesh: Bar-el Davis & Matthewson 2005; Tagalog: Dell 1987; Tamil: Pederson 2007; Thai: Koenig & Muansuwan 2000). In Hindi, the simple perfective verb form (SV) in (1) can describe eating events with arbitrary endpoints (e.g., the cookie is only partially eaten).[2]

(1) maya-ne biskuT-ko khaa-yaa par us-e puuraa nahiin khaa-yaa
Maya-Erg cookie-Acc eat-Perf but it-acc full not eat-Perf
'Maya ate a cookie but not completely.'

Note that the corresponding English sentence (*#Maya ate a cookie, but not completely*) is infelicitous.

SVs can also describe events with natural endpoints (e.g., the cookie is eaten completely). In fact, the default interpretation of the SV is that the event came to full completion, if the endstate is not canceled. This indicates that event culmination is an implicature for the SV, rather than an entailment.

[1] Kothari was a student at Stanford University when this research was carried out.

[2] These are often called non-culminating accomplishments, but we avoid that term here because the present data question the phenomenon's limitation to only accomplishment predicates.

JSAL Volume 4, Issue 1, December 2011.

Hindi has a second perfective, the complex verb form (CV) shown in (2). The main verb appears in bare root form, and is followed by a light verb in the perfective. This construction only describes events terminating at their natural endpoints.[3]

(2) maya-ne biskuT-ko khaa-li-yaa #(par us-e puuraa nahiin khaa-yaa)
Maya-Erg cookie-Acc eat-take-Perf but it-acc full not eat-Perf
'Maya ate a cookie but not completely.'

There is currently no consensus on the appropriate analysis of this phenomenon: How should we account for the availability of partially-completed interpretations, precisely when are they available and with what kinds of verbs, and what differs between the SV and CV? For Hindi specifically, we are aware of one major account, a semantic one, from Singh (1991, 1998), who proposes an aspectual distinction between the SV and CV. She invokes a new type of perfective, the neutral perfective, which is realized as the SV form. It imposes boundedness on events, but not necessarily at their natural telic endpoint. Singh assumes a homomorphism approach; parts of the affected theme object are mapped onto parts of the event such that the theme object is affected incrementally as the event proceeds (e.g., half of the cookie is eaten halfway through the eating event). For the neutral perfective, only part of the theme object need be affected. To produce the natural endpoint reading of the CV form, Singh introduces a totality operator *TOT*. When *TOT* is applied, the entirety of the object must be subjected to the event. This is the standard perfective type in English, but only arises with the presence of the light verb in Hindi. Under Singh's account, the difference between Hindi and English therefore lies in whether the neutral perfective is available.

Because it relies on homomorphism, the neutral perfective accounts beautifully for the phenomenon of partial completion interpretations, but only in accomplishment predicates with incremental themes. This account makes several clear predictions: 1) only accomplishment predicates involving an incremental theme should show the SV/CV distinction; 2) the partial completion interpretation that arises with the SV should be one in which the theme is incrementally affected; 3) all predicates with incremental themes should behave uniformly in showing this distinction; 4) clear distinctions are expected between Hindi, which has the neutral perfective, and English, which does not.

Some doubt is cast on the first prediction by data from Hindi and other languages. Several examples of this phenomenon in non-accomplishment predicates have been raised in the literature for other languages (e.g., Ikegami 1985, Koenig & Muansuwan 2000, Teng 1972, Pederson 2007). Kothari (2008) provides examples from Hindi as well:

(3) maya-ne kamiiz Taang-ii par vah Tangii nahiin
Maya-Erg shirt hang-Perf but it-acc hung not
'Maya hung the shirt, but it didn't get hung.'

(4) maya-ne saikil chalaa-yii par vah chalii nahiin
Maya-Erg bicycle ride-Perf but it-acc moved/worked not
'Maya rode the bicycle, but it didn't work.'

If there is indeed no clear-cut difference between accomplishment and other types of predicates when it comes to the SV/CV distinction, we must consider other factors that may be relevant, e.g., contextual or pragmatic factors. We hope to lend some preliminary insight into what these factors may be in the current study.

This study explores patterns of interpretation of sentences describing partially- and fully-completed events. Our goals are to explore the range of events for which speakers accept change-of-state perfective predicates as descriptions of partially-completed events, both within-language (for different syntactic constructions), and across languages (comparing Hindi and English). Our results will shed light on whether the Hindi SV/CV distinction can be explained by Singh's (1991, 1998)

[3]The CV construction is also associated with aspects of meaning beyond event completion, including affective readings and indications of surprise. See Hook (1976) for discussion.

semantic account only, or whether other factors, e.g., pragmatic and contextual ones, must also play a significant role. A full account of this phenomenon will likely include several interacting factors, including both semantic and pragmatic ones.

1.1 Previous Experimental Work

Previous experimental work has lent some insight into when speakers of Hindi-type and English-type languages permit partial completion interpretations. Pederson (2007) provides evidence that speakers of Tamil, which is similar to Hindi in having SV and CV constructions, are willing to accept SV sentences as descriptions of incomplete events. He showed native Tamil speakers video clips of scenes depicting either incomplete events (e.g., pushing on a door, only moving it a little) or fully-completed events (e.g., pushing the door to fully closed) and elicited yes/no responses to questions in SV form (e.g., *aTai-tt-aan-aa?* "did he close (it)?"). The percentage of "yes" responses varied by predicate, but the average across predicates was 39%. Pederson took this 39% acceptance rate to indicate that Tamil speakers do indeed allow SV sentences to refer to events that have not reached their endstate.

Wittek (2002) conducted a very similar task, with the same video events, in German. German, like English, does not have an SV/CV distinction, and change-of-state predicates generally entail full completion to the natural endpoint. She found, as predicted, that German speakers never accepted German perfectives as descriptions of incomplete events.

The results of these two experimental studies provide experimental support for the difference between Tamil and German. This unique approach lends important insight, and suggests a role for experimental study of this type of semantic phenomenon.

1.2 Comparison to the Current Study

We take a similar experimental approach in the current study, showing participants videos of events that arrive at different states of completion, and eliciting judgments for different sentence types. This allows us to make controlled comparisons across events, linguistic conditions, and languages.

However, we also make several changes to Wittek's and Pederson's approaches. First, instead of studying speakers' interpretations only of simple perfective forms, in Experiment 1 we elicit Hindi speakers' responses to both SV and CV sentences. This provides a natural baseline against which to compare responses. We also tested two constructions in English. Because English does not have an SV/CV distinction, in Experiment 2a we presented only simple perfective predicates, and in Experiment 2b we compared speakers' interpretations of these simple perfectives with constructions incorporating the particle "up" (e.g., *ate up*).

Second, rather than testing fully-completed events against partially-completed events in which the theme object was only barely acted upon, as Pederson and Wittek did, we instead used partially-completed events in which the intended action arrived quite close to its natural endpoint, but terminated noticeably earlier. For example, in the partially-completed version of the door-closing event in the current study, the actor executed the same motion as she did with the fully-completed version, pushing the door until it was about 80% closed. We used these almost-completed events for our partial-completion condition for two reasons. First, we hypothesized that these events would be construed as intentional acts that one might perform. For example, closing a door to keep out noise, but allow the cats to roam freely, might only require closing it most of the way, while closing it to keep out the cats entirely requires closing it all the way. Closing the door only a little, on the other hand, may be construed either as an activity or a failed attempt. Using almost-completed events therefore allowed us to control somewhat for how participants perceived the actors' intentions in the partially-completed and fully-completed conditions. Second, we surmised that these almost-completed actions would make the most stringent test case. If, despite their close perceptual and conceptual similarity, our partially- and fully-completed conditions yield differences across sentence types and/or across languages, then we have strong evidence that speakers attend closely to endstates. It is important to note, however, that scales of completion differ across verbs; in all cases we use our own intuition

to judge what kinds of events for a particular verb would best represent partial and full completion. (Stimuli are available from the first author on request.)

In addition to these important differences, our study shares an important feature with Pederson's and Wittek's studies. Like them, we chose a variety of predicate types in this study, including accomplishments with incremental themes (e.g., *eat*), as well as achievements (e.g., *pluck*). This is particularly important for our current goal of looking at patterns of interpretations across predicate types and across languages. A semantic approach to the SV/CV distinction relying on the presence of an incremental theme would predict systematic differences between our incremental theme and other predicates.

The events we used for two of our predicates differed along another dimension as well. For one of our *extinguish* events and both of our *wake* events, the videos for the fully-completed experimental condition proceeded as expected, with the event arriving at its natural endpoint. But in the partially-completed condition, the event also arrived at its natural endpoint, but then, the theme participant returned to its original state. In the *extinguish* video, the agent blew out a trick candle; it extinguished completely, but then re-lit two seconds later. For the two *wake* videos, the sleeping participant was woken by the agent, opening his eyes fully, but he closed them again and appeared asleep at the end of the video clip. We introduced this additional dimension of returning to the initial state in part because it was the most natural way to portray an incomplete or unsuccessful version for those predicates, but also because we wanted to discover whether the final state added any relevance for speakers' construals of completion. These trials allow us to investigate whether a partial completion interpretation for an SV sentence is necessarily one in which the theme object is incrementally affected. Alternatively, it could be the case that in addition to incremental partial completion, SV sentences are used in situations in which an action has not been completed to the satisfaction of the actor's perceived intention. If speakers allow SV, but not CV, sentences for these events involving a return to the initial state, then this is support for the latter hypothesis.

1.3 Goals of the Current Study

Here we report the results of two experiments investigating Hindi speakers' (Experiment 1) and English speakers' (Experiments 2a & 2b) interpretations of perfective forms describing partially- and fully-completed events. We had three goals. First, to verify the validity of our experimental approach, we wanted to ascertain whether Hindi speakers show the expected distinction between SV and CV sentences: If Hindi SV sentences can describe partially-completed events, Hindi speakers should accept SV more often than CV sentences as descriptions of these events. Second, we wanted to assess the range of predicates across which the Hindi partial-completion interpretation for SV sentences is available. Third, we wanted to explore English speakers' interpretations for the same events, to determine whether the pattern resembles that for Hindi, despite systematic cross-linguistic differences in how perfectives are realized syntactically.

2 Experiment 1: Hindi

We began by exploring Hindi speakers' judgments of SV and CV perfectives as descriptions of partially- and fully-completed events. Participants viewed events in which an actor performed an action on a theme participant, bringing about, either partially (mostly) or completely, a change of state. We presented participants with an SV or CV statement and elicited a true/false judgment as to whether the statement was an appropriate description of the event they had viewed.

2.1 Methods

2.1.1 Participants

Twenty-four adult native speakers of Hindi from Jodhpur, India, participated. All participants gave oral consent as approved by Stanford University's Institutional Review Board. Although many of the Hindi speakers were bilingual in Marwari, all were fluent speakers, readers, and writers of Hindi, who used Hindi daily.

2.1.2 Materials

Visual stimuli. For each of eight verbs (e.g., *eat*, see Table 1 for a complete list), we filmed two pairs of short video clips. Each pair featured different event participants (e.g., an actor eating a cookie, a different actor eating a chocolate bar). One video of each pair depicted a partially-completed event (e.g., eating half of the cookie) and the other depicted a fully-completed event (e.g., eating all of the cookie). The videos in each pair were edited to be the same length.

Each video ended with a still frame of the last frame of the clip, during which the audio was presented. This ensured that participants could view the final endstate of the action as they heard the sentence and gave their response.

Auditory stimuli. A native speaker recorded two sentences for each video, one describing the event using an SV form (e.g., *us-ne biskuT-ko*[4] *khaa-yaa*, "She ate a cookie") and one using a CV form (e.g., *us-ne biskuT-ko khaa li-yaa*). See Table 2 for a complete list of sentences.

TABLE 1 Description of video scenes. In the partial-completion condition, these events did not reach their natural endpoint (e.g., For Scene 1 *close*, the woman closed the door 3/4 of the way.)

Predicate	Scene 1	Scene 2
close	Woman closes a door	Woman closes a drawer
cover	Hand covers a pot with its lid	Hand covers a basket with its lid
draw	Hand draws a flower	Hand draws a circle
eat	Woman eats a cookie	Woman eats a chocolate bar
extinguish	Woman blows out a candle (In the partial-completion condition, the candle re-lights.)	Woman runs a lit piece of newspaper under the faucet
fill	Hand fills a glass with milk	Hand fills a glass bowl with marbles
pluck	Hands pluck a small twig off a larger branch	Hands pluck a banana off a bunch
wake	Woman jostles a sleeping man lying on a couch (In the partial-completion condition, his eyes open briefly but he closes them again)	Woman jostles a sleeping man seated in a chair; In the partial-completion condition, his eyes open briefly but he closes them again

2.1.3 Design and Procedure

Each participant viewed one video from each pair for each verb. For a given verb, both videos depicted either the fully-completed event or the partially-completed event, but each participant saw both fully-completed events (for four predicates) and partially-completed events (for the other four). Participants were randomly assigned to an SV-first or CV-first condition: Those in the SV-first condition heard SV sentences for the first half of the videos they saw (i.e. the first instance of each predicate), and CV sentences for the second half, and vice versa for CV-first participants. There were no differences in performance between SV-first and CV-first participants, so we collapse across these in the analysis. Videos were presented in a different pseudo random order for each participant. At the end of each video, participants heard the target sentence and were asked whether it was a true or false description of the video they had just viewed.

2.1.4 Predictions

If participants are sensitive to the SV/CV distinction, they are expected to show different responses for partially-completed events depending on syntactic condition, with SVs accepted more often

[4]The direct object of the verb bore accusative case marking in all sentences except for the "draw" sentences–in cases where the object comes into existence over the course of the event, accusative case marking is awkward.

TABLE 2 List of sentences for Scene 1 events, Experiment 1.

Predicate	SV	CV	English Translation
close	us-ne darvazee-ko band ki-yaa	us-ne darvazee-ko band kar li-yaa	She closed the door.
cover	us-ne bartan-ko Dhakaa	us-ne bartan-ko Dhak li-yaa	She covered the pot.
draw	us-ne phuul banaa-yaa	us-ne phuul banaa li-yaa	She drew a flower.
eat	us-ne biskuT-ko khaa-yaa	us-ne biskuT-ko khaa li-yaa	She ate the cookie.
extinguish	us-ne mombatti-ko bujhaa-yaa	us-ne mombatti-ko bujhaa li-yaa	She extinguished the candle.
fill	us-ne duudh-se gilaas-ko bhar-aa	us-ne duudh-se gilaas-ko bhar li-yaa	She filled the glass with milk.
pluck	us-ne Daalii-ko toDaa	us-ne Daalii-ko toD li-yaa	She plucked the twig.
wake	us-ne us-ko jagaa-yaa	us-ne us-ko jagaa li-yaa	She woke him.

than CVs. Because CV perfectives entail completion of the events they describe, we expect a near 0% acceptance rate for CV sentences as descriptions of partially-completed events. Because SV perfectives are hypothesized to allow partial completion interpretations, these should have a much higher acceptance rate. Finally, because both SV and CV sentences are felicitous descriptions of fully-completed events, we expect a 100% acceptance rate for fully-completed events, regardless of syntactic condition.

Because we are also interested in participants' responses to a range of predicates, we make a further prediction about the predicates to which the SV/CV distinction should apply. Specifically, we test the predictions of Singh's semantic account. If partial completion interpretations arise via a homomorphism between the measuring out of the event and the theme object, then we expect only the accomplishment predicates with incremental themes (*cover, draw, eat, fill*) to show the above pattern. For other predicates, both SV and CV sentences should only be acceptable for fully-completed events. But if pragmatic factors play an important role, there may not be clear differences between accomplishment and other types of predicates.

2.2 Results and Discussion

We calculated the mean proportion of "true" responses for each participant. As predicted, for partially-completed events, responses differed by syntactic condition, with a lower acceptance rate for CV sentences (29%) than SV sentences (53%). For fully-completed events, participants accepted both SV and CV sentences (99.5%).

To assess these patterns statistically, we first transformed the proportion data using an empirical logit function, and fit the transformed data using a multi-level model treating Event Completion (2: partial vs. full) and Syntax (2: SV vs. CV) as fixed factors. The beta coefficients for the models are reported in Table 3.[5] The analysis reveals that both Event Completion and Syntax are reliable predictors of participants' responses, as is their interaction. As predicted, these results indicate that speakers are indeed sensitive to the SV/CV distinction.

A glance at the results for individual predicates reveals a high degree of variability. In Table 4, for each predicate we tallied the percentage of "true" responses in each Event Completion and each Syntax condition. The results provide some insight into how different predicates behave.

Considering only the results for *partially-completed* events, we make three observations related

[5]Note that S.E. stands for "Standard Error".

TABLE 3 Experiment 1 (Hindi): Estimates of fixed effects from best-fitting multi-level model of proportion of "true" responses (empirical logit transformed)

Effect	Estimate	S.E.	*t*-value
Intercept	1.22	0.15	8.36*
Event Completion	-3.16	0.23	-13.93*
Syntax	0.62	0.23	2.74*
Event x Syntax Interaction	1.40	0.45	3.09*

$p < 0.05$ (on normal distribution)

TABLE 4 Experiment 1 (Hindi): Percent "true" responses, by condition and trial

Predicate	Partial Completion		Full Completion	
	SV	CV	SV	CV
Incremental Theme				
cover	67	33	100	100
draw	38	42	100	100
eat	83	33	100	100
fill	75	58	100	100
Non Incremental Theme				
close	25	33	100	100
extinguish				
Newspaper (partial)	20	14	99	100
Candle (return to initial state)	57	40	100	100
pluck	17	0	100	100
wake (return to initial state)	83	17	100	100

to our predictions. First, a semantic account predicts that the SV/CV distinction should arise only for accomplishment predicates with incremental themes. Therefore, a strong prediction under a semantic account would be that these predicates (*draw, cover, eat, fill*) show the distinction and other predicates (*close, extinguish, pluck, wake*) do not. This prediction was not borne out. All of the non-incremental-theme predicates showed a difference between SV and CV judgments in the expected direction. To address this statistically, we divided the predicates into incremental theme and non-incremental theme, and fit a second multi-level model on only the data for partially-completed events, including presence/absence of incremental theme and syntax as fixed factors. If only incremental theme predicates show a distinction, then we predict a significant interaction between incremental theme and syntax; that is, syntax affects responses, but only for incremental theme predicates.

The beta coefficients for the model are reported in Table 5. The analysis reveals that Incremental Theme is a reliable predictor, while Syntax is marginally reliable ($p = 0.1$), but their interaction is not. What this means is that although incremental theme predicates yield a higher overall proportion of "true" responses, the lack of an interaction with Incremental Theme reveals that the difference in responses to SV versus CV sentences is not greater for incremental theme predicates than it is for non-incremental theme predicates. In fact, the effect size of the difference between SV and CV is similar for both non-incremental-theme and incremental theme predicates (Cohen's $d = 0.58$ for non-incremental-theme predicates and 0.54 for incremental theme predicates), indicating that the effect of syntactic construction is similar for both predicate types.

Interestingly, the predicate whose results most clearly fit the expected pattern was *wake*, with a high acceptance rate for SV sentences (83%), and a low acceptance rate for CV sentences (17%). This finding for *wake* is particularly striking given the unusual events we presented in the partially-completed condition. In the partially-completed *wake* events, the event reached its complete natural

TABLE 5 Experiment 1 (Hindi): Estimates of fixed effects from best-fitting multi-level model of proportion of "true" responses (empirical logit transformed) for partially-completed events only, with Incremental Theme and Syntax as fixed factors

Effect	Estimate	S.E.	*t*-value
Intercept	0.12	0.13	0.90
Syntax	0.36	0.23	1.61†
Incremental Theme	0.49	0.16	3.05*
Syntax x Incremental Theme Interaction	0.23	0.32	0.73

$*$ $p < 0.05$ (on normal distribution)
† $p = 0.1$ (on normal distribution)

endpoint, but then retracted to an earlier state: A man was jostled into waking, opening his eyes, but then closing them again and returning to a sleep state. Not only does *wake* not have an incremental theme or a process component (cf. Butt and Ramchand 2005), but in this particular event, the actor both reaches the final state (waking), and ends the event in a state of sleep. This event is not therefore a typical partial-completion event, and questions a semantic account's reliance on an incremental progression toward completion. The results for *wake* suggest that in addition to event termination at an arbitrary endpoint, the SV/CV distinction is also relevant for events that do reach a natural endpoint, but later return to the initial state. For CVs, the theme participant must be in the relevant endstate at the time of evaluation (here, when the video ends), even if the change of state did proceed to completion at some prior point. For SVs, the endstate need not be present at the time of evaluation.

Finally, note that there is quite a bit of variability among predicates in acceptance of SVs and CVs for partially-completed events. Clearly, participants' judgments were not clear-cut, with CVs never allowing partial completion interpretations and SVs often or always allowing them. For example, while the results for *eat* and *wake* are closest to the expected pattern, *fill* yielded a very high acceptance rate for both SVs (75%) and CVs (58%), and *pluck* yielded a very low acceptance rate for both (17%, 0%, respectively). This indicates considerable gradience in speakers' judgments of these sentence-event pairings.

Taken together, these findings do provide some support for Singh's account. Overall, accomplishment predicates with incremental themes did show the SV/CV distinction, while some achievements like *pluck* showed very low acceptance rates. However, the findings also suggest that a strong semantic account cannot be the whole story. The evidence for gradience in speakers' judgments, similar effect sizes, and the unexpected responses to some predicates, like *wake*, suggest that other factors are also at play. We return to these ideas in the General Discussion (section 5). In Experiment 2, we ask how English speakers respond to these same events and predicates, despite there being no SV/CV distinction in English.

3 Experiment 2a: English

Experiment 2 explores English speakers' interpretations for the same events and translation-equivalent predicates. Unlike in Hindi, English perfective change-of-state verbs are understood to entail the achievement of their endstate. Although they resemble the Hindi SV in form, in meaning they are expected to match the Hindi CV. Therefore, English speakers should show a low acceptance rate for simple perfective sentences as descriptions of partially-completed events, parallel to Hindi speakers' low acceptance for CV sentences.

3.1 Methods

3.1.1 Participants

Twenty-four native English speaking Northwestern University undergraduates participated for course credit. All participants signed an informed consent form approved by Northwestern University's

Institutional Review Board.

3.1.2 Stimuli

Visual stimuli were identical to Experiment 1. Auditory stimuli were similar to Experiment 1, but only one sentence, a simple transitive, was recorded per video (e.g., "She ate a cookie"). We also made a slight variation from Experiment 1 for the *pluck* predicate: We chose the English verb *break* because it is more colloquial than *pluck*, but to ensure that participants interpreted the events correctly, as a removal of the object from its stem, rather than damage to the object itself, we added the particle "off" (e.g., "She broke off the banana"). For the other sentences, see the English translation column of Table 2.

3.1.3 Design and Procedure

The procedure was identical to Experiment 1, except that there was only one syntactic condition (i.e. participants heard the same sentence type throughout). Completion condition varied within-subject, as in Experiment 1.

3.2 Predictions

If English simple perfective forms are akin to the Hindi CV in meaning, then they should pattern like them. We expect acceptance rates similar to those for the Hindi CV (that is, near 100% acceptance for fully-completed events, and close to 33% acceptance for partially-completed events).

With respect to predictions for different predicate types, because English does not have the neutral perfective, Singh's account does not make any specific predictions for predicates with and without incremental themes. If, however, the pattern of acceptance resembles that found for Hindi, it may suggest that speakers of both languages are sensitive to real-world and contextual factors. That is, if English speakers ever do allow partial completion interpretations, they should do so for the very same predicates/events for which Hindi speakers do. After all, real-world and discourse contexts are held constant across experiments.

3.3 Results and Discussion

Responses were coded as in Experiment 1. For fully-completed events, the sentence descriptions were accepted 97.3% of the time, while for partially-completed events, they were accepted 46.9% of the time. A multi-level model with Event Completion (2: partial vs. full) as fixed factor revealed that Event Completion is a reliable predictor (beta coefficients in Table 6) for English as well.

TABLE 6 Experiment 2a (English): Estimates of fixed effects from best-fitting multi-level model of proportion of "true" responses (empirical logit transformed)

Effect	Estimate	S.E.	*t*-value
Intercept	2.15	0.16	13.50*
Event Completion	-3.23	0.28	-11.59*

$*\ p < 0.05$ (on normal distribution)

A glance at the results for individual predicates reveals a similar pattern as that found for the Hindi data in Experiment 1 (Table 7).[6] While in the fully-completed condition, all predicates had a near 100% acceptance, in the partially-completed condition there was a high degree of variability; acceptance rates ranged from 0% (for *break off*) to 95% (for *fill*). This range of acceptance is surprising given that change-of-state predicates in English are typically thought to entail completion of the events they describe. We interpret this as evidence that factors such as contextual information and real-world knowledge play an important role in speakers' acceptance of change-of-state predicates for partially-completed events. If extralinguistic factors are relevant, then given that Hindi and English speakers were presented with identical visual events, we expect patterns of acceptance to be

[6] We did not analyze incremental theme vs. non-incremental theme predicates for this experiment because there was only one syntax condition.

similar across languages. That is, the same event-predicate pairs that yielded high acceptance rates from Hindi speakers should yield high acceptance rates from English speakers. This prediction was largely borne out. Considering only partially-completed events, notice that *break off* had the lowest acceptance rate in both Hindi and English, with *fill* and *eat* having among the highest acceptance rates.

TABLE 7 Experiment 2a (English): Percent "true" responses, by condition and trial

Predicate	Partial Completion	Full Completion
Incremental Theme		
cover	54	96
draw	64	87
eat	67	96
fill	95	100
Non Incremental Theme		
close	13	100
extinguish		
Newspaper (partial)	8	100
Candle (return to initial state)	17	100
break off	0	100
wake (return to initial state)	75	100

Interestingly, acceptance rates were similar to those for Hindi SV sentences, rather than CV sentences. For both the English sentences and Hindi SVs, several predicates received over 50% acceptance rates. For Hindi CV sentences, only *fill* did. In this experimental task, at least, English-speaking participants were quite willing to accept these change-of-state predicates as descriptions of partially-completed events. These results cast doubt on a theory in which English perfective verbs entail event completion, while Hindi SVs do not.

However, there is an important caveat in interpreting these results. While in Experiment 1 participants were exposed to a within-subjects design, hearing both SV and CV sentences, in the current experiment participants only heard one kind of sentence: an English simple perfective. Would English speakers have responded differently if they had had the opportunity to contrast simple perfectives with another construction, as Hindi speakers did in Experiment 1? If the English perfective entails completion, then contrasting it with another construction should make no difference, but if contextual factors are relevant in determining whether a partial completion interpretation is available, then a contrast may indeed skew participants' judgments.

To explore this possibility, in Experiment 2b we devised a condition somewhat comparable to the Hindi CV: a particle construction using *up* (e.g., *fill up*). Although in English simple perfectives like *fill* are generally understood to entail completion, we suspected that the particle construction would draw even more emphasis, and therefore attention, to event completion. If English speakers notice the contrast between these completion-focused constructions and the simple perfectives, they may be even more willing to accept partially-completed events for the simple perfective, reserving their "false" judgments for the completion-focused particle construction.

4 Experiment 2b: English with *up*

Although English does not have an SV/CV distinction, particles like *up* can emphasize completion and result states. *Fill up*, for example, seems to emphasize the endstate, *fullness*, more than *fill*. In this experiment, we compared interpretations of verb-particle constructions with the simple verbs from Experiment 2. Of course, the difference between *fill* and *fill up* is not directly parallel to the difference between a Hindi SV and CV. But nevertheless, our goal in Experiment 2b was to determine if a similar pattern obtained between these two constructions as in the pattern of responses to Hindi

SV and CV sentences from Experiment 1.

We selected a subset of the predicates from the previous experiments that can take the particle *up*, yielding a construction that emphasized the result state without significantly changing the event the verb referred to: *cover (up), eat (up), fill (up)*, and *wake (up)*. Although to us, *cover up* typically means fully obscuring the covered object, rather than full placement of a lid on a container as in our videos, we included this predicate here because pilot work indicated that at least some speakers would use *cover up* to label the full completion event in our video.

4.1 Methods

We followed the methods used in Experiments 1 & 2a.

4.1.1 Participants

A different group of twelve native English-speaking Northwestern University undergraduates participated for course credit.

4.1.2 Materials

Visual stimuli were identical to Experiments 1 & 2a. There were two auditory conditions. Auditory stimuli for the simple verb condition were identical to Experiment 2a for the four predicates *cover, eat, fill*, and *wake*. The second condition incorporated the particle *up* (e.g., *She ate up the cookie.*).

4.1.3 Design and Procedure

Identical to Experiment 1, except that participants viewed just eight videos, two for each of the four predicates.

4.2 Predictions

If English simple transitives entail full completion, like Hindi CVs, adding the particle *up* may *emphasize* completion, but not significantly alter speakers' interpretations. They should be no more likely to accept simple transitives as descriptions of partially-completed events than they are of sentences with *up*.

In contrast, if *up* can play a similar role to the light verb in the Hindi CV construction, we expect a similar pattern of results as obtained for the SV/CV distinction in Experiment 1. That is, we expect sentences with *up* to have a lower acceptance rate for partially-completed events than simple transitive sentences without *up*. Further, if English speakers contrast the particle and simple verb sentences, they may be more inclined to use the simple verb sentences to take over the function of describing partially-completed events. This will result in even higher acceptance rate for the simple verb sentences than these same sentences yielded in Experiment 2a.

4.3 Results and Discussion

We entered the mean proportion of "true" responses for each participant into a multi-level model with Event Completion and Syntax as fixed factors (beta coefficients in Table 8).The analysis reveals that both Event Completion (2: partial vs. full) and Syntax (2: SV vs. CV, that is no particle vs. particle) are reliable predictors, while their interaction is marginally reliable ($p = 0.054$). However, these effects are likely mostly driven by differences between syntactic conditions in the partially-completed condition for *eat*. See Table 9.

Strikingly, the acceptance rate for the simple transitives in Experiment 2b was higher than it had been for those predicates in Experiment 2a (collapsing across those predicates appearing in both Experiments 2a and 2b, $t(70) = 2.00$, $p < .05$). This is suggestive evidence that contrasting the simple transitives with the *up* sentences caused speakers to accept partial-completion interpretations for simple transitives with even greater frequency. By drawing English speakers' attention to the availability in English of a construction that highlights completion, we were able to increase acceptance rates for partially-completed events for simple transitives. Further data will be necessary to lend clear support to this hypothesis.

TABLE 8 Experiment 2b (English): Estimates of fixed effects from best-fitting multi-level model of proportion of "true" responses (empirical logit transformed)

Effect	Estimate	S.E.	*t*-value
Intercept	1.37	0.10	13.96*
Syntax	0.51	0.20	2.60*
Event Completion	-0.54	0.20	-2.74*
Syntax x Event Completion interaction	0.76	0.39	1.93

$*$ $p < 0.05$ (on normal distribution)

TABLE 9 Experiment 2b (English): Percent "true" responses, by condition and trial

Predicate	Partial Completion		Full Completion	
	Simple Verb	Particle (*up*)	Simple Verb	Particle (*up*)
cover	83	67	83	67
eat	83	17	100	100
fill	100	83	100	100
wake (return to initial state)	100	83	100	100

5 General Discussion

We have presented the results of two experimental studies exploring Hindi and English speakers' interpretations of perfective verbs describing change-of-state events. We considered two types of events: events arriving at full completion, with the change of state fully achieved, and events for which the change of state was not fully achieved. Our goal was to discover whether speakers accepted perfective verbs as descriptions of these latter partially-completed events, and whether acceptance varied across syntactic construction, semantic class of predicate, or across language.

To address these issues, we contrasted two perfective forms in each language. In Hindi, we contrasted the SV and the CV. SV sentences are thought to be compatible with both fully- and partially-completed events, and CV sentences only with fully-completed events. Because English does not have a CV construction, in Experiment 2a we elicited speakers' judgments of transitive sentences containing simple perfective verbs. Although similar in form to the Hindi SV, they are thought to describe fully-completed events, like the Hindi CV. In Experiment 2b, we contrasted judgments of these simple perfective verbs with constructions involving the particle *up* (e.g., *fill up*) in order to introduce a contrast similar to Hindi's SV/CV distinction.

The experiments tested three main predictions. First, we predicted that if Hindi SV sentences can describe partially-completed events, Hindi speakers would accept SV sentences more often than CV sentences as descriptions of these events. Both SV and CV sentences were predicted to be acceptable for fully-completed events. We tested this prediction in Experiment 1. Hindi speakers, as expected, showed different patterns for SV and CV constructions, accepting SV sentences significantly more often for partially-completed events than CV sentences. These results support previous experimental work (Pederson 2007, Wittek 2002), but extend the findings with an experimental design which fully crosses event completion and syntactic construction within a single language.

Surprisingly, despite being accepted significantly more often than CV sentences, SV sentences were accepted just half the time, a much lower rate than expected if SVs do not entail event completion. But these results are consistent with the observation that the default interpretation for SVs is still full completion (Kothari 2008), and partial completion interpretations arise only when favored by context. We suggest that because full completion interpretations entail partial completion interpretations, the full completion interpretation is stronger, and therefore speakers may prefer it unless context strongly drives a partial completion interpretation (as in the case of explicit cancellation) (e.g., Kearns 2007).

The acceptance rate for CVs, too, was surprising; it was much higher than the near 0% acceptance we expected. We suspect that some of our partially-completed events were construed as *functionally* complete, despite not being 100% complete, and that this rendered them acceptable for both SV and CV sentences. This is most transparently true for *fill*, for which our partially-completed event depicted a glass that was approximately three-quarters full. We return to this point below.

Our second prediction stems from Singh's (1990) account of the Hindi SV/CV distinction, which predicts that partial completion interpretations for SVs arise only for accomplishment predicates with incremental themes. The findings here were surprising as well. While accomplishment predicates with incremental themes did show a difference between SV and CV constructions, so did other predicates, with a comparable effect size. Our experimental results therefore offer partial support for Singh's account, but are inconsistent with the prediction that only accomplishment predicates featuring incremental theme objects should allow partial-completion interpretations for SV sentences. The findings are particularly surprising given that the non-accomplishment predicates we tested were primarily achievements, which one would expect to be incompatible with partial completion readings of any kind. We hope that extending the experimental paradigm to a larger set of accomplishment and achievement predicates will shed light on whether there are any semantic features shared by predicates that show the SV/CV distinction, beyond Vendlerian classes.

It is worth commenting on the findings for *wake*, because these lend new insight into partially-completed event construals. Recall that the partially-completed videos for *wake* depicted a full completion of the event followed by a return to the initial state. In Hindi, these predicates showed a difference between SV and CV sentences for partially-completed events. Not only that, but the acceptance rates for this achievement predicate, in both constructions, were close to what we had initially predicted for accomplishment predicates with incremental themes: CV sentences received a very low acceptance rate (17%), and SV sentences a very high acceptance rate (83%), making it a paradigm example of the SV/CV distinction.

These results suggest that for SVs, at the time of evaluation (here, the end of the video), the endstate needn't be apparently achieved at all, as long as some change took place earlier in the event. (We suspect if the sleeping actor had not woken at all, speakers would have been much less likely to accept the SV sentence.) SV sentences can therefore be used in situations in which an action has not been completed to the satisfaction of the actor's perceived intention. But for CVs, the intended endstate must be evident at the time of evaluation. Even if the endstate was completely brought about at some earlier point during the video, speakers did not construe the CV sentence as felicitous. In judging the CV, the speaker must judge whether the actor fulfilled those intentions to an appropriate degree.[7] We do not suggest that the meaning distinction between SV and CV is reduced to one between "try to" and "succeed at", because SV sentences can describe partially-completed events that were brought about accidentally (Pederson 2007). However, perceived intention is at least one possible route to felicity of an SV sentence as a description of a partially-completed event. Although the current study only presented this event type for two predicates, we believe the findings are suggestive and warrant more detailed investigation.

Semantic class alone, then, is not adequate to explain Hindi speakers' interpretations of SV and CV sentences. While the current study does not test the predictions of a specific alternate account, we believe the results suggest promising avenues for future research. One factor that we believe to play an important role in speakers' interpretations is real-world context. Whether an event is construed to be complete is related to a contextual standard, which can differ both by the type of object (e.g., the level of a full wine glass typically does not reach as high as the level of a full water glass) and the context of use (e.g., filling a water glass to drink typically does not reach the rim of the glass, while filling a water glass in order to measure out a quantity of water may), with the relevant scale differing across verbs. (This is of course true for English as well; consider well-known examples like *John read* War and Peace *last night.* vs. *John read* Dick and Jane *last night.*) Context,

[7]The Russian perfective/imperfective distinction yields parallel interpretations here. Thanks to an anonymous reviewer for this insight.

in combination with other factors, is likely to be an important determinant of whether a predicate yields an SV/CV distinction.

An alternate hypothesis, which we also leave to future research, but which is consistent with the results of the present study, is Copley and Harley's (2010) suggestion that a force-dynamic treatment of events can account for the availability of partial-completion interpretations in languages like Hindi. They propose that a force, or input of energy, comprises the semantic weight of predicates like the ones we studied here, rather than a change of state. The CV carries a presupposition that the force was effective in bringing about the result state, while the SV does not. On this account, the actor's intention need not be relevant; all that is required is that a force applies toward the result. This kind of account accommodates contextual factors as well. The likelihood of efficacy of the net force applied in any given situation should depend on a number of real-world factors, including the scale in question and whether other factors (forces) are intervening.

Our third prediction relates to expected differences between Hindi and English. because English perfectives are traditionally assumed to entail completion, much like Hindi CV sentences, speakers' interpretations of the English sentences in Experiment 2a should pattern like the Hindi CVs in Experiment 1. But in fact, English sentences patterned more like Hindi SV perfectives than CV perfectives, with a 47% acceptance rate. These patterns do not, of course, indicate that English perfectives and Hindi SVs are identical with respect to the behavior of these predicates. It is still the case that for Hindi SVs, the cancellation of the endstate, as in (1), yields a much more felicitous utterance than its English translation. Why then did the English sentences yield such a high acceptance rate?

Further investigation is needed to understand this finding, but we speculate that it is related to our striking finding that in Experiment 2b, when speakers heard both simple transitives and *up* sentences, they gave an *even higher* acceptance rate for the simple transitives than they had for these very same sentences in Experiment 2a. This suggests that participants in Experiment 2b were comparing the sentences with and without *up* to determine which they felt was a better description of each event.[8] Once speakers' attention was drawn to the fact that there are constructions that emphasize event completion, the simple verb took on the function of describing events that have reached only partial completion. Future research will have to explore whether other constructions, too, yield this kind of contrastive effect, and how this interacts with other factors such as perceived intention.

Finally, we note that both Hindi and English showed similar patterns across the different predicate types. A glance at the acceptance rates for each predicate (Tables 4 and 7) supports this prediction, when the verbs are compared to Hindi SV sentences. For Hindi SVs, *pluck, close, draw*, and *extinguish* all had acceptance rates under 50%; for English, *pluck/break off, close*, and *extinguish* fell under this mark. In both languages, *cover, fill, wake*, and *eat* all fell above the 50% mark. Hindi CVs do show a similar pattern, with *pluck, wake*, and *extinguish* showing the lowest acceptance rates, but for CVs all eight predicates fell under the 50% mark. These striking similarities between Hindi and English suggest that Hindi and English simple verbs do not, at base, have distinctively different semantic properties; speakers of both languages allowed partial-completion interpretations for simple verbs approximately 50% of the time, and showed similar responses to different predicates. Our experimental results suggest that the differences between English and Hindi are more complex and subtle than expected if a stark aspectual difference exists between the two languages: English speakers showed an almost identical pattern to Hindi speakers, albeit somewhat attenuated. Again, contextual factors may explain this concordance across languages; after all, the events we showed and the experimental context were held constant across language groups.

These findings for English are surprising on a purely semantic account, which would have predicted that English perfectives without *up* would yield interpretations more like the Hindi CV, not the SV, and that the addition of *up* would not have had effects with respect to event completion construals

[8] They were likely comparing the constructions implicitly. Fewer than half the participants indicated that they had noticed the contrast between the sentence types when questioned during debriefing.

(although *up* may nevertheless have had emphatic or other effects). Taken together with the results from Hindi alone from Experiment 1, the results for English suggest that Singh's semantic account alone is inadequate. Although we do not suggest that there are no differences the languages, we find the similarities compelling.

Hindi as well as English perfectives do, then, describe situations with clear starting and ending points. But contrary to traditional analyses (e.g., Smith 1997), perfectives are compatible with assertions that the situation terminated before the event's natural completion point. That is, the closing bound of the event does not coincide with the bound determined by the telic point. The findings for *wake* suggest, too, that in at least some cases the event may reach that closing bound, but revert completely to the starting point, and still be described with an SV perfective.

Limitations of the current study This exploratory study is, to our knowledge, the first to use experimental data to investigate the factors underlying the interpretation of different kinds of perfective forms both within and across languages. Our findings shed light on these factors; we hope that future work will take these findings as a baseline and investigate these factors in more depth.

While the present results clearly demonstrate that incremental theme verbs are not the only ones that show a distinction between the SV and CV constructions, they do not clearly point in favor of a specific alternative hypothesis. There are likely many interacting factors to tease apart, both semantic and pragmatic. Future work must explore this phenomenon with a greater number of verbs, with systematic manipulation of event type and theme type, and a wider range of semantic categories. However, we do hope that the results shed some light on promising future directions; for example, our results suggest that the actor's perceived intentions can at least in some cases determine felicity of an SV sentence for a partially-completed event, and that real-world context needs to be both carefully manipulated, to understand its role, as well as carefully controlled, to understand the role of other factors when context is held constant.

The fact that context plays such an important role in interpreting sentences like these is both a limitation and a strength of the current study. It is a limitation because by playing videos, we gave speakers a context, and a specific event to imagine. As a result, we cannot be sure whether speakers would have provided different answers if the event had reached slightly closer to or farther from completion. For example, would speakers have been even more likely to accept Hindi SV sentences for *eat* if slightly more of the cookie had been eaten in the depicted event? Future work should probe these boundaries more closely.

But our provision of video events is also a strength of this study. Providing a specific event ensures that all speakers pictured the same events, and were evaluating the sentences with respect to those particular events that we provided. For example, we were able to elicit judgments on multiple contexts, with all other aspects of the situation held constant, such that each participant was given the same mental image of the event being described. We were also able, with the use of videos, to easily bring to mind unusual contexts, such as events involving a return to the initial state (e.g., blowing out a trick candle which re-lights). In general, the SV/CV distinction is a subtle one; because the default interpretation for a Hindi or English perfective is full event completion, even for SV sentences, it is difficult to encourage participants to consider partially-completed events. Providing them with a visual context of what such a partially-completed event looks like helps immeasurably. These results suggest that experimental studies can provide a unique kind of data, and should be pursued in tandem with other methods. We therefore hope that in addition to presenting important empirical data about the interpretation of Hindi and English perfectives, the current study also calls attention to the value of experimental studies.

Acknowledgments

We are grateful to Eve Clark, Bridget Copley, Alice Davison, Bhuvana Narasimhan, and two anonymous reviewers for their insights; all errors are of course our own. A subset of these results was presented at SALA 28, at the Interdisciplinary Workshop on Verbs: The Identification and Repre-

sentation of Verb Features, and at the South Asian Syntax and Semantics Workshop at UMass, to whose audiences we are grateful for comments. Collection of Hindi data was supported by Stanford University's Center for South Asia and Graduate Research Opportunities Award to Anubha Kothari. Collection of English data was supported by NIH R01 HD30410 (PI Sandra Waxman, Northwestern University).

References

Bar-El, Leora, Henry Davis, and Lisa Matthewson. 2005. On non-culminating accomplishments. In *Proceedings of the 35th Annual Meeting of the North East Linguistics Society*, vol. 1, pages 87–102.

Butt, Miriam and Gillian Ramchand. 2005. Complex aspectual structure in Hindi/Urdu. In N. Erteschik-Shir and T. Rapoport, eds., *The Syntax of Aspect*, pages 117–153. Oxford University Press.

Copley, Bridget and Heidi Harley. 2010. An ontology for a force-dynamic treatment of events. In B. Copley and F. Martin, eds., *Forces in Grammatical Structures*. To Appear.

Dell, François. 1987. An aspectual distinction in Tagalog. *Oceanic Linguistics* 22-23:175–207.

Hook, Peter Edwin. 1976. The Hindi compound verb: What it is and what it does. In K. Singh, ed., *Readings in Hindi-Urdu Linguistics*, pages 129–146. Delhi: National Publishing House.

Ikegami, Y. 1985. 'Activity'-'Accomplishment'-'Achievement'-A language that can't say 'I burned it but it did not burn' and one that can. In A. Makkai and A. K. Melby, eds., *Linguistics and Philosophy: Essays in Honor of Rulon S. Wells*, pages 265–304. Amsterdam: John Benjamins.

Kearns, Kate. 2007. Telic senses of deadjectival verbs. *Lingua* 117:26–66.

Koenig, Jean-Pierre and Nuttanart Muansuwan. 2000. How to end without ever finishing: Thai semi-perfective markers. *Journal of Semantics* 17:147–194.

Kothari, Anubha. 2008. Event culmination as implicature in Hindi perfectives. Ms., Stanford University.

Pederson, Eric. 2007. Event realization in Tamil. In P. Brown and M. Bowerman, eds., *Cross-linguistic Perspectives on Argument Structure: Implications for Learnability*, pages 331–355. New York: Lawrence Erlbaum.

Raja, Nasim A. 2003. Aspectual complex predicates in Punjabi. In R. Singh, ed., *The Yearbook of South Asian Languages and Linguistics*, pages 99–129. Berlin: Walter de Gruyter.

Singh, Mona. 1991. The Perfective Paradox: Or, how to eat your cake and have it too. In *Proceedings of the 17th Annual Meeting of the Berkeley Linguistics Society*, pages 469–479.

Singh, Mona. 1998. On the semantics of the perfective aspect. *Natural Language Semantics* 6:171–199.

Smith, Carlota. 1997. *The Parameter of Aspect*. Dordrecht: Kluwer Academic Publishers.

Teng, Shouxin. 1972. *A Semantic Study of the Transitivity Relations*. Ph.D. thesis, University of California, Berkeley.

Travis, Lisa. 2000. Event structure in syntax. In C. Tenny and J. Pustejovsky, eds., *Events as Grammatical Objects: The Converging Perspectives of Lexical Semantics and Syntax*, pages 145–185. Stanford: CSLI Publications.

Verma, Mahindra K., ed. 1993. *Complex Predicates in South Asian Languages*. New Delhi: Manohar.

Wittek, Angelika. 2002. *Learning the Meaning of Change-of-State Verbs: A Case Study of German Child Language*. Berlin: Mouton de Gruyter.

Mirativity in Kurtöp

Gwendolyn Hyslop, *Research Centre for Linguistic Typology, La Trobe University*

Received May 2010; Revised June 2011

Abstract

The linguistic encoding of information as being unexpected is referred to as the 'mirative'. Since DeLancey (1997) established mirativity as a cross-linguistic category, the awareness of its presence in languages of the world is growing. This is true also for the languages of South Asia, perhaps most commonly amongst the Tibeto-Burman languages. This article offers the first description of mirativity in Kurtöp, an under-described Tibeto-Burman language of Bhutan. Mirativity is encoded with unique suffixes and free forms that are also an integral facet of the verbal system in Kurtöp. Mirativity is contrasted in perfective and imperfective aspect verbal inflections, and in distinct affirmative and negative sets of equational and existential copulas.

1 Introduction

As advances in documentary linguistics provide us with in-depth descriptions of a growing number of languages, our understanding of seemingly unusual phenomena is deepened, and indeed what seemed unusual several decades in the past is now known to be more commonplace. Mirativity, described for only a few languages two decades ago, is now known to occur in dozens of different languages (if not more) from different language families, around the world. For example, since DeLancey's seminal article (DeLancey 1997), mirativity has been described for several languages in the Himalayan region alone (e.g., Grunow-Hårsta (2007) for Magar; DeLancey (1997) for Sunwar and Newar; Watters (2002) for Kham; Bashir (2010) for Shina[1]). In fact, Dickinson (2000, 380) speculates that mirativity may be a universal conceptual category. The aim of this article is to offer a description of mirativity in Kurtöp, which is grammatically encoded throughout the verbal system.

The mirative as a conceptual category is different from, but related to, evidentiality and epistemic modality, and is perhaps best understood in light of these two. Evidentiality is concerned with *source* of knowledge; epistemic modality encodes *certainty* of knowledge, while mirativity is concerned with *expectations* of knowledge. Dickinson (2000, 381) asserts that a 'mirative marker indicates psychological distancing — the speaker did not anticipate the event or state', which she contrasts with an inferential evidential marker, which 'indicates physical distancing from the event'. DeLancey (1997) defines mirativity as 'the status of the proposition with respect to the speaker's overall knowledge structure'.

Mirativity in some languages is encoded by a morpheme or construction that is also used to encode other categories, such as inference or related evidential categories. In other languages, a particular

[1]This article does more than provide an analysis of mirativity in Shina (which, in Shina, is 'one semantic development from an underspecified verbal form which can develop various meanings' (Bashir 2010, 47)); it also provides an impressive summary of mirative marking in fourteen neighboring languages of various families.

JSAL Volume 4, Issue 1, December 2011.

form or construction is devoted exclusively to mark mirativity. Perhaps the most well-known example is Turkish (Aksu-Koç and Slobin 1986, Slobin and Aksu 1982). Below, I follow Dickinson (2000, 380–381) in presenting the data and analysis from Aksu-Koç and Slobin (1986) and Slobin and Aksu (1982).[2]

(1) Kemal geld-**di**
Kemal come-Past
'Kemal came.'

(2) Kemal geld-**mış**
Kemal come-mış
'Kemal came.'

Slobin and Aksu (1982) describe three contexts for the data in (2):

a. *inference*: The speaker sees Kemal's coat hanging in the front hall, but has not yet seen Kemal.
b. *hearsay*: The speaker has been told that Kemal has arrived, but has not yet seen Kemal.
c. *surprise*: The speaker hears someone approach, opens the door, and sees Kemal — a totally unexpected visitor.

Importantly, the context described in (c) is considered a basic function of *-mış*, not an idiosyncratic or unusual use. Psychologically, the function of *-mış* is described as representing a situation in which the speaker had no "premonitory awareness" (Slobin and Aksu 1982, 196). As we will see below, this is also the analysis of the Kurtöp imperfective mirative.

Kurtöp displays many features typical of South Asian languages including SOV syntax, Differential Object Marking (Hyslop 2010), and inflectional verbal morphology. Finite clauses in Kurtöp, with the exception of imperative moods and questions, are obligatorily encoded for evidential-like values, including evidentiality, speaker's expectation of knowledge, and epistemic modality. Mirativity, as one of these categories, is encoded in perfective and imperfective aspect as well as copular clauses.

The data presented in this article come from extended fieldwork in Bhutan, mainly during 2008–2009, but also shorter trips in 2006, 2007 and 2010. My methodology has been to collect natural data by way of recorded conversations, storytelling and personal narratives. Elicitation is mainly used to fill in paradigmatic gaps in the data and cross-check for negative examples. Unless stated otherwise, the data in this article were drawn from a corpus consisting of eleven texts, representing the speech of 14 native speakers, both males and females varying in age from 20 to over 60.

2 Background

Kurtöp is a Tibeto-Burman language of the Lhüntse district of Northeastern Bhutan. The speech community begins just south of Lhüntse town, in Tangmachu, and runs north, along the Kurichu, until the Tibetan border. On the east, Kurtöp is bordered by the Dzala speaking area, to the south by Chöcangaca, and to the west by Bumthap. The location of the Kurtöp speech community is shown in Figure 1. Kurtöp has several distinct dialects and is spoken by approximately 15,000 people. The variety of Kurtöp represented in this study is that of Dungkar geok (district), where there are approximately 3,000 speakers.

Within the Tibeto-Burman language family, Kurtöp has been classified as 'East Bodish', a term first used by Shafer (1954) to identify a family of languages that was closely related to but not directly descended from Classical Tibetan. Since then, Aris (1979) and van Driem (1998, inter alia) have identified several other East Bodish languages, mainly in Bhutan. These include Dzala, Chali, Bumthap, Khengkha, 'Nyenkha, and Kurtöp. The historical placement of the East Bodish

[2] In neither Aksu-Koç and Slobin (1986) nor Slobin and Aksu (1982) are the data in (1) and (2) presented in this order. Aksu-Koç and Slobin (1986) present sentence (2) but with the subject Ahmet rather than Kemal while Slobin and Aksu (1982) present the data in (1). However, it is clear that the presentation and analysis presented by Dickinson (2000) is accurate and a concise summary of the issue.

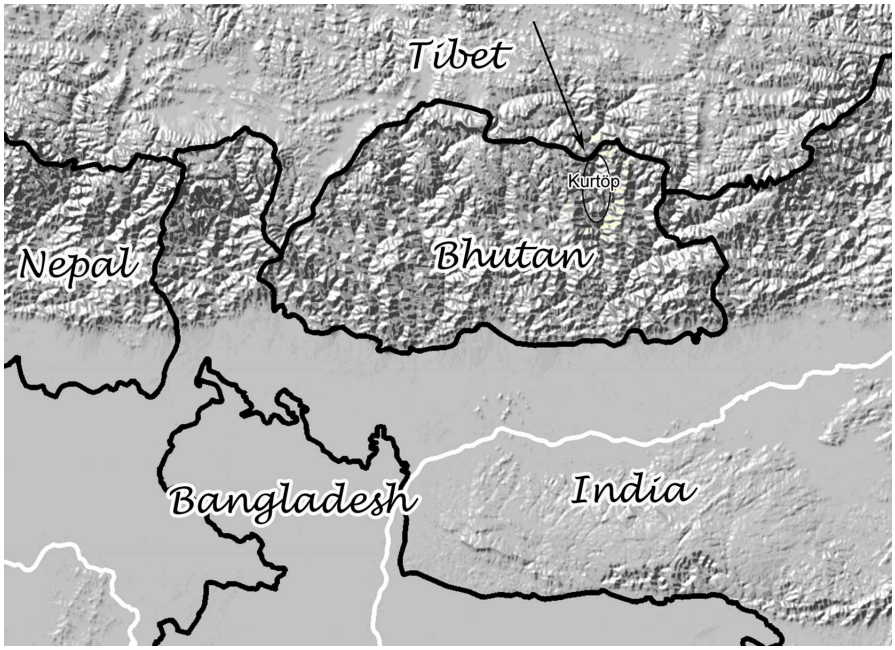


FIGURE 1 The Kurtöp language area in Bhutan

languages within Tibeto-Burman is still subject to debate, particularly in light of the fact that most of these languages are still virtually undescribed. Since Shafer, many scholars (e.g., Michailovsky and Mazaudon 1994, DeLancey 2008, Hyslop 2008) have noted the many features that the East Bodish languages share with Tibetan. However, it is possible that many of these features are borrowings over a potentially non-Bodic substrate.

2.1 Evidentiality, mirativity and speaker expectation

Evidentiality as a linguistic category encoding source of knowledge is now established as a cross-linguistically relevant category. Aikhenvald (2004) identifies numerous linguistic areas and language families well known for evidential systems, including the Balkans, areas of North America, Mexico, Amazonia, the Andes, a few languages in Africa, and Turkic, West Caucasian, Eskimo-Aleut and Tibeto-Burman language families. Recently, an entire volume of *Linguistics of the Tibeto-Burman Area* was devoted to descriptions of evidentiality in Tibeto-Burman languages, with contributions on Rgyalthang Tibetan (Hongladarom 2007), Yongning Na (Lidz 2007), Darma (Willis 2007), nDrapa (Shirai 2007), Magar (Grunow-Hårsta 2007), and Spiti Tibetan (Hein 2007).

The existence of mirativity as something different from evidentiality has long been noted (e.g., Aronson 1967, Friedman 1977, 1986) but DeLancey (1997) is credited with the establishment of mirativity as a cross-linguistic, typological category. While evidentiality is linguistic coding of source of information, mirativity is quite different, coding that the information is not expected. Despite the acknowledged difference between these categories, evidentiality and mirativity are often intertwined in the same verbal paradigms. Indeed, this is the case in Kurtöp.

2.2 The Kurtöp verbal system

Kurtöp, like almost all other Tibeto-Burman languages and the languages of South Asia, has verb-final syntax. Core arguments generally precede the verb and in the case of bivalent verbs, the A argument will precede the O argument. However, this AOV order is a generalization; in natural speech speakers may move the S, A and/or O argument to follow the verb, depending on pragmatic factors. Verbal arguments are not required overtly and in fact, are often missing in natural discourse.

Kurtöp tends toward a polysynthetic morphological profile, with many words consisting of more than one morpheme. Verbs are usually composed of two to three morphemes within three to four syllables and it is not unusual for verbs to consist of five syllables. There is one prefix in the language (the negative marker, discussed below) and the remainder of verbal morphology is comprised of suffixes and enclitics.

Leaving the matter of questions and imperatives aside, a distinction can be made in Kurtöp between clauses that end with a copula and clauses that do not end in a copula. In the former category, a copula may be used to encode typical copular functions (existence, equation, prediction, location, possession), may be used in conjunction with a clausal nominalization, or may be used with a non-final marked converb in a clause-chaining construction. Clauses that do not end in a copula will consist of minimally a verb plus a finite suffix. There are three enclitics that speakers can use at the end of the clause, either attached to the copula in the case of a copular clause, or attached to the end of the finite-marked verb, in the case of a copula-less clause.

2.2.1 Clauses involving copulas

In addition to the typical functions of copulas (cf. section 5), copulas are widely used in Kurtöp as part of the verbal system. More specifically, equational copulas can be used in conjunction with a nominalized clause, and an existential copula can be used as the final, finite verb in a clause chain.

The data in (3)[3] and (4) illustrate the basic distinction between a clause without a copula and a formally nominalized clause. In (3) the verb *ge* 'go' is suffixed with the finite egophoric perfective *-shang*. The data in (4) show the verb *ge* 'go' nominalized with the perfective nominalizer *-wala* and followed by the mirative form of the affirmative equative copula.

(3) khit ge-**shang**
3.Abs go-Pfv.Ego
'He went.'

(4) khit ge-**wala** **wenta**
3.Abs go-Nmz.Pfv Cop.Eq.Mir
'He went indeed.'

In addition to *-wala* (with allomorphs *-sala* and *-pala*), Kurtöp has an imperfective nominalizer *-khan*, a future nominalizer *-sang*, and an irrealis nominalizer *-male*. A clause nominalized by any of these can be completed with any form of the equative copula, depending on the polarity and the epistemic, evidential, or mirative value of the utterance.

Kurtöp also makes wide use of a clause-chaining construction, which, when used with a final existential copula, encodes durative aspect. A proto-typical use of the clause-chaining construction is shown in (5).[4]

(5) tsheni igu-the **co-si** boi **bi-shang**
then letter-Def make-Nf 3.Erg give-Pfv.Ego
'Then after making the letter, they gave (it).' SBC200511275.83.02-86.595KW

Here, the first clause consists of the O argument *igu-the* 'letter-Def' and the converb *co* 'make', marked with the non-final suffix. The second clause consists of the ergative-marked pronoun *boi* and the final verb *bi* 'give', which is suffixed with the egophoric perfective suffix *-shang*. Note here

[3]Data in this paper are represented in Roman letters designed for Dzongkha by George van Driem and Karma Tshering and adapted to Kurtöp. The symbols correspond to the IPA as follows: <k> [k], <kh> [kʰ], <g> [g], <ng> [ŋ], <c> [c], <ch> [cʰ], <j> [ɟ], <ny> [ɲ], <tr> [ʈ], <thr> [ʈʰ], <dr> [ɖ], <t> [t̪], <th> [t̪ʰ], <d> [d̪], <p> [p], <ph> [pʰ], <b> [b], <m> [m], <ts> [ts], <tsh> [tsʰ], <sh> [ɕ], <zh> [ʑ], <s> [s], <z> [z], <l> [l], <lh> [l̪], <r> [r], <a> [ɑ], <e> [e], <i> [i], <o> [o], <u> [u], <ö> [ø], <ü> [y], <'CV> high tone on following vowel, <ˆ> long vowel.

[4]If data are drawn from the textual database, the file name of the text, location in the text (if noted) and speaker is indicated. For example, in 'SBC200511275.83.02-86.595KW', 'SBC200511275' is the name of the recording; '83.02-86.595' indicates this utterance occurred between 83.02-86.595 seconds (time stamp given by the program Transcriber), and 'KW' indicates the speaker was KW.

that verbal arguments are shared between the two clauses, but that is not a requirement of the construction; it is also possible for each verb to have its own argument (overt or covert). This example shows one converb, but in natural speech it is quite common for several converbs to be chained together, followed by one final, finite-marked verb.

What is relevant for our purposes is the presence of an existential copula as the final, finite verb. Consider (6), showing the converb *thung* 'do' immediately followed by the existential copula. The resulting denotation is a single event, encoding durative aspect, which differs from the imperfective *-ta/-taki* in that it indicates the event continued or continues on for an extended period of time.

(6) khit chorten kora **thung-si nawala**
3.Abs stupa circumambulation do-Nf Cop.Exis
'S/he keeps circumambulating the stupa.'

Example (6) shows the clause-chain construction ending with the basic form of the affirmative existential copula, but speakers may use any of the existential copulas, depending on the particular epistemic, evidential, or mirative value they denote.

The semantically unmarked forms of the copulas are shown in Table 1. The equational copulas are *wen* and *min* while the existential copulas are *nawala* and *mû*. Syntactically, copulas differ from lexical verbs in that they do not take verbal prefixes or suffixes (however, as I show below, some of the copulas appear to have been formed diachronically with some of the verb suffixes described in sections 3 and 4). Copulas, however, can be suffixed with a subset of nominalizers (Hyslop 2011). The various epistemic, evidential, and mirative forms of the copulas are presented in section 5, where I focus the discussion on the mirative forms.

Copula	Semantic Value	Polarity
wen	Equational	Positive
min	Equational	Negative
nawala	Existential	Positive
mû	Existential	Negative

TABLE 1 Kurtöp semantically unmarked copulas

2.2.2 Clauses without copulas

In declarative matrix clauses that are not nominalized, a verb can be suffixed with one inflectional suffix as well as a negative prefix. Table 2 provides an overview of the syntagmatic analysis of the Kurtöp finite verb. The suffixes in bold, *-na* and *-ta*, are miratives in perfective and imperfective aspect, respectively, and will be discussed in greater detail below.

Negation[5]	Stem	Suffix
ma-, me-, mi- *má-, mé-, mí-*	———	*-shang, -pala, -para,* ***-na****, -mu,* ***-ta****, -taki, -male, -kina, -ø*

TABLE 2 Syntagmatic diagram of the Kurtöp verb

A stem is generally one syllable long and any stem can be negated and can take any of the suffixes. The negative prefix does not combine with the verb and all possible suffixes, however. Specifically,

[5]The form of the negative differs according to several factors. A difference in tense is denoted by the use of *ma-* versus *me-/mi-*; *ma-* denotes past tense while *me-/mi-* denotes non-past. In non-past, *me-* is used when the vowel of the stem is non-high and *mi-* is used when the vowel of the stem is high; that is, the negative prefix exhibits assimilation of height. Both past and non-past negatives also agree with the tone of the verbal stem. Verb stems obligatorily have either high or low tone (contrastive following sonorants but predictable following obstruents — voiceless conditions high tone while voiced conditions low tone; see Hyslop (2009)). This tone spreads to the prefix, so that a negative prefix has high tone if the verbal stem has a high tone, while the negative prefix has a low tone if the verbal stem has a low tone.

a verb with the suffixes *-mu* or *-male* cannot be negated. A simple example of a basic tensed verb is (7), showing the verb *ge* 'go' negated and suffixed with the egophoric perfective. Other perfectives are *-pala, -para, -na*, and *-mu*. The difference between these is evidential, mirative, or epistemic in nature, as I describe in some detail in section 3.

(7) ngat **ma**-ge-**shang**
1.Abs Neg-go-Pfv.Ego
'I didn't go.'

Two of the suffixes shown in Table 2 are used to encode imperfective aspect: *-ta* is used to encode mirativity alongside imperfective aspect while *-taki* is used in non-mirative contexts. This difference is the focus of section 3, but (8) provides a brief illustration of an imperfective marked verb.

(8) ngat ge-**taki**
1.Abs go-Ipfv
'I am going.'

Future tense in Kurtöp is encoded by one of four ways: the suffix *-male*, the suffix *-kini*, *-cina*, *-ikina*, or *-ø*. The difference between the four futures is beyond the scope of this paper, so the following example serves as an illustration of future tense:

(9) ngat ge-**male**
1.Abs go-Fut
'I will go.'

2.2.3 Verbal enclitics

Regardless of whether a clause ends with a copula or a tensed verb, any of three possible verbal enclitics may be used as well. A verbal phrase level enclitic[6] may attach to the end of a tensed verb or a copula (i.e. to the right edge of a clause). The forms and functions of these enclitics are summarized in Table 3.

Enclitic	Value
=ri	Hearsay
=sa	Counter Expectation
=mi	Tag

TABLE 3 Verbal enclitics

The examples below show the hearsay enclitic attached to a finite verb in (10), and to a copula in (11) and (12).

(10) khit [ge-shang]**=ri**
3.Abs go-Pfv.Ego=Hsy
'I (heard that) he went.'

(11) khit [ge-wala wenta]**=ri**
3.Abs go-Nmz.Pfv Cop.Eq.Mir=Hsy
'I (heard that) he went indeed.'

(12) khit chorten kora [thung-si nawala]**=ri**
3.Abs stupa circumambulation do-Nf Cop.Exis=Hsy
'(I heard that) s/he keeps circumambulating the stupa.'

A clitic may attach to an already cliticized verb or clause, so that something like (13), with a negative prefix and two enclitics is possible. In elicitation, speakers accept examples such as (13) readily, and any combination of up to two enclitics appears to be possible. For example, *gewalamiri,*

[6]These are separate from enclitics that occur on nominal elements; see Hyslop (2011) for more details about other clitics.

gewalarisa, gewalamisa, gewalasami and *gewalasari* are also possible utterances. There are no examples in the textual database with all three clitics used together and speakers do not accept such combinations as being possible Kurtöp utterances.

(13) ma-ge-wala**=ri=mi**
Neg-go-Pfv=Hsy=Tag
'I heard he went, right?'

With the relevant background into the Kurtöp verbal system in place, I can turn to an in-depth discussion of how mirativity is encoded in the language. I examine the grammatical encoding of mirativity in perfective aspect, imperfective aspect, and in the copulas.

3 Perfective aspect

In perfective aspect Kurtöp combines evidential and evidential-like categories into a five-way contrast. While a detailed description of the functions performed by these forms is beyond the scope of the present article, a brief discussion will aid in placing mirativity in its grammatical and functional context. Figure 2 summarizes the functions of the five perfective suffixes.

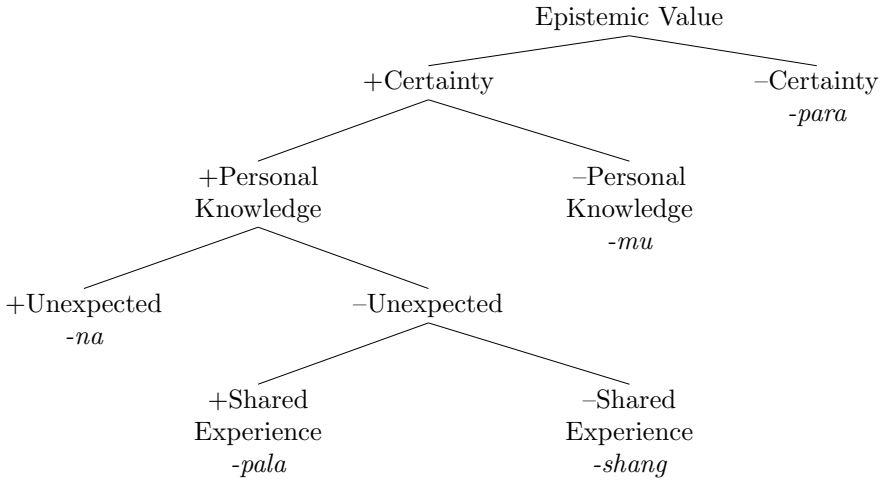


FIGURE 2 Kurtöp perfective aspect suffixes

The suffix *-shang* encodes first person knowledge, which the speaker does not expect the interlocutor to share. This is a similar category, though not identical to, what has been described for Tibetan as 'egophoric' (Tournadre 2008).[7] The suffix *-mu* encodes inference, used when a speaker has indirect knowledge of a given event. *-pala* is pragmatically unmarked in comparison to *-shang, -na*, and *-mu* and is used by default with third person subjects, though it can also be used for first person if the speaker expects the interlocutor to share his/her knowledge. The suffix *-para* encodes that the speaker is not certain of the knowledge.

To encode that knowledge is new and unexpected in perfective aspect the form *-na* is used. A simple example is (14), which was uttered by children who had been watching paragliders take off from the top of a hill. The children watched the paraglider circle around in the sky and slowly rise and fall. At one point the paraglider had disappeared from the children's view and the next time they saw him he had landed. Thus, the mirative is used to encode that the event was unexpected.

[7] I am using this term in a sense similar to that of Tournadre (2008, 295), where 'egophoric' expresses personal knowledge or intention on the part of the actual speaker. Tournadre (2008, 297) describes a contrast between narrow and broad scope of egophorics. The use here in Kurtöp (where expectation of others' knowledge is relevant) is more similar to narrow scope, though the Kurtöp category appears to be slightly different than the Tibetan category.

(14) chak-**na**
land-Pfv.Mir
'It landed!'

A more colorful example comes from a short story about an old woman and Drukpa Künle.[8] At the end of the story the woman was locked inside a room and the villagers were instructed not to open the door for seven days. However, upon the sixth day her son opened the door and discovered there was nothing but her toe remaining. As an event clearly not expected in the discourse it is encoded with the mirative form of the perfective, as shown in (15).

(15) palang=gi je=do thila-the dar-**na**=ri
bed=Gen top=Loc thumb/big.toe-Def remain-Pfv.Mir=Hsy
'On the bed remained a toe! (it is said).' KS20061212.188.829-191.112KL

Similar examples are shown in (16) and (17). The first example, (16), comes from a part of a conversation in which one speaker is relating a traveling event. During this portion of the journey he had reached Trashigang and was looking down on a temple from the top of a hill. There was an important event that day and many people had come; there were so many people, in fact, that a line had formed from the door snaking outward from the temple. The speaker was not expecting a line of people coming from the door, and thus uses the form *-na*.

(16) ko=ni yo=to jong gi-**na**
door=Abl down=Loc emerge go-Pfv.Mir
'(They) had come out of the door down there!' SBC20051127.8.052.KW

Example (17) comes from the same conversation but this time the speaker had reached a house of someone who turned out to be an old relative. The portion in (17), drawn from a longer clause, quotes the relative talking to the speaker. The relative hadn't seen the speaker since the speaker was a child and now, suddenly, the speaker arrives completely unexpectedly, as an adult.

(17) yala... wo onga tshô thrak-**na** wai
god... Prox child here arrive-Pfv.Mir wow
'God... this child has arrived! Wow.' SBC20051127.KW

The data in (18) show the mirative perfective used with second person. This example comes from a conversation between two friends, discussing visits to Rinchen Bumpa, a holy site in the area. Speaker SaT reports that he has been to Rinchen Bumpa four times, while SW has been there only once. SW is surprised to learn that SaT had visited Rinchen Bumpa so many times, and thus uses the mirative *-na*. The choice of the verb *drak* 'be better' indicates that the speaker is making a comparison between himself and his interlocutor, despite the fact the topic of the conversation was not a comparison.

(18) wit drak-**na**=mi tshene=ta
2.Abs be.better-Pfv.Mir=Tag then=Emph
'Then you were better, isn't it!' SaT.SW20081022SW

4 Imperfective aspect

Compared to the contrasts made in perfective aspect, there are a smaller number of evidential-like contrasts made in imperfective aspect.[9] The essential contrast made in imperfective aspect is similar to the difference between Turkish *-di* and *-mış*. Kurtöp *-taki* corresponds to Turkish *-di*, the non-mirative, while Kurtöp *-ta* corresponds to the Turkish mirative *-mış*. As in Turkish, the Kurtöp mirative form is also used in situations of inference and hearsay, alongside situations of surprise.

[8] Drukpa Künle is a popular figure in Bhutanese mythology. Reported to have come to Bhutan from Tibet in the 15th century, he is widely respected as the 'mad monk' and is famous for his practical jokes.

[9] There is also a durative aspect construction that involves a verb suffixed with non-final morphology (*-si*) plus an existential copula, as shown in (6). The evidential value of the clause, then, comes directly from the copula. See section 5.

In other words, the primary contrast in imperfective aspect is essentially between mirativity and non-mirativity; this is summarized in Figure 3.

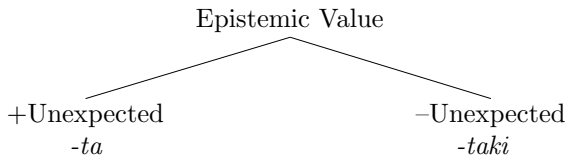


FIGURE 3 Kurtöp imperfective aspect suffixes

The non-mirative form is used canonically (though not exclusively, as I show below) in first person statements and second person questions. In (19) and (20), the unmarked imperfective *-taki* is used with first person arguments in statements. In the case of (19) the agentive argument is first person (*ngai*) while in the case of (20) the first person argument (*ngat*) is the theme. In both instances the speaker is reporting on old, intrinsic knowledge and thus the semantically unmarked form (*-taki*) is used.

(19) wakso go-ikina ngak lap-**taki** ngai
this.much need-Fut.Imm Quot tell-Ipfv 1.Erg
"'I'll need this much", I was saying.' SBC200501127.KW

(20) ngat 'Lama 'Lachung ngak-**taki** la
1.Abs Lama Lachung do-Ipfv Pol
'(They) call me Lama Lachung.' LC20081231.LC

Similarly, second person questions generally require the unmarked form of the perfective. The data in (21) show a monovalent[10] verb with a second person theme; *-taki* is used. A bivalent verb is shown in (22) with a second person agent and the semantically unmarked *-taki* is used. This latter example is slightly more complicated than the previous example in that it is not the speaker who is asking the question, but rather a character in a story told by the speaker. The character (an anthropomorphized tiger, in this instance) still uses the semantically unmarked form *-taki* in his question, as it is questioning information that would be ingrained and not unexpected.

(21) wit 'au jon-**taki** yo
2.Abs where go.Hon-Ipfv Qp
'Where are you (Pol) going?' PS20061206P

(22) 'ap barphela wit zha zus-**taki** ngak-wala wenta
Mr. frog 2.Abs what eat-Ipfv do-Nmz.Pfv Cop.Eq
"'Mr. frog, what are you eating?" (the tiger) said.' SPhTsC20081022.SPh

Third person statements and questions are common in both mirative and non-mirative form. Consider (23).

(23) wici mi=ni yo=to gor tancang kâ me-zak-**taki** wen ngaksi dasum
2.Gen eye=Abl down=Loc turn always blood Neg-drip-Ipfv Cop.Eq Quot today
zha ngâ wo ngaksi wici mi=ni yo=to gor kâ zak-**ta** ngaksi
what do Qp Quot 2.Gen eye=Abl down=Loc turn blood drip-Ipfv.Mir Quot
'Blood doesn't always drip down from your eyes (she said); today what happened? Blood is dripping down from your eyes (she said).' PS20061206.1447.408.P

In this extract from a narrated legend, there are two imperfective clauses. The first clause *kâ me-zak-taki* 'blood Neg-drip-Ipfv' is not mirative. The clause is modified by the adverb *tancang* 'always' and

[10] As discussed in Hyslop (2011), I use the terms 'monovalent' and 'bivalent' to categorize Kurtöp verbs, rather than 'transitive' or 'intransitive'. Monovalent verbs may express up to one overt argument while bivalent verbs may express up to two overt arguments.

the speaker of the clause is describing old knowledge she has about the subject. The next occurrence of the imperfective suffix comes with the same verb and argument as *kâ zak-ta* 'blood drip-Ipfv.Mir', but as a mirative. The switch to mirativity also occurs with a switch from *tancang* 'always' to *dasum* 'today'. The dripping blood in this instance is a new observation on behalf of the speaker in this clause and thus is encoded as mirative.

Another example of *-ta* from discourse is (24), where one speaker is describing an event to another speaker. There are two verbs marked as mirative imperfectives (*tun-ta* 'show-Ipfv.Mir' and *bran-ta* 'know-Ipfv.Mir') and in both instances the event was new, or not expected, in the discourse. The speaker did not know the other people were going to show photographs to the Rimpoche and further, the fact that Rimpoche himself did not know of the photographs was counter to expectation. First, given that Rimpoches are highly regarded as amongst the most knowledgeable people in Bhutan, a default expectation might be that they are aware of everything. And second, it is clear in the larger discourse the photographs being discussed in this example are of the Rimpoche himself. It seems natural to expect that someone would be familiar with photographs of himself.

(24) rimpoche=nang tun-**ta** tshe khi=ra=ya me-bran-**ta**
Rimpoche=Loc show-Ipfv.Mir Dm 3=Emph=also Neg-know-Ipfv.Mir
'They showed (the photographs) to Rimpoche and even he (Rimpoche) didn't know (the photographs).' SBC20051127.142.517KL

A similar example, in (25), comes from an interview between two Kurtöp speakers. One speaker is asking the other to give an account of rice and rice processes in the village. At this point in the narrative she is discussing the varieties of rice given by the government to the village for planting. The fact that outsiders are so involved in the process that they would be giving several varieties of rice for cultivation is not expected, and for this the speaker uses the form *-ta*.

(25) lhampa sum ble=yang bis-**ta** miri
type three four=also give-Ipfv.Mir others.Erg
'Three or four types were also being given by the others.' Rice.Harvest20081022.159.064.PS

As I mentioned above, while *-ta* generally occurs with third person subjects, this is not obligatory. In (26) the mirative *-ta* occurs with a first person subject. Though the speaker presumably has intrinsic knowledge about her ability to narrate stories, she uses the mirative to give a sense of sudden unexpectedness — the discovery of knowledge that will be incontrovertible to the hearers. Here, the speaker is asked to tell a story, begins to speak, and then utters (26) upon realizing she does not feel prepared to tell a story.

(26) me-khan-**ta** ngai=ta lap-to=rang
Neg-know-Ipfv.Mir 1.Erg=Emph tell-Inf=Emph
'I don't know at all (how) to tell (a story).' SPh.TsC20081022.TsC

Another example comes from a story, showing a mirative with first person plural. At this point in the story of Kala Wangpo, the children are being approached by hunters who have been sent by the king's wife to kill them. In the story, the children have known the hunters their whole lives and would usually feel happy to see the hunters. The experience of fear in this example is unexpected, and thus the mirative form *-ta* is used, as shown in (27).

(27) 'aci sharop wit mik thung-mo khepo tshe net pret-**ta** ngaksi
elder.brother hunter 2.Abs eye do-Ctm Foc Dm 1.Pl.Abs fear-Ipfv.Mir Quot
'Elder hunter brother, when (we) see you, well, we feel scared", (the children) said.' PS20061206P

5 Copulas

Copulas play an integral role in Kurtöp grammar; in addition to the typical copular functions (existence, equation, prediction, location, possession), copulas occur with clausal nominalizations as a way to encode finite, main clause grammar (e.g., as illustrated in (4)). Copulas are also used in

conjunction with the clause-chaining construction as a means to encode durative aspect (e.g., as shown in (6)). A detailed description of the role of copulas in main clause grammar is beyond the scope of this article; let it suffice for our purposes to illustrate the basic copular functions before turning to a discussion of mirativity in the copulas.

Kurtöp has affirmative and negative forms for equative and existential copulas, which we saw in Table 1. The equative copulas encode equation while the existential copulas encode existence, predication, location and possession. The constructions for these functions vary slightly, as illustrated in Table 4.

Function	**Construction**
Equation	[NP NP] Cop.Eq
Predication	[NP Adj] Cop.Exis
Location	[NP NP=Loc] Cop.Exis
Existence	[NP NP(=Loc)] Cop.Exis
Possession	[NP NP(=Loc)] Cop.Exis / NP=Gen NP Cop.Exis

TABLE 4 Non-verbal predication in Kurtöp

Equation is encoded by juxtaposition of two NPs followed by an equative copula. Predication is encoded with an adjective following the NP and a clause-final existential copula. The locative function is encoded with the theme NP being followed by a locative-marked NP and an existential copula. Here, locative case marking is required. This differs slightly from existence, which is encoded with two juxtaposed NPs, the second of which may optionally be marked with the locative case marker, and an obligatory clause-final existential copula. Possession may be encoded through two different constructions. The first construction is identical to that of existence: two juxtaposed NPs, the second of which is optionally marked with the locative case marker and a final existential copula. The second construction utilizes the genitive case marker in conjunction with the existential copula; the possessor is clause-initial and receives genitive marking while the possessed NP follows. The existential copula is used in this construction as well. I illustrate each of these constructions below.

An example of the equational copula encoding equations in shown in (28). The NP *ngat* '1.Abs' is followed by the NP *tshering choden*, and the clause is completed with the equational copula *wen*. The polite marker *la* also occurs in this example, but it is not a requirement of the construction.

(28) ngat tshering choden **wen** la
1.Abs Tshering Choden Cop.Eq Pol
'I am Tshering Choden.' SPh.TsC20081022.TsC

All other copular functions are encoded with the existential copula. An example of predication is shown in (29). Here, the NP *bjinlap* is juxtaposed to the adjective *chetoka*, with the existential copula serving as the predicating element.

(29) bjinlap chetoka **nawala**
holiness very.big Cop.Exis
'(The place) is very holy (lit. 'the holiness is very big').' SaT.SW20090917.1211.409SaT

A slightly more complicated example in (30) illustrates the existential copula conveying location. The theme NP *lâ zhipso ngakkhanta* 'the ones doing renovation work' is postposed to follow the copula, while the locative-marked NP *rinchen bumpa* remains in situ.

(30) tshe rinchen bumpa=ro-ya **nawala** lâ zhipso ngak-khan=ta
Dm Rinchen Bumpa=Loc-also Cop.exis work renovation do-Nmz.Ipfv=emph
'At Rinchen Bumpa there are also (people) doing renovation work.' SaT.SW20090917.SW

Existence is conveyed by one NP juxtaposed to a second NP, followed by an existential copula. The second NP may or may not be cliticized with the locative case marker. Example (31) below illustrates the NP *zhapgi zimcung* 'king's mansion' and marked with the quotative, preceding the

existential copula *nawala*. The second NP, demonstrative *yau*, is postposed to the right of the copula. There is no locative marking in this example.

(31) [[zhap=gi zimcung] ngak] **nawala** [yau]
king=Gen mansion Quot Cop.Exis Dem.up
'There is this so-called king's palace up there.' SPh.TsC20081022.3182.357SPh

In (32) I show existence predicated with a locative-marked NP. The demonstrative *wome* is cliticized with the locative marker[11] *-nang*; the second NP *shakhwi* 'hunting dog' is unmarked, and an existential copula follows.

(32) tshe [wome=nang] [shakhwi] **nawara** ngaksi
Dm Dem.down=Loc hunting.dog Cop.Exis.Pres Quot
"'The hunting dog must be down there", (they) said.' PS20061206P

In Kurtöp, there are also two ways to encode possession, both of which employ an existential copula as the predicating element. The data in (33) illustrate possession with the possessor marked as a locative while in (34) the possessor is marked a genitive. It remains unclear what conditions the use of one construction over others.[12]

(33) [net=na] [gari sum] **nawala**
1.Pl.Abs=Loc car three Cop.Exis
'We have three cars.' SBC20051127KW

(34) [neci] [am-the] **nawala** la yau
1.Pl.Gen woman-Def Cop.Exis Pol Dem.up
'We have a woman up there.' SaT.SW20090917.3126.817SaT

5.1 Existential copulas

As summarized in Table 4, existential copulas are used in predication, location, existence and possession. In addition, existential copulas may be used as the final verb in the clause-chaining construction as a means to encode durative aspect (e.g., example (6)).[13]

There are separate roots for the affirmative and negative existential copulas (*na-* for affirmative; *mu-* for negative) and within each of these there are four separate forms. In affirmative contexts the forms make a contrast between certainty and mirativity, while in negative contexts the forms contrast certainty, mirativity, and evidentiality. The diagram in Figure 4 summarizes the encoding of these contrasts.

5.1.1 Affirmative

Among the affirmative copulas there are two forms which signal uncertainty (presumption versus doubt) and two forms which signal certainty. Within the category of certainty, the contrast made is between mirative (*nâ*) situations and non-mirative (*nawala*) situations.

The copula *nâ* provides a mirative value to existential copular clauses. *nâ* contrasts with *nawala* in that *nâ* is used when the speaker has recently come across the information and was not expecting it. The latter form *nawala* is used when the knowledge is older, intrinsic, or not surprising. Consider (35) and (36).

[11] Kurtöp has two locative markers: *=to*, with allomorphs *=ro*, *=ko*, *=go*, *=ngo* and *=na*, *=nang*. The two have an overlapping distribution and the precise difference between the two remains unclear. Some differences are discussed in Hyslop (2010) and Hyslop (2011).

[12] The examples below might suggest that alienable possession is encoded with a locative while inalienable possession is encoded with a genitive. However, there are counter examples to this observation in the texts, and in elicitation speakers can accept either construction for both alienable and inalienable contexts. To further complicate the matter, the locative marker, as shown in (33), appears to be optional. Speakers accept it omitted in elicitation and in the textual database there are several examples of possession encoded by two simple juxtaposed NPs, followed by the existential copula.

[13] As I show in Hyslop (to appear) copulas and auxiliaries differ with regard to presence of the non-final suffix in a clause chain. If an auxiliary is used as the final verb in a chain, the non-final suffix may be omitted. If a copula is used, however, the non-final suffix is required.

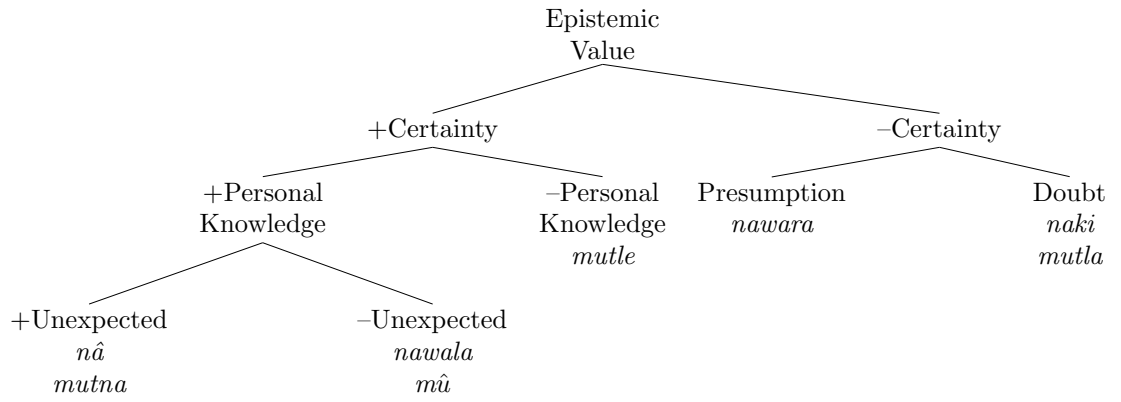


FIGURE 4 Kurtöp existential copulas. Affirmative copulas begin with *na-* while negative copulas begin with *mu-*

(35) tsakaling **nawala** jikpa okso
hat Cop.Exis big this.much
'(His Majesty's) hat was this big.' SPh.TsC20081022.1838.SPh

(36) hâ-pa-the **nâ**
Hâ-Nmz-Def Cop.Exis.Mir
'He was a Hâpa (from Hâ).' SBC20051127.KW

The first example, (35) is from an elderly villager's narration of life as he was growing up. During that period, the King had gone to the speaker's village and the speaker is describing His Majesty's arrival in this section of discourse. As part of this description, the speaker mentions His Majesty's hat, which is large, and this is as expected. The description of His Majesty's large hat can be contrasted with (36), which comes from a conversation between two speakers. While relaying events of a particular journey, the speaker in (36) describes meeting various people along the way. One of the people he meets speaks Kurtöp, but turns out to be from Hâ, an area in western Bhutan where Dzongkha is spoken natively. The speaker did not expect that the referent was from Hâ and thus uses the mirative form of existential affirmative copula.

5.1.2 Negative

A four-way contrast is made also amongst the negative affirmative copulas. The form *mû* is the unmarked negative existential; *mutle* encodes inference; *mutla* encodes doubt; and *mutna* is the negative mirative existential.

The form *mutna* can be contrasted with *mû* in that the speaker recently acquired the information and it was unexpected. *mutna* is used, for example, when the speaker suddenly notices something is not present, for example when looking in his/her wallet and realizing there is no money. If the speaker knew there was no money in his wallet and was telling someone else 'there's no money', s/he would use *mû*.

Example (37) is extracted from a narration of an older villager about what life was like during his childhood. He describes a time when there were poor yields and the living conditions were particularly bad. By using the mirative form of the copula in (37), the speaker paints a picture wherein participants suddenly notice they don't even have a piece of meat to eat. This was an unexpected turn of events.

(37) sha-the zu otor zu-male **mutna**
meat-Def eat like.this eat-Nmz.Irr Cop.Exis.Neg.Mir
'(We) didn't have a piece of meat to eat.' SPh.TsC20081022.1608.669SPh

5.2 Equational copulas

The equational copulas have two main functions: to equate one item with another in copular clauses, or to be used as the final element in a nominalized clause structure (illustrated by (4) above). In the case of either function, the speaker has a set of four affirmative or four negative copulas to choose from, depending on the evidential, mirative, or epistemic modal value s/he wants to convey. Mirativity is amongst these categories in both affirmative and negative contexts.

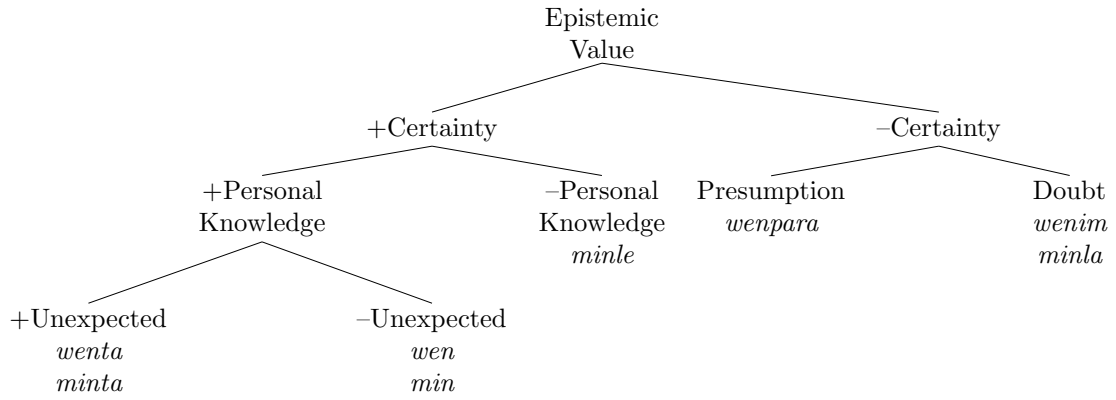


FIGURE 5 Kurtöp equational copulas. Affirmative copulas begin with *wen*(-) while negative copulas bein *min*(-)

5.2.1 Affirmative

As is found for the affirmative existential copulas (5.1.1), the four-way contrast among the affirmative equational copulas consists of two forms used when the speaker is uncertain and two forms when the speaker is certain. The form *wenpara* is used to encode presumption while the form *wenim* encodes doubt. In contexts where the speaker is certain of their knowledge, a two-way contrast is made between mirative and non-mirative utterances, with *wenta* encoding mirative instances and wen used for non-mirative circumstances.

The function of the mirative with an equational copula is similar to the mirative as I described in sections 3, 4, and 5.1. Consider (38) and (39).

(38) khesum ni 'nis **wenta**=mi
sixty and seven Cop.Eq.Mir=Tag
'It's sixty-seven, isn't it.' SPh.TsC20081022.1920.KD

(39) zongkha **wen** la
Dzongkha Cop.Eq Pol
'It is Dzongkha, Sir.' SPh.TsC20081022.3286.436.SPh

The mirative form of the copula given in (38) comes as part of an answer to a question about speakers' ages. Although the information is personal and intrinsic, age is not something the speaker in (38) thinks about frequently. When asked, the particular speaker of (38) has to stop and calculate; when he reaches sixty-seven he replies with mirativity. The use of *wenta* in (38) contrasts with *wen* in (39). Here, the speaker had given the number in Dzongkha and the interlocutor questions the speaker, confirming the number had been given in Dzongkha. (39) is the speaker's response of confirmation.

In (40), the mirative equative copula is shown encoding existence with a second person argument. This example is drawn from a speaker narrating the history of the Kurtöp-speaking region. An important religious figure in the area is Pema Lingpa, who is said to have been enlightened. The speaker in (40) relays a (purported) conversation between a lay person and Pema Lingpa. The lay person, upon realizing that Pema Lingpa is enlightened, says (40), using the mirative form of the

equative copula.

(40) wit sanji 'ngui **wenta** ngaksi
2.Abs Buddha genuine Cop.Eq.Mir Quot
"'You are a real Buddha (enlightened one)!" (he) said.' KZ20080515KZ

Perhaps because of its inherent link to surprise, the mirative equational copula occurs very frequently in storytelling. In fact, speakers report the use of *wenta* 'makes the story more interesting'. In such instances, the copula usually occurs with a formally nominalized verb in an example of a nominalized clause, as in (41). The copula *nawala*, synchronically the semantically unmarked affirmative existential, is diachronically composed of the verb stem **nak* plus the nominalizer *-pala*. Although synchronically it no longer retains any hint of nominalizing semantics, the fact that it occurs in a nominalized structure is indicative of its former status as a nominalized constituent.

(41) lungpa-the=na jepo-the nawala **wenta** la
valley-Def=Loc king-Def Cop.Exis Cop.Eq.Mir Pol
'In a village there was a King.' PS20061212.22.625P

Another example, in (42), comes from part of a narrative where an elderly speaker is describing life in the past. At this point in the narration he is describing a period of time in the distant past when Bhutanese were required to pay taxes to the King in the form of goods, such as rice, pottery, and stones for use as catapults. He deviates from the description somewhat, describing how people used to play with the catapults, and utters *wentami* to invoke the sudden surprise that would have entailed when one was hit.

(42) 'nau-gangsha ras-taki **wenta**=mi
random.thought come-Ipfv Cop.Eq.Mir=Tag
'(They) must have been shocked, right (lit. random thoughts were coming)?'
SPh.TsC20081022.Sph

Consider examples (43) and (44), showing the switch from the unmarked copula to the mirative copula in a narrative. These data come from a story about the legendary king Kala Wangpo. At this point in the story the narrator is reporting speech between the king's servants, who have gone looking for the hunting dog, and an elderly couple the servants meet in a remote region. In (43), the king's servants assert that the hunting dog must be in the elderly couple's possession, using a nominalized clause plus unmarked copula.

(43) neci khwi khepo nin=gi 'lom-pala **wen** ngaksi
1.Pl.Gen dog Foc 2.Pl=Erg hide-Nmz.Pfv Cop.Exis Quot
'You guys have hidden our dog (they said).' PS20061206

They repeat themselves, in (44),[14] but switch to the mirative form of the copula. The narrator has the king's servants use the mirative here as a way to convey an added emotion or force to the utterance. Even though the servants do not actually see the hunting dog, the use of mirative suggests they do and therefore makes it more difficult for the elderly couple to argue against them.

(44) neci shakhwi khepo wo=na 'lom nawala **wenta** ngak nin=gi
1.Pl hunting.dog Foc Prox=Loc hide Cop.Exis Cop.Eq.Mir Quot 2.Pl=Erg
'You guys have hidden our hunting dog (they said).' PS20061206

5.2.2 Negative

As with the negative existential copulas (Figure 4), a four-way contrast is made amongst the negative equational copulas. If the speaker is uncertain, s/he will use the form *minla*; if the speaker gained

[14]This example also shows the agent, *ningi* '2.Pl.Erg', postposed from its canonical position before the verb (more specifically, at the beginning of the clause). Presumably, this movement indicates a particular pragmatic effect, such as topicality or focus, but more research is required to understand the precise function of this movement. This particular example is unusual in that the postposed agent occurs at the far right edge of the clause, following the quotative, almost as an afterthought on behalf of the speaker.

his/her knowledge through inference, then the form *minle* will be used; the form *minta* is used to encode mirativity, and *min* encodes certainty without encoding indirect evidence or mirativity.

The most common use of the mirative negative equational *minta* is when a speaker self-corrects. Consider (45).

(45) **minta** 'yui mendrelgang=ta dara no jepo kut-khan
Cop.Eq.Neg.Mir village Mendrelgang=emph now younger.brother king appoint-Nmz.Ipfv
khepo=nang wenta
Foc=Loc Cop.Eq.Mir
'Oh, not Mendrelgang village, now it is (the place where) the younger brother was appointed king.' PS2006.1751.925.P

The example in (45) is extracted from a section of a narrative where the speaker suddenly realizes she has made a mistake. In the previous clause she mentions the village Mendrelgang but immediately after realizes that it is not the correct place in the story. She speaks to herself, using *minta* 'Cop.Eq.Neg.Mir'. As she thinks out loud she mentions information about the place she is supposed to be referring to; note in this case she uses the affirmative version of the mirative copula.

A similar example is in (46). Here again the speaker self-corrects. He had mistakenly referred to a group of 'two' and upon realizing he was incorrect, he says *minta* 'Cop.Eq.Neg.Mir' and gives the correct number.

(46) net zon **minta** net sum Pema Drakpa net sum
1.Pl two Cop.Eq.Neg.Mir 1.Pl three Pema Drakpa 1.Pl three
'Oh, not the two of us three of us, (with) Pema Drakpa (there were) three of us.' SaT.SW20090919.1120.559.SaT

6 Summary and conclusions

A contrast in mirativity is grammatically encoded in Kurtöp copulas and perfective and imperfective morphology. In perfective aspect the mirative perfective *-na* is one of five possible forms that encode epistemic modality, evidentiality, and expectation of others' knowledge. In imperfective aspect only a two-way contrast is made: mirative clauses are contrasted with non-mirative clauses. In the affirmative existential and equational copulas a four-way contrast is made between presumption, doubt, mirativity, and non-mirativity. This differs somewhat when compared to the contrast made in the negative existential and equational copulas, where mirativity contrasts with doubt, indirect evidence, and non-mirativity.

It may be of theoretical interest that mirativity is more prevalent in Kurtöp than evidentiality. That is, while mirativity is contrasted in perfective aspect, imperfective aspects, affirmative copulas and negative copulas, evidentiality ('source' of information) is contrasted only in perfective aspect and the negative copulas. Note that oral source of information can also be encoded by a hearsay enclitic, as illustrated in Table 3, but is not part of a verbal paradigm in the same way mirativity is. Whether there are any functional motivations for mirativity to be more primary than evidentiality remains unknown and is beyond the scope of this article.

	Aspectual Sufixes	**Copulas**
-na	Perfective	
-ta	Imperfective	
nâ		Affirmative Existential
mutna		Negative Existential
wenta		Affirmative Equational
minta		Negative Equational

TABLE 5 Kurtöp mirative forms

Table 5 displays all the forms used in marking mirativity in Kurtöp. Through a brief comparison of

the forms it quickly becomes apparent that there are two roots involved in mirativity in Kurtöp: *-na* and *-ta*. The former is used to encode mirativity in perfective and existential contexts while the latter is used in imperfective and equational contexts. There is little doubt that these represent two distinct roots which have since grammaticalized into their respective positions. The precise source of these forms remains unknown,[15] as does the motivation for one form to grammaticalize into perfective aspect and existential contexts, while the other form would grammaticalize into imperfective aspect and equational contexts.

7 Acknowledgements

Research on Kurtöp has been generously funded by the National Science Foundation, the Hans Rausing Endangered Languages Project, and the Association for Asian Studies. This research and analysis would not have been possible without the countless hours of assistance Kuenga Lhendup has given to transcribe and translate texts, or the countless hours Karma Tshering has contributed to assistance in data collection and analysis. I am also grateful for comments and feedback from Scott DeLancey, Lila San Roque, William O'Grady, Henriette Daudy and two anonymous reviewers. Nonetheless, I am solely responsible for errors and misinterpretations.

References

Aikhenvald, Alexandra. 2004. *Evidentiality*. Oxford: Oxford University Press.

Aksu-Koç, Ayhan A. and Dan I. Slobin. 1986. A Psychological Account of the Development and Use of Evidentials in Turkish. In W. Chafe and J. Nichols, eds., *Evidentiality: the Linguistic Coding of Epistemology*, pages 159–167. Norwood: Ablex.

Aris, Michael. 1979. *Bhutan, the early history of a Himalayan kingdom*. Warminster, England: Aris & Phillips.

Aronson, Howard. 1967. The grammatical categories of the indicative in the contemporary Bulgarian literary language. In *To honor Roman Jakobson: Essays on the occasion of his seventieth birthday, 11 October 1966*, vol. 1 of *Janua linguarum*, pages 82–98. The Hague: Mouton.

Bashir, Elena. 2010. Traces of mirativity in Shina. *Himalayan Linguistics* 9(2):1–55.

DeLancey, Scott. 1997. Mirativity: The grammatical marking of unexpected information. *Linguistic Typology* 1:33–52.

DeLancey, Scott. 2008. Kurtöp and Tibetan. In B. Huber, M. Volkart, and P. Wildmer, eds., *Chomolangma, Demawend und Kasbek: Festschrift für Roland Bielmeier zu seinem 65. Geburtstag*, pages 29–38. International Institute for Tibetan and Buddhist Studies GmbH.

Dickinson, Connie. 2000. Mirativity in Tsafiki. *Studies in Language* 24(2):379–422.

Friedman, Victor. 1977. *The grammatical categories of the Macedonian indicative*. Columbus: Slavica.

Friedman, Victor. 1986. Evidentiality in the Balkans: Bulgarian, Macedonian, and Albanian. In W. Chafe and J. Nichols, eds., *Evidentiality: the linguistic coding of epistemology. Advances in discourse processes 20*, pages 168–187. Norwood: Ablex.

Grunow-Hårsta, Karen. 2007. Evidentiality and mirativity in Magar. *Linguistics of the Tibeto-Burman Area* 30(2):151–194.

Hein, Veronika. 2007. The mirative and its interplay with evidentiality in the Tibetan dialect of Tabo (Spiti). *Linguistics of the Tibeto-Burman Area* 30(2):195–214.

Hongladarom, Krisadawan. 2007. Evidentiality in Rgyalthang Tibetan. *Linguistics of the Tibeto-Burman Area* 30(2):17–44.

Hyslop, Gwendolyn. 2008. Kurtöp and the classification of the languages of Bhutan. In *Proceedings from the 44nd meeting of the Chicago Linguistic Society*, pages 141–152. Chicago: University of Chicago Press.

Hyslop, Gwendolyn. 2009. Kurtöp tone: a tonogenetic case study. *Lingua* 112:827–845.

[15] One possible scenario is that *-na* is a grammaticalization from a former existential verb *nak* 'to be at', which is clearly the etymological source of the existential copular base in Kurtöp and still present as a lexical verb in Bumthap and Khengkha, Kurtöp's closest linguistic relatives. The verb *tak* 'become', still a lexical verb in Kurtöp, is a possible source for the mirative *-ta* suffix.

Hyslop, Gwendolyn. 2010. Kurtöp case: The pragmatic ergative and beyond. *Linguistics of the Tibeto-Burman Area* 33(1):1–40.

Hyslop, Gwendolyn. 2011. *A grammar of Kurtöp*. Ph.D. thesis, University of Oregon.

Hyslop, Gwendolyn. to appear. The Kurtöp -si construction: converbs, clause chains and verb serialization. In *Functional-Historical Approaches to Explanation: A Festschrift for Scott DeLancey. (Typological Studies in Language Series)*. Amsterdam: John Benjamins.

Lidz, Liberty. 2007. Evidentiality in Yongning Na (Mosuo). *Linguistics of the Tibeto-Burman Area* 30(2):45–88.

Michailovsky, Boyd and Martine Mazaudon. 1994. Preliminary notes on languages of the Bumthang group. In *Tibetan Studies: Proceedings of the 6th Seminar of the International Association for Tibetan Studies*, vol. 2, pages 545–557. Fagernes: The Institute for Comparative Research in Human Culture.

Shafer, Robert. 1954. The linguistic position of Dwags. *Oriens, Zeitschrift der Internationalen Gesellschaft für Orientforschung* 7:348–356.

Shirai, Satako. 2007. Evidentiality and evidential-like categories in nDrapa. *Linguistics of the Tibeto-Burman Area* 30(2):125–150.

Slobin, Dan I. and Ayhan A. Aksu. 1982. Tense, Aspect and Modality in the Use of the Turkish Evidential. In P. J. Hopper, ed., *Tense-Aspect: Between semantics & pragmatics*, page 185 ff. Amsterdam: John Benjamins.

Tournadre, Nicolas. 2008. Against the concept of 'Conjunct'/'Disjunct' in Tibetan. In B. Huber, M. Volkart, and P. Wildmer, eds., *Chomolangma, Demawend und Kasbek, Festschrift für Roland Bielmeier*, pages 281–308. International Institute for Tibetan and Buddhist Studies GmbH.

van Driem, Geoge. 1998. *Dzongkha*. Languages of the Greater Himalayan Region 1. Leiden: Research School CNWS.

Watters, David. 2002. *A grammar of Kham*. Cambridge: Cambridge University Press.

Willis, Christina. 2007. Evidentiality in Darma (Tibeto-Burman). *Linguistics of the Tibeto-Burman Area* 30(2):89–124.