Peripheral and Clause-internal Complementizers in Bangla: A Case for Remnant Movement^{*} Tanmoy Bhattacharya Universität Leipzig

The purpose of this paper is to show that the notion of what is *not* a Phase is equally important as the notion of what constitutes a Phase. Since the notion of a Phase is one particular (albeit an emphatic) instance of the notion of constituency, a *non*-Phase or an incomplete Phase is predicted to be a non-constituent. This paper looks at a curious geometrical puzzle involving clauses with internal Comps in Bangla (=Bengali) and show that such clauses are incomplete phases. In particular, it is shown that the C and its complement are not merged in sequence, nor can they be spelled out as a Phase during the course of the derivation. The claim that the C and its complement do not form a constituent challenges the familiar notion of constituency by showing that an internal C has a non-linear relation to what has been traditionally considered to be its complement.

This challenge is inspired by Kayne's (1998a,b, 1999) demonstration that P-Comps do not form constituents with their complements. Although Kayne's algorithm accounts for a set of unresolved problems involving P-Comps in Romance, it has not yet been tested for Cs in general. This algorithm, if followed verbatim, is shown to derive the unmarked order of constituents but fails to derive the puzzling Ginternal order in Bangla. Another goal of this paper therefore is to present a revised Kaynean algorithm, which, by way of solving the puzzle, is shown to provide crucial evidence for derivation by Phase (Uriagereka 1997, Chomsky 1999). This is a particularly welcome result as it brings two different research strands together.

The implicit claim of the overall analysis is that consistent leftward XP movement derives various surface orders in verb-final languages (Bhattacharya 1998 et seq).

1 The Puzzle itself

In the unmarked case, the complement clause is postverbal and the Comp is in the initial periphery in Bangla, exactly as in English: (1) John jane [je ma kal rate oSudh kheyeche]ⁱ
John knows [that mother last night.LOC medicine ate]
'John knows that mother took medicine last night'

However, if the complement clause is moved to a pre-verbal position (from its *Nachfeld*, which is the post-verbal position for finite clauses in this language), then curiously the Comp can no longer remain in the initial position of the complement clause:

(2) John [ma **je** kal rate oSudh kheyeche] jane John [mother that last nightLOC medicine ate] knows

If the complement were to precede the subject, the same configuration obtains:

(3) [ma je kal rate oSudh kheyeche] John jane

If for some speakers (2) is preferable over (3), this is because specific subjects in Bangla seem to behave like left dislocated subjects (Bhattacharya 2000b, Simpson & Bhattacharya 2001). By all accounts though, the fact that whenever the complement CP moves the C cannot remain in the initial position is a puzzling phenomenon, one that is not readily attested in the world's languages.ⁱⁱ

2 A Na?ve "Disturbed Move" account

A descriptive, therefore naïve, account of the data above leads to two distinct possibilities listed below as options A and B.

(A) Movement induced by "disturbance", i.e., if something within the complement is re-arranged then the clause as a whole must also move.

This option is supported by data such as follows which show that if the complement were to remain in the canonical postverbal position, the C cannot be non-initial:

- (4)a. *amra jantam [ma je aSbe] we knew mother that come.will 'We knew that mother will come'
 - b. *John dekhlo [Robin je khacche]
 John saw Robin that eating
 'John saw that Robin is eating'
 - c. *John bhablo [Sue hEmleT **je** poReche] John thought Sue Hamlet that read 'John thought that Sue has read Hamlet'

(B) Since the clause must move the Comp cannot remain in the initial position.

The data supporting this option is as follows, which shows that if the complement is in a preverbal position, the C within that complement cannot be in the initial position:

(5) a.	*amra [je ma_aSbe] jantam
b.	*John [je Robin khacche] dekhlo
c.	*John [je Sue hEmleT poReche] bhablo

These two options are configurationally represented as follows:

(6)a. $V + *[_{CP} ...C...]$ b. $*[_{CP} C...] + V$

Judging by the supporting data above, it might seem that options (A) and (B) are variants of each other, however, as I shall point out in section 4, there is a real difference here.

3 A Question of Typology

The geometry of the phenomenon that we have witnessed so far raises the following question:

(7) Why must an initial element move inside a clause XP to enable that clause to move inside another clause YP?

This question, in spirit, is actually quite similar to a Greenbergian universal such as the following:

(8) If a language is comp-final, the language is OV.

This universal implies that the internal order within the CP (C-finality) is keyed in to the internal order within the VP (V-finality) and that precisely is the phenomenon we are dealing with. In other words, (7) and (8) are saying the same thing. 'The puzzle' therefore seems to be related to broader issues.

Pre-empting the analysis somewhat at this stage, it may be pointed out that the typological similarity of our question in (7) to a Greenbergian universal has no theoretical importance in the analysis advanced since the question of head-finality has no place in an analysis based first on the LCA and second, on the notion that C and its complement do not actually form a constituent.

4 Phase

A word or two about the notion of PHASE that has already come to occupy an important place in syntactic theorization will set the relevance of the analysis proffered here. The genesis of the concept lies in the importance of derivationality as a crucial algorithm in the way syntax has come to be seen to proceed. This, in return, led to the concern for reduction of complexity since Chomsky (1998) but anticipated since at least Uriagereka (1997). In the latter's work, the idea of Multiple Spell-Out (MSO) appeals to a Dynamically Split Model in which a derivation spells out different chunks of structures in steps. Once a particular unit is spelled out to an intermediate PF (and LF) sequence, it is no longer possible to access its internal constituent structure. It can nonetheless participate in further Merge but only as an inaccessible whole unit. This model therefore provides a reduction of the derivational workspace in the true sense. In Chomsky's formulations, a natural syntactic object is loosely defined as the syntactic equivalent of a proposition in the "meaning side". This corresponds to either a full clause or a verb phrase with all theta-roles assigned, i.e., a CP or a vP. Chomsky calls this unit a PHASE and proposes the following cyclicity condition:

(9) The head of a PHASE is "inert" after the PHASE is completed, triggering no further operations. (Chomsky 1998:20)

This, and the MSO model, virtually ensures that fragments of syntactic objects are inaccessible once the computation is locally complete. In Uriagereka (1997, 1999), this is shown via the classic CED case in (10).

- (10)a. [who did you see [a critic of t]]
 - b. *[who did [[a critic of t] see you]]

Since subjects constitute one single derivational space, extraction out of it is impossible once it derivationally spelled out.

In Chomsky (1998) it is suggested that one of the empirical basis for the concept of Phase is that Phases seem to have a degree of phonetic independenceⁱⁱⁱ. In this connection, one renowned test is the Nuclear Stress Rule (NSR) of Bresnan (1972). For Bresnan, the domains of application for the NSR are S and NP. In view of the current notion, NSR can be argued to be applicable at the level of the verb phrase. Consider (11).

- (11)a. The parable shows what (suffering men) can create. 1
 - b. The parable shows (what suffering) men can create.

In (11a), *what* is the object of the embedded verb *create*. Bresnan shows that indefinites like *what* cannot bear primary phrasal stress even when final in the verb phrase. Instead, the primary stress is assigned to the rightmost element which can bear the stress, the verb *create*. In (11b) the object of the

embedded verb is *what suffering* where *suffering* is assigned the primary stress. When the wh-phrase is moved to [Spec,CP] on the subsequent phase, *suffering* carries its primary stress with it. NSR therefore seems to treat the vP as a domain of operation, i.e. a phase.

A return now to the two options introduced in section 2 will reveal that option B involves a violation of the impenetrability of a Phase implied in (9) above. This is so because the second movement in (12b), representing option B, takes place after the XP has been spelled out.





5 Tuck-in and Remnant Movement

The question (7) raised in section 3, has a theoretical answer. First, the question itself can be translated as follows (ignoring the matrix subject position):

(13) V [a [b c]] \Rightarrow [b [a c]] V

I.e., the relative (precedence) order of the Comp a and the embedded subject b is reversed when the complement moves out of the postverbal position. Fortunately, a combination of available syntactic operations allow this geometry to be derived theoretically. In Bhattacharya (2001), it is briefly shown that, if *Tuck-in* is enforced on Remnant Movement, the combination will have the desired effect of inverting the precedence relation between the two elements a and b. First, a few words about these operations.

5.1.1 Tuck-in

Tuck-in as in (14) is derived from Richards' (1997) study of multiple Wh fronting in Balkan languages who proposes that the Whs must involve crossing rather than nesting paths in their movement to multiple specifiers of a single head.

(14) $Tuck-in^{iv}$

Later XP movement targets inner specifiers, i.e., they tuck in.

5.1.2 Remnant Movement (RM)

This operation involves movement of a category which includes the remnant trace of another category. Therefore in the following, first, YP moves out of XP to the Spec of ZP leaving a trace t_{YP} . Then XP which contains this remnant trace moves out of ZP to the Spec of WP.



RM or Incomplete Category Fronting is exhibited by the following in German where the pre-V2 participle includes a trace of the direct object:

(16)	$[VP t_i Gelesen]_j$	hat [_{IP}	keiner	[_{IP} das]	Buch] _i t _j]]	
	read	has	no-one	the b	book	(Müller 1998)

5.1.3 A combination of Tuck-in and Remnant Movement

If we now enforce *Tuck-in* on RM then the desired goal of inverting the order between a and b in (13) is achieved. In (17b), where move 2 involves RM violates *Tuck-in* since Q (which includes the trace of b) moves to an outer spec of P whereas in the case of (17a) the movement of Q is to the inner spec of P according to the condition in (14). The effect, as can be readily observed, is that in the latter case the precedence order of a and b is reversed, as desired.



In terms of real data, following the above derivation, the puzzling order of (2) (a shorter version of that example) can be now derived as follows:

(18)a.
$$V_{\text{matrix}} [_{CP} je (ma aSbe)]$$

that mother come.will
b. $ma V_{\text{matrix}} [_{CP} je (t_{ma} aSbe)]$
c. $ma [_{CP} je (t_{ma} aSbe)] V_{\text{matrix}} t_{CP}$
 $RM by Tuck-in$

The step in (18c) pertains to the combination of the two operations. After the movement of the embedded subject *ma* out of the CP to an outer spec position, the CP with the remnant trace moves to an *inner* spec position respecting *Tuck-in*. Let us call this the pied piped solution (since the CP pied pipes after the embedded subject moves out).

There are some problems with this solution. First, the set of movement in steps b and c are unmotivated. Secondly, pied piping of the CP in step c is still a descriptive account (as any pied piping account is) and finally, the extraction of the subject from the CP violates derivation by Phase.

6 Topicalization

Question (7) also has an empirical answer which has to do with the fact undisclosed so far that examples like (2) actually have a topicalized meaning. This is clear from the following example:

(19) John [ma je kal aSbe] jane John mother that tomorrow come.will knows
'As for the fact that mother will come tomorrow, John knows it'

Since topicalization is a root phenomenon, this partly explains why the complement must move up. A part of the derivation in (18) therefore can be rescued by appealing to the fact that it is 'mother' which really carries the topic feature and that Pied Piping results in a topicalized meaning of the whole complement when it is moved to a pre-verbal position. However, the other problems remain unaddressed.

7 Kayne's Algorithm

A more interesting solution may be advanced if we consider Kayne's (1998a,b, 1999) radical idea, briefly reviewed below, that the C and its complement does not form a constituent. Kayne demonstrates this via the P-Comp di in (20). The P-Comp in this model does not form a constituent with the infinitival complement IP *cantare*.

(20) Gianni ha tentato *di cantare* John has tried to sing-INF Rather, the derivation proceeds as follows:

(21)a.	Merge matrix V with IP:	tentato + cantare
b.	Merge Comp with (a):	$di + \{tentato, cantare\}$
c.	Comp attracts IP to its Spec:	cantare, di {tentato, t _P }
d.	A new head W is merged and C	adjoins to it:
	di+W	$\{ cantare, t_{di} \{ tentato, t_{\mathbb{P}} \} \}$
e.	Comp(+W) attracts remnant VP	to [Spec,W]:
	{tento	$t_{\rm IP}$, $t_{\rm IP}$, $di+W$ { $cantare$, $t_{\rm di}$, $t_{\rm VP}$ }

The step in (21b) crucially implies that *di* and *cantare* do not form a constituent. Kayne addresses a good many unresolved problems in Romance syntax by letting the derivation proceed in this manner. I direct the reader to the original sources for details. For more immediate concerns, let us see if this algorithm holds water for the problem at hand.

7.1 Je as an attractor

By following the algorithm verbatim, we predict and derive the base order of complements in Bangla, i.e., the order in (1) or (22):

(22)	John jane [je ma aSbe]
	John knows that mother come.will

(23)a. Merge V_{matrix} with complement IP: { $_{\text{VP}}$ jane, { $_{\mathbb{P}}$ ma, aSbe}}

- b. Merge the Comp with (a): $\{je, \{jane, \{IP ma, aSbe\}\}\}$
- c. Comp attracts IP: $\{IP ma, aSbe\}\{je, \{VP jane, tIP\}\}\}$
- d. C to a higher head: $\{je, \{_{IP}ma, aSbe\} \{t_{COMP}, \{_{VP}jane, t_{IP}\}\}\}$
- e. C attracts VP: $\{v_{P} jane, t_{IP}\}\{je, \{I_{P} ma, aSbe\}\{t_{COMP}, t_{VP}, t_{IP}\}\}\}$

Aditionally due to step d, the algorithm also predicts the following:

(24)*John [ma aSbe je] jane

However, it cannot derive the crucial order of (2) and some other orders that I do not discuss here.

7.2 A Revised Kaynean Model

One way to apply this algorithm to our case is by proceeding as follows:

- c. $\{VP, \{C, \{V, \{IP Sub t_{VP}\}\}\}\}$
- d. {C, {VP, { t_C , {V, { $_{IP}$ Sub t_{VP} }}}}
- $\{\{I_{IP} Sub t_{VP}\}, \{C, \{VP, \{t_{C}, \{V, t_{IP}\}\}\}\}\}$ e.

I.e., instead of the IP, the lower VP is attracted in step c and in the last step the remnant IP is attracted. In terms of actual data, the derivation proceeds as follows:

(26)a.
$$je V_{matrix} [IP ma [VP aSbe]] \rightarrow$$

b. $aSbe je V_{matrix} [IP ma t_{VP}]] \rightarrow$
c. $je [aSbe [t_{je} V_{matrix} [IP ma t_{VP}]]] \rightarrow$
d. $[IP ma t_{VP}] je aSbe t_{je} V_{matrix} t_{IP}$

However, some of the problems with the pied piping solution remain here because no motivation has yet been given for the various movement.

7.3 Comp as a Contrast marker

The C-internal clauses in addition to the topicalized meaning seem to set up some kind of contrast^V with the remainder of the complement as well (i.e. the complement without the subject), especially in cases of longer complements:

- (27) John [ma je kal rat-e phOl kheyeche] janto John mother that last night-LOC fruit eaten knew
 - a. 'As for the fact that mother ate fruit last night, John knew it'
 - b. 'As for the fact that mother ate fruit (and not drink wine) last night, John knew it'

This tantamounts to the observation that at the same time as the whole complement is topicalized, part of it gets a contrastive meaning. The VP attraction is (26a) is thus justified as triggered by the need to check a focus-like feature of contrast.

Secondly, perception and intonation experiments show that speakers identify and produce the same intonation contour for both contrastive topic and focus. Speaker B has produced 4 contrasts which all have similar vertical excursions on the pitch accent.

(28) A: Vaši deti uže vzroslye?

'Are your children already adults?'B: Pocti, Andrej ucitsja v universitete, a Vova \u03b8 v gimnazii

'Almost, Andrej studies at universite, a vova ov gimnazi (Mehlhorn, in preparation)

I.e., at some level topic and focus seem to be related to a more general notion of contrast. Syntactically, a clearer proof of this obtains in Bangla where the particle je can induce a clefted meaning (italicised in the

translation) as well whenever it is not in the second position in these C-internal clauses:

(29) John [ma phOl je kheyeche] janto
John mother fruit that eaten knew
'As for the fact that *it was a fruit* that mother ate, John knew it'

This roughly indicates that *je* can carry a general feature of contrast at some level of derivation which subsumes both a topic and a focus feature. This probability can now be used to account for the movement of C, left unmotivated in Kayne's original algorithm, in step (26b) above. This head movement is based on the need for the C to release its topic-like contrast feature in the next step. The last remaining movement, that of the remnant IP in step (26c) is for checking the overall Topic or a Ground feature (i.e. whatever remains after taking out Focus, Vallduví 1992) against the recently moved C head.

This account thus does not rely upon a descriptive mechanism like Pied Piping and it accounts for the fact that the whole complement, and not the subject alone, gets a topicalized meaning.

In addition, this account now provides crucial support for derivation by Phase since the extraction of the embedded VP takes place from a non-phase like IP in step (26a). This possibility, in the first place, is created *because* in the Kanynean algorithm there is no embedded CP to begin with. If there had been an embedded CP, extraction out of it would violate the Phase impenetrability condition. A surprising result of this way of deriving the puzzling order therefore is that Kayne's algorithm, proposed independently of Chomsky's derivation by Phase, provides evidence for the latter.

8 Final Peripheral Comp

Finally, I provide confirmation of the analysis presented from final peripheral Comp cases. Bangla typically employs clause final C *bole* (a form of verb 'to say') also:

(30) amra [ma kal aSbe bole] janiwe [mother tomorrow come.will C] know'We know that mother will come tomorrow'

The complement cannot, in this case, be in a postverbal position. Notice that no topicalized meaning obtains in this case although the complement is in a preverbal postion. Bole is used as a causal marker elsewhere in the language:

(31)a.	mollika aSbe	bole,	anondo	murgi	reMdheche
	Mollika come.v	will because	Anondo	chicken	cooked
b.	robbar bole.	dokan bOn	dho		

Sunday because shop closed

I will therefore assume that a version of a causal feature is carried over when *bole* is used as a Comp. However, there is no feature of contrast involved with this Comp. Unlike *je* therefore, *bole* can allow at most one movement across it since it has only one feature. This prediction is borne out. The derivation for (30) proceeds as follows:

 $(32) \quad bole \ V_{matrix}[_{IP} \ ma \ (_{VP} \ kal \ aSbe))] \ \rightarrow \ ma \ kal \ aSbe \ bole \ V_{matrix} \ t_{IP}$

Next, unlike in the Kaynean model, the C does not obligatory head move in this case. This is due to the fact that *bole* unlike *je* does not contain one or more features of contrast and can only attract the whole IP once to its spec. This suggests that in this case at least a Kayne-like IP attraction for the second step will do the job.

The assertion that *bole* does not carry any feature of contrast can be easily verified from the ungrammaticality of the following:

(33) *amra [ma bole aSbe] jani we mother C come.will know

I.e., *bole* must always be clause final.

Notes

^{*} Thanks to audiences at WECOL 2000, Fresno, at CIEFL, Hyderabad, and at Großbothen for questions and comments on presentations based on a version of this paper. I also wish to thank Richard Kayne, Norvin Richards and Juan Uriagereka for discussions on the problem at length. ⁱ Transcription key: T D R = retroflex td t; S = palato alveolar $\int E O = mid$ vowels $\mathfrak{B} \circ \mathfrak{O}$

ⁱⁱ Though see Bayer (1984) who discusses data like (i) in the Bayarian dialect of German:

⁽i) [xP Da Xaver daß an Mantl kafft hot] hot neamad glaubt

the Xaver t hat a coat bought has, has nobody believed

However, as Bayer points out this possibility is allowed in Bavarian as opposed to standard German because the former's disrespect for the Doubly-filled-Comp filter. I.e., the complement XP is assumed to have moved to the [Spec,CP] position in (i). Bayer's concern therefore is not the position of the C within the complement but with that of the whole complement itself and of the discovery that although other cases of doubly filled Comps in Bavarian (e.g., (ii) below) allow the complement to be at the *Nachfeld* position, XPs of the type in (i), do not. (ii) I woaß ned [xP wer daß des doa hot]

I know not who that this done has

Note that this is simply not possible in Bangla where, if the complement were to remain at the *Nachfeld*, it must have the C in the initial position:

⁽iii)*ami jani na [ke je eTa koreche]

I know not who that this done

¹¹¹ It is also suggested that Phases are reconstruction sites which is why reconstruction takes place to an A-movement trace position rather than to a PRO-site (see (ia)). The phonetic independence issue is also related to the observation that control cases as opposed to raising cases pattern with CPs in being phonetically isolable (see (ib):

(i) a. [one interpreter each] {was assigned t_i/*planned PRO_i to speak} to the diplomat
b. It is to go home (every evening) that John prefers/*seems

^{iv} In Bhattacharya (1999, forthcoming), it is shown that *Tuck-in* applies within the DP in Bangla.

^v Thanks to Probal Dasgupta for judgement on this point.

References

Bayer, Josef. 1984. "COMP in Bavarian syntax", *The Linguistic Review*, 3: 209-274
Bhattacharya, Tanmoy. 1998. "DP-internal NP movement", *University College London Working Papers in Linguistics* 10: 225-252

- Bhattacharya, Tanmoy. 1999. *The structure of the Bangla DP*. London: University College London dissertation
- Bhattacharya, Tanmoy. 2000a. "In search of the vague 'One", ConSOLE 7, ed. by Tina Cambier-Langeveld, Anikó Lipták, Michael Redford and Erik Jan van der Torre, 33-48. Leiden: SOLE
- Bhattacharya, Tanmoy. 2000b. "XP movement in DPs and CPs", Paper presented at the *The Antisymmetry Workshop*, Cortona (May)
- Bhattacharya, Tanmoy. 2001. "The Puzzle of Bangla Comp-Internal Clauses", Snippets 3
- Bhattacharya, Tanmoy. (forthcoming). "Numeral/ Quantifier-Classifier as a Complex Head", *Semi-Lexical Heads*, ed. by Henk van Riemsdijk and Norbert Corver. The Hague: Mouton de Gruyter
- Bresnan, Joan. 1972. "On Sentence Stress and Syntactic Transformations", *Contributions to Generative Phonology*, ed. by Michael Brame, 73-107. Austin: University of Texas Press
- Chomsky, Noam. 1998. "Minimalist Inquiries: The Framework", MIT Occasional Papers in Linguistics, Cambridge, MA: MIT
- Chomsky, Noam. 1999. "Derivation by Phase", *MIT Working Papers in Linguistics* **18**, Cambridge, MA: MIT
- Kayne, Richard S. 1998a. "Overt vs. Covert Movement", Syntax 1: 128-191.
- Kayne, Richard S. 1998b. "A Note on Prepositions and Complementizers", *Internet Celebration for Noam Chomsky*, MIT Press.

Kayne, Richard S. 1999. "Prepositional Complementizers as Attractors", *Probus*, **11**: 39-73

- Mehlhorn, Grit. (in preparation). Empirische Evidenz für die Beschreibung kontrastierter Konstituenten im Russischen, Doctoral dissertation, Universität Leipzig
- Müller, Gereon. 1998. Incomplete Category Fronting: A derivational approach to remnant movement in German. Dordrecht: Kluwer
- Richards, Norvin. 1997. What Moves Where When in Which Language? Cambridge, MA: MIT dissertation
- Simpson, Andrew and Tanmoy Bhattacharya. 2001. "Obligatory *Wh* Movement in a *Wh* in situ Language", *NELS* **30**, Amherst: GLSA

So, though the Bavarian data reported in (i) is superficially similar to the data in Bangla, their relation to postverbal complements in general are different and the interesting fact of the puzzling position of the C inside the complement CP in (i) (same as in Bangla) remains to be analysed. It may also be pointed out that Bavarian (i) above is more like (3) in the text, and not (2), which is the marked case for Bangla.

Uriagereka, Juan. 1997. "Formal and Substantive Elegance in the Minimalist Program (On the Emergence of Some Linguistic Forms)", *The Role of Economy Principles in Linguistic Theory*, ed. by Chris Wilder, Hans-Martin Gaertner and Manfred Bierwisch, 170-204. Berlin: Akademie Verlag

Manfred Bierwisch, 170-204. Berlin: Akademie Verlag Uriagereka, Juan. 1999. "Multiple Spell-Out", *Working Minimalism*, ed. by Norbert Hornstein and Samuel Epstein, Cambridge, MA: MIT Press

Vallduvì, Enric. 1992. The Informational Component. Garland: New York

wecoltb_circ Filename: D:\NIJER Directory: C:\Documents and Settings\Administrator\Application Template: Data\Microsoft\Templates\Normal.dot Title: Tanmoy Bhattacharya Subject: Author: t. knosche Keywords: Comments: 2/17/2001 9:02 PM Creation Date: Change Number: 2 2/17/2001 9:02 PM Last Saved On: Last Saved By: Tanmoy Total Editing Time: 1 Minute Last Printed On: 5/23/2005 1:13 PM As of Last Complete Printing Number of Pages: 13 Number of Words: 3,438 (approx.) Number of Characters: 19,602 (approx.)