

Remarks and Replies

Obligatory Overt *Wh*-Movement in a *Wh*-in-Situ Language

Andrew Simpson
Tanmoy Bhattacharya

Bangla has commonly been assumed to be an SOV *wh*-in-situ language. Here it is suggested that both of these characterizations are incorrect and that Bangla actually has obligatory overt *wh*-movement from a basic SVO word order. This is disguised by a conspiracy of factors but revealed in restrictions on *wh*-scope and certain apparently optional word order possibilities with complement clauses. Adopting a different perspective on the SOV status of Bangla allows for a simple explanation of the patterns observed and raises the possibility that other “*wh*-in-situ” languages may also have (obligatory) overt *wh*-movement.

Keywords: *wh*-movement, *wh*-in-situ, South Asian languages

1 Introduction

Bangla (Bengali) is a South Asian Indo-Aryan language that has always been taken to be strongly head-final and underlyingly SOV (see, e.g., Bayer 1996). Bangla has also commonly been taken to be a *wh*-in-situ language, as there would not appear to be any overt *wh*-movement in regular questions such as (1) and (2).

- (1) jɔn kon boi-ʈa poʃlo
John which book-CL read
‘Which book did John read?’
- (2) jɔn [CP ke cole gæche] bollo
John who left gone said
‘Who did John say left?’

Here we will argue that the *wh*-in-situ characterization of Bangla is incorrect and that in fact, overt *wh*-movement obligatorily takes place in all Bangla question forms. Such *wh*-movement is frequently heavily disguised in Bangla, but revealed (among other places) in restrictions on *wh*-

This article was originally presented in 1999 at NELS 30 (Rutgers University), the University of Southern California, and UCLA. Our thanks to the audiences in these venues for their suggestions and comments.

scope and certain apparently optional word order possibilities with complement clauses. We argue that overt *wh*-movement in Bangla has gone unnoticed in the past because of a conspiracy between two major factors. First, we will suggest that Bangla's underlying word order is not SOV, but SVO. Second, we will argue that the landing site of *wh*-movement is frequently not a fully S-clause-initial position but is instead often hidden lower in the clause by other operations of movement/base generation, causing *wh*-movement to take place undetected.

2 The Position of Clausal Objects and a Restriction on *Wh*-in-Situ

Though Bangla is regularly described as being an SOV language, the positioning of object complement clauses raises suspicions about this description. Finite complement clauses do not necessarily occur to the left of their selecting verb, but may also be found to its right. That is, either position shown in (3) is possible.¹

- (3) jɔn ([_{CP} meri cole gæche]) bollo ([_{CP} meri cole gæche])
 John Mary left gone said Mary left gone
 'John said that Mary left.'

While such pre- or postverbal positioning of the object CP might sometimes seem to be optional, in certain cases the alternation is not free and there is an important restriction relating to the occurrence of *wh*-in-situ in embedded clauses. If a *wh*-phrase occurs in an embedded clause and is intended to have matrix clause scope, the CP must occur in preverbal position, as in (4) and the gloss in (4a).

- (4) ora [_{CP} ke aʃbe] ʃuneche Sub [_{CP} ... wh ...] V
 they who come.will heard
 a. Who have they heard will come?
 b. They have heard who will come.
 (Bayer 1996)

In (5), where the same CP follows the verb, it is no longer possible for the *wh*-subject to take matrix scope and only the indirect reading in gloss (5b) is possible.

- (5) ora ʃuneche [_{CP} ke aʃbe] Sub V [_{CP} ... wh ...]
 they heard who come.will
 a. #Who have they heard will come?
 b. They have heard who will come.
 (Bayer 1996)

If the embedding matrix clause verb does not permit a question as a complement, as in (6),

¹ The presence of an overt complementizer adds certain complications to the distribution of finite CPs. As these complications are arguably not directly relevant for the phenomenon under discussion, they are not investigated here for reasons of space. See Bayer 1996 for much discussion.

postverbal positioning of a CP containing a *wh*-element is ungrammatical, since embedded indirect scope is not available as an option.

- (6) *tumi bhabcho [_{CP} ke baṛi korbe] Sub V [_{CP} . . . wh . . .]
 you think who house make.will
 (Bayer 1996)

Similar patterns occur in Hindi (Mahajan 1990, Srivastav 1991) where, just as in Bangla, a simple *wh*-element with matrix scope cannot occur in an embedded tensed CP located to the right of the verb. A significant difference between Bangla and Hindi is that Hindi does not allow finite complement clauses to occur in preverbal position at all and so an equivalent to Bangla (4) is not possible in Hindi.

This apparent restriction on *wh*-in-situ clearly requires explanation. The first accounts of this phenomenon (in Hindi; Mahajan 1990, Srivastav 1991) argued for an analysis in terms of LF *wh*-movement being blocked. Both authors suggested that postverbal CPs in Hindi are extraposed to their surface position from a regular preverbal object position and that this extraposition creates a barrier for LF movement of the *wh*-phrase. Postverbal CPs are assumed to be adjoined to the matrix clause when they are extraposed, and LF *wh*-extraction from such adjunct constituents is then simply taken to be blocked by Subjacency applying at LF.

Despite the initial plausibility of such an account, more recently the extraposition analysis of postverbal CPs in Hindi and Bangla has come under certain criticism, and there are reasons to believe that some other explanation of the *wh*-patterning should therefore be given. Bayer (1996), for example, points out that it is possible for a matrix clause indirect object to bind a pronoun in a postverbal CP in Bangla as in (7). It is argued that such a bound variable interpretation should not be available if the CP is extraposed and adjoined to a position higher than the indirect object, as the indirect object should then not be able to c-command the pronoun inside the CP.

- (7) tumi prottek-ṭa chele-keṭi bolecho [_{CP} ke ta-keṭi durga pujo-y notun jama
 you each-CL boy-ACC said who he-ACC Durga Puja-LOC new shirt
 kapoṭ debe]
 clothes give.will
 'You told each boy who will give him new clothes at Durga Puja.'

Mahajan (1997) presents similar binding-theoretic arguments against an extraposition analysis in Hindi and notes that R-expressions and other elements in postverbal CPs clearly pattern as if they are c-commanded by preverbal VP-internal indirect objects. As such c-command relations should not exist if postverbal CPs are extraposed and adjoined to higher positions, the conclusion is again that extraposition does not in fact underlie Hindi [Subject V CP] sequences.

Assuming therefore that a simple extraposition analysis is inappropriate to account for the *wh*-patterns in (5)–(6), Bayer (1996) presents a rather different restructuring analysis, which suggests that postverbal CPs are first base-generated as adjuncts but later in the same derivation restructured as rightward complements. Being selected at LF in a noncanonical rightward direction (under the assumption that Bangla is a head-final language), such postverbal CPs are argued to

be barriers for LF extraction of *wh*-elements contained inside them, thus explaining the unacceptability of examples such as (6). Although this proposal captures the binding patterns in (7) and avoids other problems associated with an extraposition account, it is seriously challenged by evidence that rightward CPs are actually not barriers for movement. As shown in (8), overt extraction from a rightward CP is actually fully grammatical. It therefore seems difficult to maintain that the same basic structure should block LF *wh*-movement in other cases as Bayer proposes.

- (8) kriṣṇo [mæleria-te]_i bhablo [_{CP} ram t_i mara gæche]
 Krishna malaria-LOC thought Ram die gone
 ‘Krishna thinks that Ram died of malaria.’

Given these and other criticisms of the restructuring and extraposition accounts discussed in Simpson and Battacharya 2000, we will now propose an alternative analysis that critically does not assume any LF *wh*-movement.

3 Development of an Alternative: Overt *Wh*-Movement

3.1 *Wh*-Movement to a Clause-Internal Licensing Position

The basic patterning observed with complement clauses in Bangla is illustrated in (9) and (10). Regular finite CPs can occur either preverbally or postverbally, whereas CPs containing *wh*-elements with scope higher than the containing CP can occur only in preverbal position. Therefore, the important restriction that needs to be accounted for is why *wh*-elements with higher scope do not seem able to occur in postverbal CPs, as in (10b).

- (9) a. Sub [_{CP} . . .] V
 b. Sub V [_{CP} . . .]
- (10) a. Sub [_{CP} . . . wh . . .] V
 b. *Sub V [_{CP} . . . wh . . .] (bad with matrix scope)

Previous accounts have assumed that the (b)-forms in (9) and (10) are necessarily derived from the (a)-forms in some way, because Bangla is an SOV language. Here we suggest that a straightforward alternative account of the *wh*-patterns in (9) and (10) is available if one simply considers them in the opposite way. Instead of assuming that the (b)-forms are derived via extraposition from SOV-type (a)-forms, we suggest that the (a)-forms are derived from the (b)-forms via raising of the CP from an underlying SVO base structure. An SVO base hypothesis is already supported by the binding phenomena observed in (7) (and in Mahajan 1997), which indicate that postverbal CPs are low in the clausal structure and therefore most naturally in their base positions. If it is now hypothesized that forms such as (10a) are actually derived from an SVO base such as (10b), this alternation can significantly be argued to show that CP *wh*-movement in fact takes place and that in (10a) the CP as a *wh*-phrase raises from a postverbal base position to a *wh*-position located below the subject, resulting in licensing of the *wh*-phrase as schematized in (11).

- (11) Sub [_{CP} wh . . .]_i V t_i

Although Bangla has commonly been assumed to be a *wh*-in-situ language, we now suggest that this is incorrect and that such a perception of Bangla has arisen because there has simply been a tendency to look for *wh*-movement in the “wrong place”—in clause-initial position—as well as to assume that Bangla must be underlyingly SOV. If one now entertains the possibility that the *wh*-licensing position might in fact lie under the regular surface position of the subject instead of being fully S-initial and that Bangla is underlyingly SVO, very soon one can see that *wh*-movement can be assumed to occur overtly in *all* regular “*wh*-in-situ” questions, such as (1) and (2). In cases where a full *wh*-CP is suggested to have raised to preverbal position, as in (2)/(4)/(10a)/(11), such movement of a full clausal *wh*-element will be similar to the cases of *wh*-CP movement that have been observed in languages such as Basque (12) and Quechua (13).²

- (12) [_{CP} Nor etorriko d-ela bihar]_i esan diozu Mireni t_i?
 who come AUX-C tomorrow said AUX Mary
 ‘Who did Mary say will come tomorrow?’
 (de Urbina 1990)

- (13) [_{CP} Ima-ta_i wawa t_i miku-chun-ta]_k Maria t_k muna-n?
 what-ACC child.NOM eat-TNS-Q Maria.NOM want-TNS.3
 ‘What does Maria want that the child eat?’
 (Hermon 1985)

The major surface difference between the Basque/Quechua and Bangla cases of *wh*-CP raising will simply be that the landing site of *wh*-movement is to the right of the regular subject position in Bangla, a property reexamined in section 4.

Concerning the alternation in (10a–b), this patterning can now be straightforwardly explained as follows. If it is assumed that *wh*-movement is forced to take place overtly in Bangla, it can be argued that this is successfully effected in the fully acceptable (2)/(4)/(10a)/(11), where the CP as a *wh*-phrase raises from its postverbal base position to the linearly postsubject *wh*-licensing position, but fails to occur in the ungrammatical (6)/(10b), and it is this failure of obligatory overt *wh*-movement that causes the latter structures to crash (just as similar unraised *wh*-forms are unacceptable in English: **He asked she met who*). In such a fairly simple approach there is clearly no need to invoke any kind of LF *wh*-movement or differing restrictions on overt and covert movement to rule out such forms as ungrammatical, and their unacceptability is explained as a simple lack of overt *wh*-licensing/feature checking.

² LF movement of certain IP clauses containing an in-situ *wh*-phrase has also been suggested to occur as an instance of QR (Quantifier Raising) in *wh*-expletive constructions in Hindi (Mahajan 1990).

Bangla is therefore now hypothesized to have the properties in (14):

(14) *Typological properties of Bangla*

- a. Bangla is a language with obligatory overt *wh*-movement.³
- b. Bangla is an SVO language.⁴
- c. The *wh*-Q-licensing position in Bangla is not necessarily clause-initial, but follows the regular surface position of the subject.

3.2 Long Wh-CP Movement

Once one starts pursuing the line of thought that Bangla has obligatory overt *wh*-movement to a postsubject *wh*-position from an SVO base, it is possible to find other good evidence supporting this hypothesis. One particularly strong argument for overt *wh*-CP movement comes from three-clause structures. If the most deeply embedded third clause contains a *wh*-phrase and the only *wh*-licensing position is in the matrix clause (because the verbs in the matrix and the second clause are selected so as not to embed questions), the structure is licensed, as anticipated under the overt *wh*-movement hypothesis, by long *wh*-CP movement of the lowest CP containing the *wh*-phrase to the matrix clause.

- (15) *tumi* [_{CP} *ke cole gæche*]_i *habcho meri bollo t_i*
 you who left gone thought Mary said
 ‘Who do you think Mary said left?’

Also significant in (15) is that the natural landing site of this long *wh*-CP movement is precisely the postsubject position where the *wh*-licensing position is suggested to be located. Importantly, then, such examples show that a *wh*-CP occurs in exactly the same postsubject position where *wh*-CPs occur in biclausal *wh*-questions, but here the CP is clearly not an argument of the matrix verb ‘think’ and therefore can have reached the postsubject position only via movement from a

³ In connection with the claims made for Bangla here, note that Mahajan (1997:209, fn. 9) suggests that the resistance of *wh*-in-situ elements in Hindi to ‘stranding’ in postverbal positions in monoclausal sentences might support the view that such elements have moved to some fixed position in the overt syntax. Such an observation about stranding (also made to us by Hajime Hoji about Japanese (personal communication)) might seem to buttress the suggestions made here; we thank a reviewer for bringing this to our attention.

⁴ We have suggested that clausal objects undergo raising from an SVO base if this is forced for reasons of *wh*-licensing/feature checking, and later we will argue that there is also a non-*wh* focus trigger for CP-raising. When the object of a verb is however a DP, we suggest that movement to preverbal position is caused by an overt Case-licensing requirement, as such DP-raising is fully automatic (except possibly in cases of ‘stranding,’ which we do not attempt to analyze here; see Mahajan 1997 for much relevant discussion). In connection with this, if Bangla/Hindi nonfinite clauses are analyzed as DPs/nominalizations as argued in Bhattacharya 1994, a Case-licensing approach to the preverbal positioning of object DPs will provide a natural account for why nonfinite clauses in these languages are almost always preverbal (even in the absence of a *wh*-focus interpretation).

lower position. Consequently, it is not unnatural to assume that the surface postsubject position of other *wh*-CPs in biclausal *wh*-questions such as (4) may also be the result of similar *wh*-CP movement from an underlying SVO form.

3.3 *Wh-DP Movement*

More evidence for the claim that Bangla has obligatory overt *wh*-movement comes from patterns involving *wh*-DP movement rather than *wh*-clausal pied-piping. Just as Basque and Quechua allow *wh*-CP raising alongside more regular *wh*-DP movement (see (12) and (13)), many speakers of Bangla allow a second strategy involving the raising of *wh*-DPs or PPs as an alternative to *wh*-CP raising. In addition to the hypothesized *wh*-clausal pied-piping in examples such as (16), the structure in (17) is also possible, in which the CP occurs to the right of the verb and a *wh*-DP from this CP occurs raised in the postsubject *wh*-licensing position (equivalents to (17) are also attested in Hindi).

(16) jɔn [CP ke cole gæche]_i bollo t_i
 John who left gone said
 ‘Who did John say left?’

(17) jɔn ke_i bollo [CP t_i cole gæche]
 John who said left gone
 ‘Who did John say left?’

In the present account it can be suggested that in (17) the CP simply remains in its base-generated position instead of raising into the matrix, and a *wh*-DP from inside the CP is raised to the matrix *wh*-licensing position. The existence of such *wh*-DP raising alongside *wh*-CP raising, and the clear parallels with Basque and Quechua, add further good support to the *wh*-clausal pied-piping hypothesis. It should also be noted that, significantly, the targeted landing site of the object *wh*-DP is again most naturally the postsubject position, precisely where *wh*-CPs are suggested to raise to.

By contrast, SOV accounts of similar patterns in Bangla and Hindi cannot assume such a simple analysis of these patterns, as the postverbal CP in all such accounts is suggested to be an island for extraction, either because it is extraposed and an adjunct or because it is selected as a complement in the noncanonical direction. Concerning Hindi, Davison (1988) suggests that in fact no movement is involved in examples similar to (17) and that the *wh*-phrase is base-generated in the matrix clause as an inner topic. Following Davison’s approach, Bayer (1996) proposes a similar nonmovement account for Bangla as well. However, there is simple evidence in Bangla that such an account cannot be maintained. Specifically, the case particle/postposition occurring on the *wh*-phrase is directly linked to the predicate in the embedded clause, so that if the latter is changed, the former also automatically has to change. This indicates clearly that the *wh*-phrase has indeed been moved from the embedded clause rather than base-generated in the matrix with some default case particle/postposition.

- (18) tumi [ki ɔʃukh-e/*-er]_i bhable [_{CP} ram t_i mara gæche]
 you which illness-LOC/*-GEN thought Ram die gone
 ‘Of which illness did you think that Ram died?’
 (Bayer 1996)
- (19) tumi [kon ɔʃukh-er/*-e]_i bhable [_{CP} t_i kono cikitsa nei]
 you which illness-GEN/*-LOC thought any treatment be-not
 ‘For which illness did you think that there is no treatment?’

Second, the relation of the *wh*-phrase to the gap in the postverbal CP is critically island-sensitive and cannot cross into adjunct or relative clauses, again indicating that movement has occurred, linking the two positions.

- (20) *tumi ke_i kādcho [karon t_i mara gæche]
 you who weep because die gone
- (21) *tumi [kon ɔʃukh-e]_i bhable [_{CP} ram [mohila-t_i [je t_i mara gæche]]
 you which illness-LOC thought Ram woman-CL who die gone
 take] jane
 her knows

3.4 Focus-CP Movement

Additional support for the CP-raising/SVO account of Bangla also comes from a reconsideration of the positioning of (finite) non-*wh* CPs. As with *wh*-CPs there are two patterns, with CP complements occurring either (a) preverbally ([Subject CP V]) or (b) postverbally ([Subject V CP]). The current SVO analysis of Bangla would assume that the (a)-forms should be derived from SVO (b)-forms via CP-raising. Reconsidering the interpretation of such structures provides evidence suggesting that the preverbal positioning is associated with *contrastive focus*, thus offering clear potential support for a CP-raising analysis.

First, if a complement CP does contain a contrastive focus, it indeed most naturally occurs in preverbal rather than postverbal position.

- (22) jɔn [_{CP} or BABA aʃbe] ʃone ni
 John his father come.will heard not
 ‘John didn’t hear that his FATHER will come.’

Second, the most natural position for a CP containing an answer to a *wh*-question is preverbal rather than postverbal position (hence the [Subject CP V] order rather than the [Subject V CP] order in (3) as an answer to the question in (2)). Third, intonation patterns on pre- and postverbal CP forms differ. If the CP follows the verb, then the verb is commonly stressed and prominent, whereas if the CP precedes the verb, prominence and stress on the verb are reduced and sentence prominence on the CP itself is increased. Fourth, cases of ‘‘long CP-movement’’ are possible, where a deeply embedded complement CP is clearly raised to the matrix clause postsubject position from a lower clause argument position in examples very similar to long *wh*-CP movement

(illustrated in (26b) below). It can therefore be suggested that the preverbal positioning of non-*wh* complement CPs commonly results from the raising of such CPs from postverbal positions for reasons of focus.⁵ Like *wh*-CP movement, focus-CP pied-piping has indeed been attested in other languages, for example Basque.

- (23) [JON_i etorriko d-ela t_i bihar]_k esan diot Mireni t_k.
 John come AUX-C tomorrow said AUX Mary
 ‘I have told Mary that it is JOHN that will come tomorrow.’
 (de Urbina 1990)

It is also well documented (e.g., Culicover 1992, Simpson 2000, de Urbina 1990) that focus and *wh*-movement appear to target the same clausal position in many languages. Given this observation, it is rather natural to assume that the preverbal positioning of non-*wh* CPs in Bangla results from a focus-raising operation that in some ways resembles *wh*-CP movement. Indeed, noting that both focused and *wh*-complement clauses appear to occur in the same position following the subject, we suggest that this position in Bangla is not just a *wh*-licensing Q-position but a more general polarity-type phrase (in the spirit of Culicover 1992) that can host and license either *wh*-features or simple focus features and therefore attract either *wh*- or just plainly focused CPs or DPs. However, to this we should add, with emphasis, that although *wh*- and focus movement are suggested to target the same basic functional projection, *wh*- and focus movement in Bangla nevertheless differ significantly and are clearly distinct types of movement. Critically, elements with a purely focused interpretation can raise to and be licensed in the focus position that is available in essentially every clause. Hence, in a three-clause structure like the one in (24), the DP ‘*Hamlet*’ from the lowest clause can be raised into the focus position of either the lowest clause (24a), the intermediate clause (24b), or the matrix clause (24c).

- (24) a. jɔn bhablo [_{CP} meri bollo [_{CP} su [HÆMLET]_i poɾeche t_i]]
 John thought Mary said Sue *Hamlet* read
 ‘John thought Mary said it was *HAMLET* that Sue read.’
 b. jɔn bhablo [_{CP} meri [HÆMLET]_i bollo [_{CP} su poɾeche t_i]]
 ‘John thought it was *HAMLET* Mary said Sue read.’
 c. jɔn [HÆMLET]_i bhablo [_{CP} meri bollo [_{CP} su poɾeche t_i]]
 ‘It was *HAMLET* that John thought Mary said Sue read.’

⁵ Note too that when a *wh*-phrase occurs in a preverbal CP and is interpreted as having *indirect/embedded* scope, as in (4b), it is commonly associated with a contrastive focus-type interpretation. In (4b) there is consequently a preference for the interpretation ‘They have heard *who* will come, not when/why/how they will come’. As there is no parallel common preference for a contrastive focus reading when a *wh*-phrase occurs in a postverbal CP, as in (5), it can be argued that the latter is the neutral, base position for complement CPs and that preverbal CPs occur in a nonneutral derived position, resulting from focus-CP raising. Note also that similar types of contrastive focus reading naturally arise when a CP containing a *wh*-phrase is raised in English; for example:

(i) [What Mary bought]_k I want to know t_k, not when she bought it.
 As (23) shows, Basque has similar phenomena.

However, *wh*-phrases cannot be licensed in these same focus positions, as shown in (25a–b); and in a parallel three-clause structure a *wh*-phrase base-generated in the lowest clause is forced to raise to the *wh*-licensing position in the matrix clause, as shown in (25c). Consequently, the obligatory overt *wh*-movement that has been discussed here is *not* the same as focus movement. If it were simply focus raising, the relevant focus-feature checking should be able to take place in any potential focus position; yet this is clearly not the case.

- (25) a. *jɔn bhablo [_{CP} meri bollo [_{CP} su [ki]_i poɾeche t_i]]
 John thought Mary said Sue what read
 b. *jɔn bhablo [meri [ki]_i bollo [su poɾeche t_i]]
 John thought Mary what said Sue read
 c. jɔn [ki]_i bhablo [meri bollo [su poɾeche t_i]]
 John what thought Mary said Sue read
 ‘What did John think Mary said Sue read?’

It should be noted that the same basic patterns observed here with *wh*-DPs and focused DPs also hold for *wh*-CPs compared with focused CPs, and whereas a focused CP can be licensed in the focus position available in any clause, as shown in (26a–b), a *wh*-CP cannot. In (27a) the *wh*-CP cannot be licensed in the lower clause focus position and is instead forced to raise to the *wh*-licensing position available only in the higher matrix clause, as in (27b). Again, then, the evidence shows that Bangla *wh*-CP/-DP movement has to target specifically *wh*-interrogative positions and is not simply raising for the licensing of any pure focus features associated with *wh*-phrases.

- (26) a. meri bhablo [_{CP} jɔn [_{CP} RAM aʃbe]_i bollo t_i]
 Mary thought John Ram come.will said
 ‘Mary thought it was RAM John said would come.’
 b. meri [_{CP} RAM aʃbe]_i bhablo [_{CP} jɔn bollo t_i]
 Mary Ram come.will thought John said
 ‘It is RAM that Mary thought John said would come.’
- (27) a. *tumi bhable [_{CP} jɔn [_{CP} ke eʃeche]_i bollo t_i]
 you thought John who came said
 b. tumi [_{CP} ke eʃeche] bhable [_{CP} jɔn bollo]
 you who came thought John said
 ‘Who do you think John said came?’

4 The Location of the *Wh*-Licensing Position

4.1 Adjuncts and the Position of *Wh*-Elements

In this section we consider more closely where in the clause the *wh*-licensing position may actually be located. In section 3 we suggested that *wh*-elements target a licensing position that seems to be below the regular surface position of the subject. Here we also note that it is common for a

wh-phrase to occur to the right of any adjuncts present, as in (28), which again might seem to suggest that the *wh*-licensing position is unusually low in the clause in Bangla.

- (28) jɔn borders-e kal [kon boi-ʔa]_i kinlo t_i
 John Borders-LOC yesterday which book-CL bought
 ‘Which book did John buy yesterday at Borders?’

However, whereas (28) represents the most regular ordering of adjuncts with respect to a *wh*-phrase, it actually only constitutes a common preference, and it is also possible for adjuncts to intervene between the *wh*-phrase and the verb as in (29).

- (29) jɔn [kon boi-ʔa]_i borders-e kal kinlo t_i
 John which book-CL Borders-LOC yesterday bought
 ‘Which book did John buy yesterday at Borders?’

This may indicate that the *wh*-licensing position is actually higher than the regular base position of adjuncts and that the positioning of the adjuncts in (28) results from either scrambling these elements to the left of the *wh*-position or simply base-generating them in a second, higher adjunction site. This conclusion is also what one would expect on more general grounds. Given that adjuncts of all types may occur as *wh*-phrases, it has to be assumed that these can all be potentially base-generated below the *wh*-licensing position so as to raise to this position for feature checking; and if the *wh*-focus position were instead to be lower than the regular base position of adjuncts, possibly just above the VP, such elements would incorrectly be expected never to occur licensed as *wh*- or focused elements.

The suggestion that adjuncts are regularly scrambled/base-generated to the left of the *wh*-position is also supported by evidence from focus sentences where an element *je* appears optionally attached to DPs that are focused, as in (30).

- (30) jɔn [CP meri borders-e kal hæmleʔ(-je) kineche] jane
 John Mary Borders-LOC yesterday Hamlet-JE have.bought knew
 ‘John knew that it was *Hamlet* that Mary bought yesterday at Borders.’

Bayer (1996) points out that diachronically this element *je* is likely to be derived from a homophonous element *je* that occurs elsewhere as an embedding C⁰.

- (31) jɔn bollo [CP je meri cole gæche]
 John said JE Mary left gone
 ‘John said that Mary left.’

Synchronically, however, Dasgupta (1980) suggests that focus *je* is not a C⁰, but an enclitic that attaches to the right-hand side of a contrastively focused DP. This is consequently (for Dasgupta) why it may occur on a DP apparently low in the clause in examples like (30) and not obviously in the clause-initial C-domain.

Despite this apparently low position of focus *je* in cases such as (30), there is evidence that focus *je* actually occupies a high clausal functional head position in such examples. If focus *je*

were a simple clitic/focus suffix attached to DPs, one would expect that it could be attached wherever a DP can be focused. However, this is not so: *je* cannot occur with a focused DP in matrix clauses.

- (32) *jɔn HÆMLET_i-je bhablo [su porɛche t_i]
 John *Hamlet*-JE thought Sue read

The ungrammaticality of cases such as (32) is quite unexpected if *je* were simply a DP focus suffix, because in that case *je* should be able to raise to the matrix clause together with the focused DP element. However, the restriction on *je* in cases such as (32) is easily understood if *je* is instead a functional head derived from the embedded clause C-domain subordinator *je*. If *je* is regularly base-generated in a head position in embedded clauses attracting focused DPs to its specifier position, *je* and the focused DP ‘*Hamlet*’ in its specifier in (32) will not form a constituent and hence will not be able to undergo long movement to the matrix clause together. Forms such as (32) are therefore expected to be ungrammatical.⁶

This patterning with *je* now allows important conclusions to be drawn about the surface positioning of adjuncts both in focus sentences and in *wh*-questions. In focus sentences *je* frequently occurs with a focused DP in a position that linearly follows the subject and other adjuncts, as shown in (30). Importantly, however, it is also possible for *je* to appear with a focused DP in a position *preceding* other adjuncts.

- (33) jɔn [CP meri hæmleɫ-je borders-e kal kineche] ʒuneche
 John Mary *Hamlet*-JE Borders-LOC yesterday bought heard
 ‘John heard that it was *Hamlet* that Mary bought at Borders yesterday.’

As the patterns discussed above indicate that *je* and the focused XP in its specifier do not form a constituent and so cannot undergo any movement, it can be assumed that the alternations in (30) and (33) instead result from movement/base generation of the adjuncts in different possible positions in the clause. This then leads to the conclusion that the lowest potential base-generated position of the adjuncts is to the right of the focus position instantiated by *je* and that there is simply a common tendency to reposition such adjuncts leftward to the front of the clause when focus is present.

Returning now to the issue of *wh*-questions, in section 3.4 we suggested that both *wh*- and focus movement target the same general polarity-type head. If it is assumed that the *wh*-licensing position is then the same basic head position as the focus-licensing position, and that the occurrence of *je* provides a good indication of the location of the focus position, this allows for the conclusion

⁶ Note that if *je* could be base-generated as a focus head directly in the matrix clause, one might expect that examples such as (32) would be acceptable, with raising of the focused DP from the lower clause to the specifier of this matrix clause *je*. As (32) is ungrammatical, however, this seems to indicate that *je* is still restricted to occurring only in embedded clauses because of its early origin as a subordinating element. If this is so, one might anticipate that *je* would not be accepted in simple monoclausal focus structures consisting only of a matrix clause, and this expectation is indeed borne out. See also Bhattacharya 2002 for much further investigation of Bangla *je*.

that the *wh*-licensing position is indeed also located above the regular (nontopic) base position of adjuncts.

Significantly, then, reasonable arguments can be made that the *wh*-licensing position is actually not as low as might be thought on the basis of common examples such as (28) where the subject and adjuncts all naturally precede the raised *wh*-phrase object. The patterns observed indicate that there is a common strong tendency to reposition adjuncts to the left of the *wh*-/focus-licensing position in *wh*-questions and focus sentences in Bangla and that this leftward positioning of adjuncts heavily disguises the occurrence of *wh*-movement, in part explaining why the occurrence of obligatory *wh*-movement has essentially gone unnoticed in the past.

These conclusions now suggest that one should also reconsider where the subject is positioned in *wh*-questions such as (28). If the above discussion of *wh*/focus and focus *je*'s origin as a C^0 indicate that the *wh*-/focus-licensing position may be located naturally high in the clause, quite possibly in the C-domain, this might in turn suggest that the subject in *wh*-questions regularly occurs in a high clausal topic-like position, the *wh*-position being located under this topic position. Evidence that this is so comes from the observation that only referentially definite or specific elements occur as subjects preceding *wh*-phrases in the “subject position.” As shown in Bhattacharya 1999, specificity in the Bangla DP is encoded in the relative order of N(P) and numeral + classifier, [[NUM-CL] NP] sequences being interpreted as nonspecific and [NP [NUM-CL]] sequences as specific. The contrast in (34) shows that subjects in *wh*-questions are necessarily specific (also definite) and a DP subject with the nonspecific [[NUM-CL] NP] order cannot occur.

- (34) a. chele du-ʈo [kon boi-ʈa]_i poʈlo t_i
 boy two-CL which book-CL read
 ‘Which books did the two boys read?’
 b. *du-ʈo chele [kon boi-ʈa]_i poʈlo t_i
 two-CL boy which book-CL read

It can therefore be argued that the *wh*-movement hypothesized to take place in Bangla is ultimately *not* exceptional in targeting a low clausal position, and that the *wh*-licensing position in fact occupies a much more regular C-domain location after all. What is therefore particularly interesting and unusual about Bangla is that the surface occurrence of *wh*-movement is frequently heavily disguised by the preference for positioning other non-*wh* arguments and adjuncts in still higher positions in the clause.⁷ This interesting observation may now call for reconsideration of other languages that have been assumed to be *wh*-in-situ languages, and it raises the question whether similar displacement phenomena might be hiding overt *wh*-raising more widely in such

⁷ Time and place adjuncts such as those in (28) and the like optionally occur as topics in S-initial position in many languages, as in the similar English *wh*-question (i).

(i) Yesterday, at Borders, what books did you buy?

Cinque (1999) suggests that such “circumstantial” adverbs/adjuncts are generated either in a low base position or in higher topic positions. Alternatively, they might be assumed to be always generated in a single low base position and then optionally repositioned higher as topics via a defocusing/topicalization movement.

languages. Quite generally, it can be suggested that the pressure to place non-*wh* elements above the *wh*-position in Bangla may be a consequence of the fact that non-*wh* arguments/adjuncts in *wh*-questions regularly constitute presupposed background information contrasting with the new, focused information value of the *wh*-phrase, and as presupposed information the non-*wh* elements are naturally placed in topic-like positions. If similar pressure to partition the clause overtly into old and new information in this way may be present and operative in *wh*-questions in other “*wh*-in-situ” languages besides Bangla, it is clear that *wh*-movement could in fact be concealed in parallel ways in possibly large numbers of languages.

Finally, it can be noted that the placement of *topic-like* subjects in a position preceding the *wh*-phrase in examples such as (34) contrasts with the positioning of *quantified/QP* subjects, which are generally forced to occur to the right of a *wh*-phrase.

- (35) a. ka-ke kew/sudhu meri voṭ dæy-ni
 who-DAT anyone/only Mary vote gave-not
 ‘Who did no one vote for?/Who did only Mary not vote for?’
 b. *kew/sudhu meri ka-ke voṭ dæy-ni
 anyone/only Mary who-DAT vote gave-not

Because quantified elements frequently resist topicalization (e.g., **As for no one/everyone/only Mary, which book did he/she buy?*), QP subjects in Bangla might naturally be expected to remain lower in the clause than other nonquantificational subjects that occur in the topic position. The fact that *wh*-phrases must occur to the *left* of such QP subjects is now simply explained if it is assumed that these phrases are forced to undergo overt raising to a position higher than the nontopic/QP subject position (i.e., [Topic [_{whP} wh_k [Subject (_{QP/nontopic}) . . . t_k . . .]]]). Such patterns consequently provide more good evidence both for the occurrence of overt *wh*-movement in Bangla and for the claim that the *wh*-licensing position is indeed located high in the clause, as in other languages.⁸

⁸ Alternations similar to those in (35a) and (35b) have been noted in other SOV “*wh*-in-situ” languages (e.g., Korean, Japanese), and the unacceptability of *[QP_{subject} wh_{object} V] sequences has been attributed to LF *wh*-movement being blocked by the presence of the QP—an “intervention effect.” “Scrambling” of the *wh*-phrase resulting in an acceptable [wh_{object} QP_{subject} t V] sequence has then been assumed to allow LF *wh*-raising to occur from a position that is not c-commanded by the QP (Beck and Kim 1996). Given the broader patterning found in Bangla, a rather different interpretation of such intervention effect paradigms is now possible in Japanese/Korean-type languages, and it can be suggested that *[QP_{subject} wh_{object} V] sequences may be ungrammatical, not because of restrictions on LF *wh*-movement caused by the c-commanding QP, but because *wh*-movement may have to take place overtly in such languages to a position located above the position of QP subjects, just as in Bangla. In *[QP_{subject} wh_{object} V] forms, this *wh*-movement fails to take place, but it does occur in the well-formed “scrambled” [wh_{object} QP_{subject} t V] sequences. Bangla, with its wide range of evidence for overt *wh*-movement, may therefore indicate that intervention effects can be interpreted as one subtype of evidence revealing overt *wh*-raising and that the various *wh*-in-situ languages that show intervention effects may actually be overt *wh*-movement languages similar to Bangla.

For an extension of the analysis proposed here into the patterns of multiple-*wh* questions, see also Bhattacharya and Simpson, in preparation.

5 Concluding Remarks

We have shown in this article that a language previously assumed to be a *wh*-in-situ language instead actually has regular overt *wh*-movement to a high C-domain *wh*-position. Past failure to detect this *wh*-movement is due to two factors: the common (re)positioning of other non-*wh* elements in clausal positions above the *wh*-position, masking the *wh*-movement; and the assumption that Bangla is underlyingly SOV.

Our main conclusions support four recent influential suggestions. First, to the extent that the account of *wh*-patterns proposed here suggests an SVO rather than an SOV analysis of Bangla, the conclusions provide empirical support for Kayne's (1994) idea that there may be a universal SVO order underlying other surface forms such as SOV. Second, the conclusion that the *wh*-focus-licensing position in Bangla is not exceptionally low in the clause but high in the C-domain as in other languages supports the Universal Base Hypothesis defended in Cinque 1999 that languages have a largely invariant underlying clausal structure and that there is a fixed universal ordering of functional projections. Third, the suggestion that a scope-creating operation of overt movement may be disguised by other factors is fully in line with similar ideas expressed by Kayne (1998) concerning the overt licensing of negative elements. Finally, a new awareness of the fact that *wh*-movement may be subtly concealed by additional operations of movement/base generation of other non-*wh* elements in higher positions opens up the interesting possibility that other so-called in-situ languages might similarly be found to have overt *wh*-movement if a broader range of evidence is reexamined, and it could turn out that *wh*-in-situ is not such a common option as previously assumed.

References

- Bayer, Josef. 1996. *Directionality and Logical Form*. Boston: Kluwer.
- Beck, Sigrid, and Shin-sook Kim. 1996. On *wh*- and operator scope in Korean. *Journal of East Asian Linguistics* 6:1–56.
- Bhattacharya, Tanmoy. 1994. Long-distance agreement in Hindi. Ms., University of Hyderabad.
- Bhattacharya, Tanmoy. 1999. The structure of the Bangla DP. Doctoral dissertation, University of London.
- Bhattacharya, Tanmoy. 2002. Peripheral and clause-internal complementizers in Bangla. In *WECOL 12*, ed. by Vida Samian, 1–25. Fresno: California State University, Department of Linguistics.
- Bhattacharya, Tanmoy, and Andrew Simpson. In preparation. Multiple *wh* syntax and Bangla *wh*-movement. Ms., University of Delhi and University of London.
- Cinque, Guglielmo. 1999. *Adverbs and functional heads: A cross-linguistic perspective*. Oxford: Oxford University Press.
- Culicover, Peter. 1992. Topicalization, inversion, and complementizers in English. In *Going Romance and beyond*, ed. by Denis Delfitto, Martin Everaert, Arnold Evers, and Frits Stuurman, 54–87. OTS Working Papers in Linguistics 91–002. Utrecht: Utrecht Institute of Linguistics.
- Dasgupta, Probal. 1980. Questions and relative and complement clauses in a Bangla grammar. Doctoral dissertation, New York University.
- Davison, Alice. 1988. Operator binding, gaps and pronouns. *Linguistics* 26:181–214.
- Hermon, Gabriella. 1985. *Syntactic modularity*. Dordrecht: Foris.
- Kayne, Richard. 1994. *The antisymmetry of syntax*. Cambridge, Mass.: MIT Press.
- Kayne, Richard. 1998. Overt vs. covert movement. *Syntax* 1:128–191.

- Mahajan, Anoop. 1990. The A/A' distinction and movement theory. Doctoral dissertation, MIT, Cambridge, Mass.
- Mahajan, Anoop. 1997. Rightward scrambling. In *Rightward movement*, ed. by Dorothee Beerman, David LeBlanc, and Henk van Riemsdijk, 186–214. Amsterdam: John Benjamins.
- Simpson, Andrew. 2000. *Wh-movement and the theory of feature-checking*. Amsterdam: John Benjamins.
- Simpson, Andrew, and Tanmoy Bhattacharya. 2000. *Wh-CP pied piping in Bangla*. In *NELS 30*, ed. by Masako Hirotsu, Andries Coetzee, Nancy Hall, and Ji-yung Kim, 583–596. Amherst: University of Massachusetts, GLSA.
- Srivastav, Veneeta. 1991. Subjacency effects at LF: The case of Hindi *wh*. *Linguistic Inquiry* 22:762–769.
- Urbina, Jon Ortiz de. 1990. Operator feature percolation and clausal pied-piping. In *Papers on wh-movement*, ed. by Lisa L.-S. Cheng and Hamida Demirdache, 193–208. MIT Working Papers in Linguistics 13. Cambridge, Mass.: MIT, Department of Linguistics and Philosophy, MITWPL.

(Simpson)

Department of Linguistics

SOAS

University of London

Russell Square

London WC1H 0XG

United Kingdom

as4@soas.ac.uk

(Bhattacharya)

Department of Linguistics

University of Delhi

Delhi 110007

India

tanmoy@du.ac.in

Copyright of Linguistic Inquiry is the property of MIT Press and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.