

Does Hindi-Urdu have feature driven wh movement to Spec vP?

Veneeta Dayal
Rutgers University

dayal@rci.rutgers.edu

Department of Linguistics

18 Seminary Place

New Brunswick, NJ 08901 (USA)

To appear in Linguistic Inquiry 2016

Abstract: Manetta (2010) argues that Hindi-Urdu has feature driven overt wh movement to Spec vP, a position from where wh expressions can take clausal scope. This is argued to provide a unified account of the following aspects of question formation in the language: the tendency of wh expressions to occur in preverbal position, the possibility of overt long movement, and the use of scope marking to question out of finite complements. It is shown that this view of Hindi-Urdu question formation is untenable. We first probe some hidden assumptions in Manetta’s account, revealing crucial gaps in the argumentation. We then discuss new facts, as well as some old ones, that challenge central components of the proposal. Taken together, they establish that the move to capture question formation in Hindi-Urdu through wh movement to Spec of vP does not deliver the promised results.

Key words: feature driven wh movement, vP Phase, wh scrambling, extraction, wh expletives/scope marking, scope freezing

I. Introduction

1.1 Questioning in Hindi-Urdu

Hindi-Urdu (henceforth H-U) is an SOV language with relatively free word order. Finite complements come to the right, often associated with an overt pronominal expression in the preverbal position (Davison 1984, 1988, Gurtu 1985, Mahajan 1990, Srivastav/Dayal 1991, 1996, among others). The core facts are illustrated below:

- 1a. anu-ne **kalam/kyaa** khariidaa
A-ERG pen/what bought
“Anu bought a pen” / “What did Anu buy?”
- b. anu **gaaRii/kyaa** chalaanaa jaantii hai
A car what drive-INF knows
“Anu knows (how) to drive a car.” / “What does Anu know (how) to drive?”
- c. anu (**yeh**) jaantii hai ki **ravi-ne kalam/kyaa** khariidaa
A this knows that Ravi pen/what bought
“Anu knows that Ravi bought a pen” / “Anu knows what Ravi bought.”

A point worth noting in the examples above is the scope of the wh expressions with respect to finite complements. The wh expression in (1c) cannot be interpreted with matrix scope, leading to the view that the finite clause is an island for covert scope taking. For present purposes, the specific accounts that have been proposed for these facts are not as important as the empirical generalization they reveal.¹

Given the core facts in (1), there are two questions that arise. If word order is relatively free, what is the status of wh movement in H-U? If the scope of wh expressions in finite complements is confined to the local domain, how does H-U express long-distance wh dependencies?

With regard to the first question, H-U was initially taken to be a wh in situ language but this does not fit well with the observation that wh expressions tend to occur in the preverbal position. This, however, is not a hard constraint, as can be seen below for questions over subject and indirect object, respectively. The (a) examples have the wh in preverbal position, the (b) examples have them in the neutral position for subjects and indirect objects:

2a. yeh kavitaā **kis-ne** likhii?

This poem who-ERG wrote

“Who wrote this poem?”

b. **kis-ne** yeh kavitaā likhii?

3a. tum-ne paisaa **kis-ko** diyaa

You-ERG money who-DAT gave

“Who did you give the money to?”

b. tum-ne **kis-ko** paisaa diyaa

Both orders given above are acceptable. In fact, sometimes the base order may seem more natural than the one in which the wh expression is in preverbal position. This is so in (3), for example, where (3b) is arguably more natural sounding than (3a). Absent a fully controlled study of the discourse conditions under which variations arise, however, it seems a reasonable working hypothesis to take wh phrases to preferentially appear in the preverbal position. It has been claimed by Kidwai (2000), and following her Manetta (2010), that the pre-verbal position is a focus position to which wh expressions move, with the alternative orders being derived through scrambling.²

With regard to the second question, there are two ways of questioning into finite complements, the first of which involves overt displacement of a wh expression:

4a. siitaa **kaun** soctii hai [ki __ aayegaa]

Sita who thinks that will-come

“Who does Sita think will come?”

b. **kaun** siitaa soctii hai [ki __ aayegaa]

Note that in (4a) the extracted *wh* is in the matrix preverbal position, while in (4b) it is in clause-initial position. The possibility of such displacement is contingent on there being no pronoun *yeh* “this” in the matrix object position (cf. 1c).

The second strategy for questioning into complements is what has been variously called “scope marking”, “partial *wh* movement” and “*wh* expletive strategy” and has by now been attested in a large number of languages (see Dayal 2013 for a recent survey). Its hallmark is that there is an invariant *wh* expression corresponding to *what* in the matrix clause, but the question seems to be about a *wh* expression in the embedded clause:

5. *siitaa kyaa soctii hai [ki kaun aayegaa]*
Sita what thinks that who will come
“Who does Sita think will come?”

There are two possible approaches to this phenomenon, dubbed *direct* and *indirect* dependency approaches by Dayal (1994a). Very briefly, the direct dependency approach aligns scope marking with overt extraction of the kind seen in English, whereby the embedded *wh* expression comes to have matrix scope. It is predicated on the view that the matrix *wh* expression has no semantic role beyond marking scope. In the indirect *wh* dependency approach, on the other hand, the matrix *wh* expression plays the same role that it does in a simple question *What does Sita think?*, marking quantification over propositional variables. The question denoted by the CP associate of *what* is the restriction on this variable, telling us which set of propositions this variable should draw its value from: *Which proposition in the set denoted by who will come does Sita think?* Under this view, the closest English correlate of (5) is the English sequential scope marking construction (Dayal 1996), as given in (6), rather than the translation given in (5):

6. What does Sita think? Who will come?

It is hard to argue that *who* takes scope out of a syntactically separate clause. For such structures, it seems relatively uncontroversial that the indirect dependency approach has to be adopted (Dayal 1996). The question is whether the same should not also apply to standard cases of scope marking, such as (5).

1.2. The Case for feature driven *wh* movement to Spec vP

Manetta (2010), building on Rackwoski and Richards (2005), takes as her starting point the view that the preverbal position in H-U, which hosts *wh* expressions, is Spec vP. Her account, in brief, posits the derivations in (7) and (8) for simple monoclausal questions. (7) and (8) illustrate questioning over subject and object positions, respectively:

- 7a. *hamid-ko kis-ne maaraa*
Hamid-ACC who-ERG hit
“Who hit Hamid?”

- b. [CP C... Hamid-ACC [vP who-Nom [vP v ____ hit]]]
*i*Q *u*Q *uwh*
uwh *iwh* EPP

The *v* head probes its domain, which includes its specifier position, values its uninterpretable [wh] feature as well as its EPP feature by interacting with the wh expression in Spec via *Move*. The C head then probes its domain and values its uninterpretable [wh] feature by interacting with the wh phrase at the left edge of the vP through *Agree*. This also values the uninterpretable Q feature on the wh phrase, leading to an LF that can be interpreted as a direct question. Subsequent scrambling of the object results in the observed word order, allowing the wh expression to surface in the preverbal position.

Questions over objects follow the same path but for the fact that it is the subject, originally merged into Spec, vP, that must be scrambled to yield the right word order:

- 8a. Hamid-ne **kyaa ciiz** dekhii
Hamid-ERG what thing saw
“What thing did Hamid see?”

- b. [CP C... Hamid-ERG [vP what thing [vP v ____ saw]]]
*i*Q *u*Q *uwh*
uwh *iwh* EPP

Generalizing to the long-distance cases, Manetta gives the following derivations for “extraction” and “wh expletive” strategies. We will probe the appropriateness of these descriptive labels in section 2. For now, let us see how the observed effects are derived:

- 9a. siita-ne **kis-ko** socaa ki ravi-ne ____ dekhaa
Sita-ERG who-ACC thought that Ravi-ERG saw
“Who did Sita think Ravi saw?”

- b. [CP C... [vP wh-XP [v... ____] [CP C... [vP ____ [v... ____]]]]]
uwh *u*Q *uwh* *uwh*
*i*Q *iwh* EPP EPP

Here again, the full wh expression *who-ACC* has an uninterpretable Q feature and an interpretable [wh] feature. The probe *v* in the embedded clause interacts with it through *Move*, raising it to the embedded Spec, vP, simultaneously valuing its uninterpretable [wh] feature and satisfying its EPP feature. Manetta gives several arguments for allowing matrix *v* to probe into the embedded CP to the edge of embedded vP, where it can value its uninterpretable [wh] feature and its EPP feature, by moving the wh to Spec of matrix vP. The wh phrase gets its Q feature valued by matrix C and ends up with matrix scope.

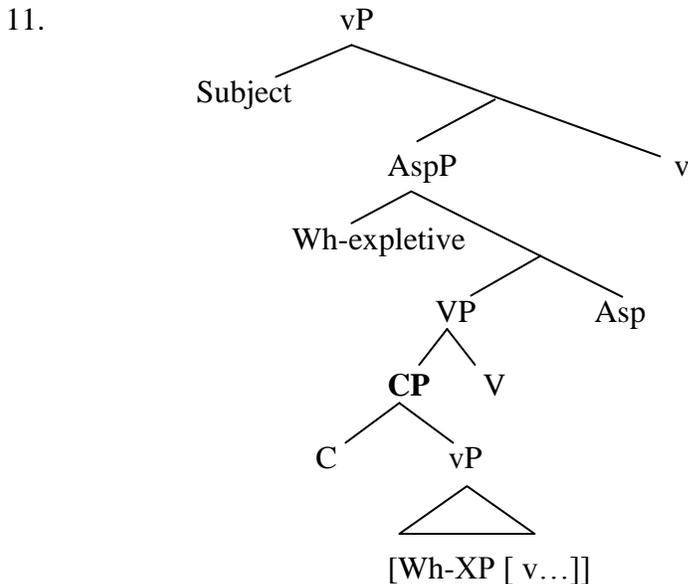
The expletive strategy follows along the same lines up to the embedded clause:

- 10a. siitaa-ne **kyaa** socaa ki ravi-ne **kis-ko** dekhaa
 Sita-ERG what thought that ravi-ERG who-ACC saw
 “Who did Sita think Ravi saw?”

- b. $[_{CP} C \dots [_{VP} \text{wh-expl} [v \dots __]] \ [_{CP} C \dots [_{VP} \text{wh-XP} [v \dots __]]]]]$
 $\begin{matrix} uwh & uQ & uwh & & uQ & uwh \\ iQ & & EPP & & iwh & uQ \\ & & & & & EPP \end{matrix}$

The main difference between (10b) and (9b) is that here the matrix v’s EPP feature is satisfied by a wh-expletive in the numeration. Since it has to get accusative case, the expletive is generated in Spec, AspP where it is assigned case by transitive v and moves into Spec vP to satisfy the EPP feature on v. Being an expletive it has no wh feature. This requires the matrix C head to have its uninterpretable [wh] feature valued by a wh phrase in some accessible position within its domain, its own phase or the edge of the immediately lower phase. It first probes the vP phase in the matrix clause and finds the expletive with no wh feature at the edge and continues to probe. Since the [wh] feature on the matrix v is still unvalued, matrix v gets its feature valued by the wh phrase at the edge of the lower vP. The matrix v then values the wh feature of matrix C, giving the embedded wh matrix scope.

Crucial to Manetta’s account in (9) and (10) is the transparency of the embedded C in allowing matrix v to probe into embedded vP. Here Manetta departs from Rackowski and Richards (2005). For them, the embedded CP is assigned accusative case by matrix v, which then makes it transparent and allows the v to continue probing down to the lower phase, the embedded vP. Manetta’s assumptions about H-U phrase structure are a bit different. As (10b) shows, the wh expletive is generated inside the domain of transitive v, in Spec, AspP where it is assigned accusative case and satisfies the EPP feature on v by *Move*. This, then, raises the question of the CP associate, which Manetta takes to be the true complement of V: $[_{vP} \text{Subject} [[_{VP} \text{CP V}] v]]$. Putting the two together, we get:



- b. $[\text{CP } C \dots [\text{vP non-wh-expl } [v \dots __]] [\text{CP } C \dots [\text{vP wh-XP } [v \dots __]]]]]$
 $\begin{array}{ccccc} uwh & & uwh & & uQ & uwh \\ iQ & & EPP & & iwh & uQ \\ & & & & & EPP \end{array}$
- c. $[\text{CP } C \dots [\text{vP wh-expl } [v \dots __]] [\text{CP } C \dots [\text{vP wh-XP } [v \dots __]]]]]$
 $\begin{array}{ccccc} uwh & uQ & uwh & & uQ & uwh \\ iQ & & EPP & & iwh & uQ \\ & & & & & EPP \end{array}$

In this derivation, matrix *v* needs to value its *uwh* feature, which it can do by probing into the embedded clause. Matrix *C* can now use the features of *v* to value its own *uwh* feature. Given the logical possibilities available within Manetta’s system, then, the prediction is that H-U should allow *wh* expressions inside finite complements to have matrix scope, a prediction that the facts do not support.

As we can see, Manetta leaves a crucial gap in her account by not nailing down the status of the finite complement as a scope island. Since all accounts of long-distance *wh* dependencies in H-U build on this fact, the omission is not a trivial one.

II. The Three Facets of Question Formation in Hindi-Urdu

Let us turn now to the strategies H-U employs in forming questions. We will start with local *wh* movement, which is known to favor the preverbal position. We will see that Manetta’s account does not capture this fact. We will then consider strategies used to override the ban on questioning out of finite clauses, long movement and scope marking. Manetta argues that the first involves extraction and the second an expletive *wh* strategy. We will see that these positions are problematic when the full range of facts, new as well as old, is taken into consideration.

2.1. Local Wh Movement

We have seen the role that the theory of feature checking and the role of Spec of *vP* plays in the analysis of H-U question formation proposed by Manetta. Let us focus on local *wh* movement to see if it captures the adjacency of the *wh* expression and the verb, the primary goal of the paper. As Manetta herself notes, positing a *vP* phase with the EPP feature that attracts a *wh* expression does not guarantee this. A theory of obligatory scrambling is needed to derive the relevant adjacency. Spelling out such a theory is not a trivial enterprise, as can be appreciated by taking even a cursory look at Kidwai (2000) where this is undertaken. Without a similarly explicit specification of the details by which scrambling would be ensured, Manetta’s explanation for the preverbal position for *wh* expressions in terms of *wh* movement to Spec *vP* in H-U is simply a promisory note.

2.2. Long-distance Scrambling

Let us turn now to cases involving finite complements. Gurtu (1985), and following her Mahajan (1990), brought into the public discussion examples of overt wh movement in H-U of the kind discussed in section 1 under the term “extraction”. As noted in Srivastav (1991) and Dayal (1996) such questions had not initially been accepted by many speakers. It was also noted, however, that the sentences become fully acceptable under certain intonational contours. Interestingly, Manetta herself notes a similar problem with Kashmiri in her ft 11 but does not mention this fact for H-U. At any rate, the conclusion proffered in Srivastav/Dayal for the resistance to these data was that the cases under discussion were not the neutral/primary strategy for long-distance dependencies in the language, as we might expect them to be if they represented extraction of the kind seen in English. Instead, they were characterized as cases of long-distance scrambling.

Briefly, there are two pieces of evidence that support the analysis of long-distance wh displacement as scrambling. First, such questions are not neutral requests for information but show the kind of discourse sensitivity typically associated with scrambling. The question in (14), read with stress on the matrix subject, makes the question acceptable but also infuses it with a contrastive reading. The information sought is about the thoughts of the addressee as opposed to someone else:

14. **kaun** TUM socte ho ki ___ aayegaa
Who you think that will-come
“Who do YOU think will come?”

The second piece of evidence, also discussed in Dayal (1996), has to do with the fact that while a direct question interpretation is possible for questions like (14), it is not required. In the context of people wondering what Anu is doing, as in (15), an indirect question interpretation is natural. This is a well-attested fact about scrambling (Saito 1992):

15. Speaker A: anu **kyaa** kar rahii hai
Anu what is-doing
“What is Anu doing?”
Speaker B: **kaun** anu soc rahii hai ki ___ aayegaa
who Anu thinking that will-come
“Anu is wondering who will come.”

Let me emphasize that (15) is not, in and of itself, an argument against Manetta’s account of wh movement in H-U. It merely establishes the existence of long-distance scrambling in the language, that is, of instances of discourse-sensitive long wh movement that do not necessarily lead to matrix scope interpretations. And in doing so, it sets the stage for evidence that does argue against Manetta’s claim that long-distance movement is to the matrix Spec, vP position. Put simply, long wh movement does not show the preference for preverbal position that the intended analysis is designed to capture. In this it differs from local wh movement.

(16a) shows that it is possible, if somewhat marginal, to move a wh phrase out of the complement of a ditransitive verb like *bataanaa* “tell”. Notably, this is only possible if the wh expression moves to clause-initial position, not if it immediately precedes the matrix verb (or even the indirect object and the matrix verb). Note that a pause after the wh phrase helps the acceptability of (16a), but no such modulation redeems (16b) or (16c):⁵

- 16a. ? **kaun**, anu-ne umaa-ko bataayaa ki ___ yehaaN rahtaa hai
 Who Anu-ERG Uma-DAT told that here lives
 “Who did Anu tell Uma lives here?”
 b. *anu-ne umaa-ko **kaun** bataayaa ki ___ yehaaN rahtaa hai
 c. * anu-ne **kaun** umaa-ko bataayaa ki ___ yehaaN rahtaa hai

Another, perhaps clearer, contrast emerges in (17) where the displaced expression is the nominative marked object of an embedded dative subject construction:

- 17a. ? **kaun**, sitaa-ne kahaa ki umaa-ko ___ pasand hai
 Who Sita-ERG said that Uma-DAT is appealing
 “Who did Sita say Uma likes?”
 b. * sitaa-ne **kaun** kahaa ki umaa-ko ___ pasand hai

Similarly telling examples are given in (18)-(19). The embedded agentive subject, with the same case as the matrix subject, can be scrambled to the clause-initial position, as shown in the (a) cases, but not to the matrix preverbal position, as shown in the (b) cases:

- 18a. **kis-ne** sitaa-ne socaa ki ___ ravi-ko dekhaa
 Who-ERG Sita-ERG thought that Ravi-ACC saw
 “Who did Sita think saw Ravi?”
 b. * sitaa-ne **kis-ne** socaa ki ___ ravi-ko dekhaa

- 19a. **kaun** sitaa soctii hai ki ___ yehaaN rahtaa hai
 Who Sita thinks that here lives
 “Who does Sita think lives here?”
 b.* sitaa **kaun** soctii hai ki ___ yehaaN rahtaa hai

Finally, the preference for scrambling to clause-initial position is corroborated by the pairs of sentences in (20) and (21), due to Rajesh Bhatt (p.c.). In (20) the matrix subject and the scrambled embedded object are both marked with *-ko*, though the matrix wh involves dative case and the wh the homophonous accusative case. By varying the form of the dative subject from *tum-ko* to *tum-heN* in (21) we eliminate the possibility that the contrast in (20) is a low-level surface OCP type effect:

20a. **kis-ko** anu-ko lagtaa hai ki ravi ___ pasand kartaa hai
 Who-ACC Anu-DAT seems that Ravi likes
 “Who does it seem to Anu that Ravi likes?”

b. ???anu-ko **kis-ko** lagtaa hai ki ravi ___ pasand kartaa hai

21a. **kis-ko** tumheN lagtaa hai ki ravi ___ pasand kartaa hai
 Who-ACC you-DAT seems that Ravi likes
 “Who does it seem to you that Ravi likes?”

b. (?)tumheN **kis-ko** lagtaa hai ki ravi ___ pasand kartaa hai?

While (21b) is much improved, it remains somewhat marginal. The important point, though, is that in every case the order in which the displaced phrase is clause-initial is perfect, modulo intonation. The one in which it is in preverbal matrix position, instead, is either completely unacceptable or degraded to some extent. It appears that the default position for long wh movement is to the matrix clause-initial, not preverbal, position.⁶

Given what we have seen here, long-distance wh movement in H-U is clearly not analogous to extraction in English as it allows for indirect question readings. And it is clearly not to Spec of matrix vP, as it shows a preference for a clause-initial over a preverbal matrix position.⁷ Any account used to explain local wh movement cannot be extended to cases of long wh movement as they are clearly distinct phenomena.

2.3. Scope Marking

Manetta’s account of scope marking, as we saw in section 1.2, is a direct dependency account. However, it bears mentioning that there is no inherent conflict between indirect dependency and the clausal architecture that Manetta assumes. In fact, Rackowski and Richards (2005), on which Manetta’s analysis is based, endorse the indirect dependency for scope marking. Manetta raises several objections to the indirect dependency approach to H-U scope marking. Most of them stem from fundamental misconceptions about the connection between the semantics of indirect dependency and its syntactic manifestations but to go into them here would take us too far afield.⁸ I will restrict my discussion therefore to the phenomenon of scope freezing, generally thought to follow from the indirect dependency approach. Manetta, however, offers a solution within the direct dependency approach she proposes that merits consideration.

To appreciate what is at stake, consider the extraction structure in (22a). Kroch (1989) noted that such questions are ambiguous between a referential reading (*what is the size of the set of books that Bill said John read?*) and an amount reading (*what is the number such that Bill said John read that many books?*). Lahiri (2002) made the important observation that H-U scope marking structures only have the amount reading. In other words, it has the reading that monoclausal amount questions have. Here we can demonstrate the contrast using the English extraction and sequential scope marking structures in (22a)-(22b):

- 22a. **How many books** did Bill say John read?
 b. **What** did Bill say? **How many books** did John read?

As Lahiri argued, the scope freezing effect follows from an indirect dependency approach since the question in (22b) effectively asks: *which proposition in the set denoted by how many books did John read? is such that Bill said that proposition.*

Manetta addresses the challenge posed by scope freezing for the direct dependency account she proposes in (10).⁹ She adopts Reinhart's (1998) choice functional analysis of long-distance dependencies, where the *wh* expression stays inside the embedded clause. She suggests that overt extraction allows for ambiguity while covert scope taking does not. While this seems plausible enough on the surface, it does not hold up under scrutiny. Consider a language like Japanese, which mirrors precisely the type of LF that Manetta posits for H-U scope marking, one where the *wh* expression remains in the embedded clause and is bound at a distance by an (unselective) Q-binder in matrix Spec, CP:¹⁰

- 23a. [John-ga [Bill-ga **nan-satsu-no hon-o** yonda to] omotteiru -ka]
 J-Nom B-Nom how many book-Acc read C think Q
 (watashi-wa siritai)
 (I - Top want to know.)
 "How many books does John think Bill read?"

- b. Ken-wa [Sato-sensei-ga kinoo **nannin-no kanja-o** shinsatsu-shita]-to
 Ken-top Sato-Dr. yesterday how many-Gen patient-acc see-did-comp
 omotteiru-no?
 think-Q
 "How many patients does Ken think that Dr. Sato saw yesterday?"

The crucial point about these examples is that they do not display the relevant scope freezing effect. (23a) and (23b) are ambiguous in exactly the way that the corresponding English extraction question in (22a) is. We can conclude, then, that the scope freezing in H-U scope marking is evidence of indirect dependency at work. Manetta's account in terms of a long distance dependency is not empirically supported.

III. Conclusion

A unified analysis of seemingly diverse properties of a language is a worthy goal but I have shown here that the attempt in Manetta (2010) to provide such an analysis of H-U question formation falls short. There are assumptions on which the account rests, such as the status of finite clauses as scope islands in H-U, that are not clearly articulated. Focusing on the specific issue of *wh* movement to spec vP, I have shown that an explicit account of the correlation between obligatory scrambling and preverbal focus position in H-U is needed. A proper evaluation of the role of movement to the edge of the vP in H-U must therefore await these elaborations. I have also addressed specific claims about long-distance *wh* dependencies in H-U, long *wh* movement as well as scope marking. I have

shown that long movement is not an instance of standard extraction. And I have shown that the embedded wh in scope marking does not take matrix scope. These conclusions were based on what is already known about these structures as well as on the basis of new facts.

To end this discussion, we might ask: *does H-U have feature driven wh movement to Spec vP?* The answer is that it well may as far as local wh movement is concerned, but we have no evidence for or against it at this point. As far as wh dependencies across clauses is concerned, we have seen evidence against long movement being to Spec vP in one case and against a direct long distance dependency in the other.

Acknowledgments: I am indebted to Mark Baker, Josef Bayer, Rajesh Bhatt, Miriam Butt, Lisa Cheng, Ayesha Kidwai, the participants of Syntax 3 at Rutgers in Fall 2013 and three anonymous reviewers for much useful feedback. All remaining errors and omissions are my own.

References

Davison, Alice. 1984. Syntactic Constraints on Wh in-situ: Wh Questions in Hindi-Urdu. Paper presented at the Annual Meeting of the Linguistic Society of America, Boston.

Dayal, V. (1994a). "Scope Marking as Indirect Wh-Dependency." *Natural Language Semantics* 2.2, 137-170.

Dayal, V. (1994b). "Binding Facts in Hindi and the Scrambling Phenomenon" in M. Butt, T. King and G. Ramchand (eds.) *Theoretical Perspectives on Word Order Issues in South Asian Languages*, CSLI.

Dayal, V. (1996). *Locality in WH Quantification, Questions and Relative Clauses in Hindi*. SLAP 62. Kluwer, Dordrecht.

Dayal, Veneeta (2000). *Scope Marking: Cross-Linguistic Variation in Indirect Dependency*. In: *Wh-Scope Marking*. Uli Lutz, Gereon Müller, and Arnim von Stechow (eds.), 157–193. Amsterdam: John Benjamins.

Dayal, Veneeta (2013). *The Syntax of Scope and Quantification*. Den Dikken (ed), *The Cambridge Handbook of Generative Syntax*. Cambridge: Cambridge University Press.

Gurtu, Madhu. 1985. *Anaphoric Relations in Hindi and English*. Ph.D. thesis, CIEFL, Hyderabad.

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

Kroch, Anthony (1989). *Amount Quantification, Referentiality and Long Wh-Movement*. Unpublished manuscript, University of Pennsylvania.

Lahiri, Utpal. 2002. On the proper treatment of “expletive *wh*” in Hindi. *Lingua* 112:501–540.

Mahajan, Anoop. 1990. The A/A-bar distinction and movement theory. Doctoral dissertation, MIT, Cambridge, MA.

Manetta, Emily (2010). Wh-Expletives in Hindi-Urdu: The vP Phase. *Linguistic Inquiry* 41.1: 1-34.

Rackowski, Andrea, and Norvin Richards. 2005. Phase edge and extraction: A Tagalog case study. *Linguistic Inquiry* 36:565–599.

Reinhart, T. (1998). “Wh-in-situ in the Framework of the Minimalist Program” *Natural Language Semantics*.

Saito, Mamoru (1992). Scrambling as Semantically Vacuous A’ Movement. In M. Baltin and A. Kroch (eds) *Alternative Conceptions of Phrase Structure*. University of Chicago Press: Chicago.

Srivastav, Veneeta. 1991. *Wh Dependencies in Hindi and the Theory of Grammar*. Ph.D thesis, Cornell University.

¹ Let me state at the outset that these facts are uncontroversial and hold across dialects of H-U. The same holds for other facts discussed in this paper.

² Manetta also notes the possibility of focused non *wh* phrases appearing in other positions in the sentence. The data here highlight that the same holds for *wh* phrases.

³ Manetta (2010:20) states “following Rackowski and Richards’s (2005) approach, the matrix *v* head must have some feature that requires it to Agree with the embedded *C*, just as it might interact and agree with a direct object. In Tagalog, the presence of this feature has overt morphophonological consequences; in Hindi-Urdu, it does not. A consequence of this relation is that the phase boundary of the embedded *C* becomes transparent to *v*, and *v* can continue probing down to the next phase edge.”

⁴ It is not entirely clear to me whether matrix *v* needs to have the feature *uQ*. Manetta takes *iQ* to correspond to an unselective existential binder over choice function variables. Given standard assumptions, such binding can occur across islands. At any rate, regardless of whether *uQ* feature should be on matrix *v*, the unavailable reading is predicted within the terms of Manetta’s account. Thanks to an anonymous reviewer and Rajesh Bhatt for helpful discussion.

⁵ Let me reiterate a point made at the beginning, this time with respect to (16) – (21). These data have been verified by several speakers of H-U from different parts of India and Pakistan. The contrasts hold across speakers and dialects. An anonymous reviewer notes, for example, that the contrast between (16a) and (16b) is more of a contrast than a ? versus * but agrees that (17) is clearer and furthermore that the conclusion drawn on the basis of these data is valid.

⁶ Another anonymous reviewer agrees that (16)-(21), as well as the other observations, are robust but suggests that they may be compatible with long wh movement to Spec of matrix vP. The suggestion is that the propensity for wh to appear in clause-initial position may be due to further movement being triggered by C-related features. This would prevent the embedded wh from stopping at Spec of vP. Note though that this move distinguishes long movement in H-U from extraction, a move that goes counter to Manetta’s account where movement to clause initial position is not required for scope. Note also that it would not explain the possibility of the indirect question interpretation we find in (15).

The reviewer also suggests that movement triggered by C-related features may explain the well-known incompatibility of matrix *yeh* “this” and long wh movement:

- (i) **siita** (***yeh**) **kaun** soctii hai [ki ___ ayegaa]
Sita this who thinks that will come
(ii) **kaun** siitaa (***yeh**) soctii hai [ki ___ ayegaa]
“Who does Sita think will come?”

I believe this conclusion is premature. Consider the following perfectly acceptable sentences in which there is both a *yeh* “this” associated with the finite complement and a wh expression which is an argument of the matrix verb:

- (iii) **kis-ne yeh** siitaa-ko bataayaa ki ravi nahiiN aayegaa
Who this Sita-DAT told that Ravi not will-come
(iv) **kis-ne** siitaa-ko **yeh** bataayaa ki ravi nahiiN aayegaa
“Who told Sita that Ravi won’t come?”

Under Manetta’s account the wh expression *kis-ne* and the so-called expletive *yeh* would both have a claim to the preverbal position. Whatever mechanisms would be harnessed to navigate these cases would apply equally to the unacceptable (i)-(ii), and whatever mechanisms are used to rule out (i)-(ii) would apply equally to the acceptable (iii)-(iv).

⁷ The landing site of long movement is an old problem in the literature on scrambling. See Mahajan (1990), Dayal (1994b) and Kidwai (2000) for discussion specific to Hindi. Here I restrict myself to the obvious conclusion that movement is not to the final scope position but to a position from which reconstruction to the embedded Spec, CP is possible (cf. 15).

⁸ For example, Manetta classifies Lahiri’s account of variable binding in scope marking using skolem functions as a semantic workaround. She also finds the restriction on predicates and complements mysterious under the indirect dependency approach. The solutions presented for these facts within the indirect dependency approach are quite explicit and are all independently motivated in the semantics of natural language. Furthermore, there is greater flexibility in the syntax of scope marking within indirect dependency than Manetta recognizes (see, in particular, Dayal 2000 on this).

⁹ Manetta presents the following paradigm to argue against indirect dependency:

(i) mujhe [**yah khabar** [*ki ve log nahiiN aayeNge*]] kal milii
 I-DAT this news that those people not will c-ome yesterday got
 “I got this news that those people won’t come yesterday.”

(ii) mujhe [**yah khabar** kal milii [*ki ve log nahiiN aayeNge*]]

(iii) *sita-ne [**kyaa** [*ki ravi-ne kisko dekhaa*]] socaa
 Sita-ERG what that Ravi-ERG who-ACC saw thought

(i)-(ii) are the non-wh counterparts of scope marking. The CP complement can occur inside the noun phrase in preverbal position (i), or at the right edge of the clause (ii). The scope marking construction (iii) is ungrammatical with the CP complement inside the noun phrase. Since the base form is missing, argues Manetta, there is no source under indirect dependency for the attested form where the CP occurs at the right edge:

Manetta’s paradigm is incomplete. The non wh counterpart of (i) is unacceptable, just like (iii). The CP complement is bad in preverbal position, good at the right edge:

(iv) maiN [**yah** [**ki ve log aayeNge*]] jaantii thii [*ki ve log aayeNge*]]
 I this that those people will come knew that those people will come

The true generalization about scope marking is that a full noun-complement structure is not possible with the wh counterpart, regardless of position:

(v) *sita-ne [**kyaa khabar/baat** [*ki ravi-ne kisko dekhaa*]] socaa
 Sita-ERG what news/matter that Ravi-ERG who-ACC saw thought

(vi) *sita-ne [**kyaa khabar/baat**] socaa [*ki ravi-ne kisko dekhaa*]]

Why *yeh* and *kyaa* differ in this way is an interesting independent issue and needs to be settled before implications for direct versus indirect dependency can be evaluated. It may well pose a problem also for Manetta’s conception of H-U complementation (cf. 11). As a final aside, note that (vii) with a genitive possessive and an NP head is possible:

(vii) tumhaaraa kyaa khyaal hai ki kaun ayegaa
your what thought is that who will come
“What are your thoughts about who will come?”

¹⁰ Thanks to Satoshi Tomioka and Kunio Kinjo for help with these data. Note that to get the relevant ambiguity the embedded wh must not be of the form *nan-satsu hon-o* which has a floated numeral quantifier, because they tend to be non-specific and favor narrow scope, though Satoshi Tomioka points out that this is just a tendency, not a hard fact. The data in (23) does not pose such problems. (23b) can mean *What is the size of the group of patients that Ken said Dr. Sato saw?* or *What is the number such that Ken said Dr. Sato saw that many patients?*