Chapter 4: Wh-expletives in Hindi-Urdu: the *v*P phase

1. Introduction

In this chapter I will address wh-displacement and wh-expletive constructions in the Indic language Hindi-Urdu. These constructions bear many resemblances to their counterparts in Kashmiri, but are distinct in several crucial respects. I will try to show that the approach to wh-expletives that has been developed in previous chapters of this dissertation can be extended to account for Hindi-Urdu, and that a unified account of the various strategies of forming long-distance wh-dependencies in the two languages can be constructed.

There are two ways to form a long-distance wh-dependency in Hindi-Urdu. In the first, depicted in (1), the full wh-phrase which originated in the subordinate clause has been displaced into the matrix clause, and is found in the canonical preverbal wh-position. The semantic correlate of this configuraion is interpretation as a root question.

(1) Sita-ne **kis-ko** socaa ki Ravii-ne _____ dekhaa [Hindi-Urdu]

Sita-erg who-acc thought that Ravi-erg saw

'Who did Sita think that Ravi saw?'

In the second way of forming a long-distance wh-dependency, depicted in (2), the full wh-phrase appears in the canonical preverbal wh-position in the subordinate

clause. In the matrix clause a minimal wh-word *kyaa* appears in the preverbal position. The result is also a matrix reading.

(2) Sita-ne **kyaa** socaa ki Ravii-ne **kis-ko** dekhaa [Hindi-Urdu]

Sita-erg expl thought that Ravi-erg who-acc saw

'Who did Sita think that Ravi saw?'

Superficially, (1)-(2) might appear very similar to long-distance wh-dependencies in Kashmiri below in (3)-(4). However, keep in mind when considering this data that Hindi-Urdu is typically analyzed as a verb final, wh-*in-situ* language, and Kashmiri as a verb-second language with full wh-movement. In the next section we will explore these differences.

Recall that in Kashmiri, the full wh-phrase originating in the subordinate clause can be displaced into the matrix clause, as in (3), or the wh-expletive *k'a:* can appear in the matrix clause, while the full wh-phrase remains in the canonical pre-verbal wh-position in the subordinate clause, as in (4). In both cases the result is an interpretation as a root question.

(3) tse **k´m'** chu-y ba:sa:n ki mohn-as dits __ kita:b [Kashmiri]

you who aux think that Mohan-dat gave book

'Who do you think gave Mohan the book?' (PK 9/21/04)

(4) tse **k'a**: chu-y ba:sa:n ki mohn-as **k´m'** dits kita:b [Kashmiri]

you expl aux think that Mohan-dat who gave book

'Who do you think gave Mohan the book?' (Wali and Koul 1997: 18)

The earliest treatments of wh-movement assumed that such dependencies were of a potentially 'unbounded' span. Later in the course of research on wh-movement, it was assumed that moved constituents were required to move through a sequence of intermediate positions, specifically through the CP layer of each containing clause. Recent research has questioned whether CP is the only intermediate landing site. The concept of the 'phase' (Chomsky, 2000), affords equal status to CP and *v*P, and therefore suggests that the edge of *v*P should be a forced stopping point for long movement as well. An observation supporting this local view of wh-movement is that expletive elements in certain languages can be found in exactly those positions proposed as intermediate stopping-off points, specifically in the specifier of CP. If this is an accurate understanding of wh-movement and its association with the appearance of expletives, then we should expect similar effects wherever intermediate stopping-off points are postulated – in particular, at the edge of *v*P.

In order to develop an understanding of constructions of the type (1)-(2) in Hindi-Urdu, I will bring together three current strands of work on wh-dependencies. The first strand of work is that pursued in the chapters above, as well as in previous work (Simpson 2000, Manetta 2005a, Manetta 2006). This line of research argues that wh-expletives, much like expletives in the A-system (Chomsky 2000) are featurally deficient elements in that they possess only uninterpretable features, lacking an interpretation. When these wh-expletives are merged into the specifier of a functional head, the wh-features on that head are then left free to Agree with an

unraised wh-phrase in its domain. I will claim here that Hindi-Urdu wh-expletives have precisely these properties and effects.

The second strand of research relevant here is a specific view of the *v*P phrase. This strand argues, as mentioned above, that insofar as *v*P is a phase and *v* a phase-defining head, the specifier of *v*P may be as crucial an intermediate stopping off point in the course of long-distance wh-movement as the specifier of CP (Chomsky 2005, Rackowski and Richards 2005). That is, the *v* head has features relevant to wh-movement, and any wh-phrase occurring within the *v*P phase must move first to the phase edge (Spec, *v*P) before interacting with any higher head. We will see what impact this assumption will have both on our previous account of Kashmiri and on our approach to Hindi-Urdu.

Finally, the third strand of research that will prove to be of interest also concerns the heads that are active in wh-movement. The account proposed in this chapter will lend support to the relatively recent proposal that in some significant part, intra-language variation can be attributed to the featural properties of the phase-defining heads (Chomsky 2005, Manetta 2005b). That is, it is the organization of wh-related features on specifically the phase-defining functional heads C and *v* which determine many of the characteristics of questions crosslinguistically.

We will endeavor here to bring these three strands of research to bear on the contrasts in the syntax of wh-dependencies in Hindi-Urdu and Kashmiri. This micro-comparison in the following sections will elucidate differences between (1)-(2) on the one hand and (3)-(4) on the other. Combining the research above with the account

already outlined in this dissertation, we will be able to provide a unified approach to long-distance wh-dependencies in the two languages.

The central claim of this chapter is that the sets of features driving wh-movement and wh-expletive constructions in the two languages are precisely the same. The contrasts between the two will be accounted for as a difference in the properties of the phase-defining heads in each language. We will claim that the features involved in question formation are located on the phase-defining head C in Kashmiri, but on the phase-defining head *v* in Hindi-Urdu.

Section 2 of this chapter outlines the basic sentence structure of Hindi-Urdu, and then turns to describe question formation with a particular focus on long-distance wh-dependencies. I will compare these constructions in Hindi-Urdu with their counterparts in Kashmiri, and based on these observations I will formulate the basic intuition we wish to capture. Section 3 continues in this vein, highlighting a set of contrasts between Hindi-Urdu and Kashmiri which lend empirical support to the observations expressed in Section 2. The fourth section turns to a version of the feature-based approach to Kashmiri familiar from chapters above, to ascertain how this approach can be extended to the facts in Hindi-Urdu. Finally, Section 5 develops an approach to long-distance wh-dependencies in Hindi-Urdu, which takes into account that the specifier of *v*P may be a stopping off point for successive-cyclic wh-movement in some languages. This section also contains a review of alternative views of wh-expletive constructions in Hindi-Urdu and related languages that have not yet been addressed in this dissertation, including Mahajan (1990, 2000), Simpson and

Battarcharya (2003), and a further discussion of Dayal (1996). In this subsection I will argue that while the proposals introduced here incorporate many aspects of previous approaches, they provide a better overall understanding of the facts internal to Hindi-Urdu, allow a better understanding of the contrasts with Kashmiri, and are better integrated theoretically. I also emphasize that the present account will offer a solution to the longstanding puzzle of wh-*in-situ* in Hindi-Urdu – why the wh-expletive construction exists at all.

2. Wh-dependencies in Hindi-Urdu

2.1 Comparing Hindi-Urdu and Kashmiri

Unmarked word order in Hindi-Urdu[1] is verb final, and in a transitive sentence the subject typically precedes the object.

(5) Hamid-ne pani piya [Hindi-Urdu]

Hamid-erg water drank

'Hamid drank water.'

[1] 'Hindi' and 'Urdu' are the names given to a standard variety of a Western Indic language spoken in India and Pakistan. Historically, the common language was referred to as 'Hindustani'. The term 'Urdu' was the name given by the Mughals to the language of the court, and Indo-Mustlim literary figures referred to the language of the subcontinent as 'Hindi' as early as the 17th century (Kachru 1990). In essence, the two terms currently differentiate distinct literary vocabularies and orthographic systems. Their status as separate 'languages' was institutionalized under British rule and, following partition, by the divided states of India and Pakistan. Popularly, these languages are viewed as distinct, though their syntax is identical. As Aisha Kidwai writes "Because the distinction between the two languages does not derive from ethnicity, religion, or culture … I refer to this common language as Hindi-Urdu" (Kidwai 2000: 153). I will follow Kidwai in this choice.

(6) a. bacci-ne mehmaan-ko phul pesh kiye [Hindi-Urdu]

child-erg guest-acc flowers present aux

'The child presented flowers to the guest.' (Schmitt 1999)

b. main sab se pehle is bande-ko tang karta hu: ñ

I all before this guy-acc bother do aux

I bother this guy first (before all others). (3/31/06, Classfellows)[2]

Kashmiri, as we have seen in previous chapters, is a verb-second language, with a variety of phrase types potentially appearing before the verb. (7a), for instance, is an unmarked word order, and (7b-e) are also grammatical (Wali and Koul, 1997: 89).

(7) a. aslaman <u>dits</u> mohnas kita:b ra:mini kh´:trÆ ra:th [Kashmiri]

aslam-erg gave Mohan-dat book Ram-dat for yesterday

'Aslam gave Mohan a book for Ram yesterday.'

b. mohnas <u>dits</u> aslaman kita:b ramini kh´trÆ ra:th [Kashmiri]

c. kita:b <u>dits</u> aslaman mohnas ramini kh´trÆ ra:th [Kashmiri]

d. ra:mini kh´:trÆ <u>dits</u> aslaman mohnas kita:b ra:th [Kashmiri]

e. ra:th <u>dits</u> aslaman mohnas kita:b ra:mini kh´:trÆ [Kashmiri]

Hindi-Urdu is often described as a wh-*in-situ* language, meaning that the wh-word does not appear to be dislocated from its original position. More accurately, in constituent questions, the unmarked position for the interrogative phrase is

[2] (http://www.freewebs.com/zukhruf/classfellows.htm)

immediately before the sentence-final verb, regardless of the grammatical role that it bears (Bhatt 2003b, Schmidt 1999, Kidwai 2000).

(8) kitaab-ko kis-ne parhi? [Hindi-Urdu]

book-acc who read

'Who read the book?'

(9) a. Hamid-ne kyaa parha? [Hindi-Urdu]

Hamid-erg what read

'What did Hamid read?'

b. abhi kis-ko dekhta haiñ?

now who-acc look aux

Who are you looking at now? (3/31/06, Paklinks)[3]

(10) aap chai kab pite haiñ? [Hindi-Urdu]

you tea when drink aux

'When do you drink tea?'

Hindi-Urdu is a language that permits relatively free scrambling of constituents, so for examples like (5)-(6) and (8)-(10) a variety of other word orders are possible, bearing various marked interpretations. These alternatives will be of less interest to us here than the unmarked order, but they will be addressed to some extent below.

As is familiar from the chapters above, Kashmiri wh-phrases immediately precede the second-position verb. A contrastive topic may also appear pre-verbally just when a wh-phrase is present (Bhatt 1999).

[3] (http://www.paklinks.com/gs/archive/index.php/t-25313.html)

(11) a. rajan kÇmis he:v nev kita:b? [Kashmiri]

Raj whom showed new book

'To whom did Raj show his new book?' (Wali and Koul:12)

b. kÇm' he:v shi:las nev kita:b ra:th [Kashmiri]

who showed Sheila new book yesterday

'Who showed a new book to Sheila yesterday? (Wali and Koul: 12)

Embedded clauses in Hindi-Urdu have the same word order as main clauses, and are optionally preceded by the subordinating particle *ki*. Finite embedded clauses appear obligatorily to the right of the verb, although all noun complements appear to the left (Dayal 1996).

(12) a. Vo jaantii hai [ki anu aayii] [Hindi-Urdu]

She know aux that Anu come

'She knows that Anu came.'

b. *Vo [ki anu aayii] jaantii hai. [Hindi-Urdu]

She that Anu came know aux.

Intended: 'She knows that Anu came.'

(13) Sonia ... tum jaanti ho yeh revolver khaali hai, mein jaanta hun ki yeh

Sonia…. you know aux this revolver empty is, I know aux that this

revolver khaali hai ..... magar police nahin jaanti ki yeh revolver khaali hai

revolver empty is ….. but police not know that this revolver empty is

[Hindi-Urdu]

'Sonia, you know this revolver is empty, I know that this revolver is empty,

but the police don't know that this revolver is empty.'

(Hindi Film, Don, 1978)

There has been much discussion of these clauses in the literature, particularly with respect to their status as complements. It has often been argued that finite clauses appearing to the right of the selecting verb in Hindi-Urdu does not originate in that position, but have somehow been extraposed (Davison 1988) or otherwise adjoined to CP/IP (Dayal 1996). We have examined this issue in some detail above in Chapter 3, however what is clear is that these clauses behave in all respects as though they are c-commanded by material in the matrix clause. For instance, even when a wh-expletive is present in the matrix clause, a quantifier in the matrix clause can bind a pronoun in the embedded clause (Mahajan 2000). If the apparent CP complement is adjoined, this should not be possible since the wh-expletive *kyaa* then occupies the sole complement position associated with the verb. In (14) the quantifier *har aadmii* 'each man' in the matrix clause can bind the pronoun *us-ne* 'he' in the second CP (Mahajan, 2000).

(14) Har a:dmi:-ne$_i$ kya: soca: ki us-ne$_i$ kis-ko dekha:. [Hindi-Urdu]

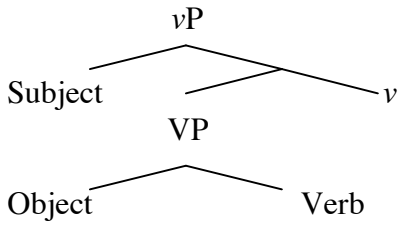
each man-erg expl thinks that he-erg who saw

'Who did every man$_i$ think that he$_i$ saw?'

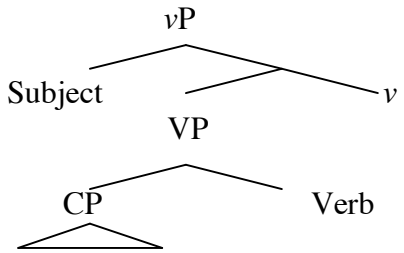
Bayer (1996) discusses this fact for both Bangla and Hindi-Urdu and also notes that while displacement of wh-phrases is possible from complement CPs, it is impossible for CPs that are clearly adjuncts. For these reason and others articulated by Davison (1993), we will follow Bayer in claiming that finite clauses are complements in Hindi-Urdu, and that they exceptionally appear on the right.

Given the facts above, I will assume the phrase structures in (15)-(16) for basic Hindi-Urdu sentences. In the case of (15), both the transitive light verb (*v*) and the VP projection are head final, with the verb taking a noun complement to the left. In the case of (16), we will assume that in the process of post-syntactic linearization, the CP complement must be aligned to the right edge of the vP phase (see Fox and Pesetsky, to appear).

(15)



(16)



When a verb does take a finite complement clause in Hindi-Urdu, as in Kashmiri, it is possible for interrogative phrases originating in the embedded clause to take matrix scope using one of two strategies: the question phrase originating in the embedded clause itself can be displaced, appearing in the main clause, or can remain in its usual position in the embedded clause while the apparently meaningless question word *kyaa* appears in the main clause. These two possibilities were exhibited in (1)-(2), repeated here.

(1) Sita-ne kis-ko socaa ki ravii-ne dekhaa? [Hindi-Urdu]

Sita-erg who-acc thought that Ravi-erg saw

'Who did Sita think that Ravi saw?'

(2) Sita-ne kyaa socaa ki Ravii-ne kis-ko dekhaa? [Hindi-Urdu]

Sita-erg expl thought that Ravi-erg who-acc saw

'Who did Sita think that Ravi saw?'

It is certainly the case that a variety of scrambled word orders are possible in Hindi-Urdu constituent questions, as was described above. For instance, (8), repeated here as (17a), is also possible with the orders (17b-c) (Bhatt 2003b).

(17) a. kitaab-ko kis-ne parhi? [Hindi-Urdu]

book-acc who-erg read

'Who read the book?'

b. kis-ne kitaab-ko parhi [Hindi-Urdu]

wh-erg book-acc read

c. kitaab-ko parhi kis-ne [Hindi-Urdu]

book-acc read who-erg

However, the wh-expletive *kyaa* may appear only in the pre-verbal position where interrogative phrases are canonically located, and cannot be separated from the verb by other constituents (Mahajan 2000). This discrepancy between the wh-expletive and the full wh-phrase will fall out naturally in the account below.

(18) *Sita-ne kyaa abhii abhii socaa ki Ravii-ne kis-ko dekhaa? [Hindi-Urdu]

Sita-erg expl now now thought that Ravi-erg who-acc saw

'Who did Sita think just now that Ravi saw?' (Mahajan 2000)

Returning to Kashmiri, finite complement clauses have identical word order to matrix clauses. I repeat here the two primary types of long-distance wh-dependencies, to exhibit the range of grammatical word orders in both the matrix and and complement clauses.

(19) Mira: **k'a**: chi yatsha:n ki su gotsh anun ____. [Kashmiri]

Mira what aux wants that he should bring

'What does Mira want that he should bring?' (Wali and Koul: 20)

(20) Mira: **k'a:** chi yatsha:n ki **k'a:** su gotsh anun. [Kashmiri]

Mira expl aux wants that what he should bring

'What does Mira want that he should bring?' (Wali and Koul: 20)

(21) tse **kem'** chu-y ba:sa:n ki mohn-as dits __ kita:b [Kashmiri]

you who aux think that Mohan-dat gave book

'Who do you think gave Mohan the book?' (PK 9/21/04)

(22) tse **k'a**: chu-y ba:sa:n ki mohn-as **kem'** dits kita:b [Kashmiri]

you expl aux think that Mohan-dat who gave book

'Who do you think gave Mohan the book?' (Wali and Koul: 18)

Let us now turn to additional data which helps us identify the distinct positions of the tensed verb in clauses in Kashmiri and in Hindi-Urdu. Unlike in Kashmiri, the verb in Hindi-Urdu is clearly not in second position. For instance, sentential negation follows the fronted tensed verb in Kashmiri, but precedes the tensed verb in Hindi-Urdu.

(23) anu-ko hindii-urdu nahiiN aatii hai [Hindi-Urdu]

Anu-dat Hindi-Urdu not come aux

'Anu doesn't know Hindi-Urdu.' (Schmidt)

(24) wo pahle sabzi xor nahin tha [Hindi-Urdu]

he before vegetarian not was

He was not a vegatarian before. (Schmidt)

(25) raath khyav-na larRkav batI [Kashmiri]

yesterday eat-not boys food

'The boys did not eat the food yesterday'. (Bhatt 1999)

Also, in Hindi-Urdu the main and tensed auxiliary verbs must appear finally and adjacent to one another, while in Kashmiri it is the tensed auxiliary verb that takes second position, while the main verb appears in final position.

(26) a. Jaun yeh soctaa hai [Hindi-Urdu]

John that thinks aux

'John thinks that.'

b. *Jaun hai yeh soctaa [Hindi-Urdu]

(27) a. laRk ch-u dohay sku:l gatsh-aan [Kashmiri]

boy aux daily school go-perf

'The boy goes to school every day.' (Bhatt 1999)

b. *laRk dohay skuul gatsh-aan ch-u [Kashmiri] (PK 9/21/04)

2.2 Differences Observed

Let us now turn to make some initial observations about the differences we see between Hindi-Urdu and Kashmiri in the data above. First, though both Hindi-Urdu and Kashmiri are underlyingly verb final (Kashmiri, for instance, features word-final order in non-finite clauses (see Bhatt 1999)), Kashmiri has verb-second word order on the surface. In fact, Kashmiri exhibits so far the typical profile of a verb-second language, familiar from the extensive literature on this language type among European languages. Adopting the account from Chapters 2 and 3, we will assume here that the verb in Kashmiri is in the C head.

Wh-material (wh-expletives and wh-phrases) in both Kashmiri and Hindi-Urdu appears in the preverbal position. We know that in Kashmiri, this position is an A-bar position above the C head, specifically the specifier of CP. However, it is clear from the comparison above that the preverbal position of wh-material in Hindi-Urdu

must be significantly lower than CP. If the verb itself is within *v*P (and we have no reason to assume otherwise), then wh-material is in an A-bar position relatively close to the verb.

At this point, I will introduce a hypothesis that will serve as a major theme for the remainder of this chapter. I will suggest that Hindi-Urdu wh-material is not *in-situ*, as has usually been claimed, but instead in a distinguished position in the overt syntax which is lower than Spec, CP. I will propose that this position is the specifier of *v*P. Among those who have argued that the *v*P phase is a possible stopping-off point for long-distance wh-movement are Rackowski and Richards (2005). I want to begin with this basic intuition: a wide range of facts about Hindi-Urdu and its wh-dependencies can be understood naturally if we take the specifier position of *v*P to be the position in which Hindi-Urdu wh-material appears.

3. The Position of wh-material in Hindi-Urdu

The account of Hindi-Urdu presented in this chapter will argue that wh-material in Hindi-Urdu is not *in-situ*, as is commonly assumed, but instead moves to the position Spec, *v*P. This is a surprising hypothesis; in this section I will offer three kinds of empirical evidence in its support.

The evidence presented below will be comparative in nature, contrasting the structure of Kashmiri with that of Hindi-Urdu. The first type of evidence involves the position of focused constituents in the two languages relative to other constituents. The second set of data involves the position of adverbs that are typically analyzed as

adjoined to *v*P. The third section makes use of relatively new work on sluicing in the two languages, particularly with respect to the well-known property of sluicing that it appears to amnesty island violations.

3.1 Focused Constituents

The grammaticalized focus position in Hindi-Urdu is the position immediately preceding the verb in linear word order.

(28) mai-ne kamre-me [inhii tiin larkon-ko] bheja [Hindi-Urdu]

I -erg room-to [these-foc three boys-acc] sent

I sent these three boys to the room (Butt and King 1996)

(29) kitabeN kal maiN laya tha [Hindi-Urdu]

books yesterday I brought aux

I brought the books yesterday (It is I who brought the books yesterday)

(Kidwai 2000)

There is a thus a fixed syntactic position for focused constituents, whether interrogative or non-interrogative.[4] This word order pattern is consistent with that fixed position being the specifier of *v*P in Hindi-Urdu.

Analogously, in Kashmiri the position for both interrogative and non-interrogative focus is the same. Both types of constituents appear in the preverbal

[4] Hindi-Urdu has very flexible word order, and much work has been done on scrambling in this language. I won't discuss this phenomenon here, but I will assume that after a focused constituent has valued it features in the specifier position of *v*P, it can subsequently scramble (see Bhatt 2003b).

position. As discussed above, the Kashmiri focus particle *–ti* can only appear suffixed to constituents in this immediately preverbal position (all Bhatt, 1999).

(30) hu:n-ti chu behna broNh panin ja:y goD sa:f kara:n [Kashmiri]

dog-foc aux seat before self's place first clean do

'Even the dog cleans his place before sitting.'

(31) *? panin ja:y chu hu:n-ti behna broNh goD sa:f kara:n [Kashmiri]

self's place aux dog-foc seat before first clean do

Intended: 'Even the dog cleans his place before sitting.'

(32) *gary-ti kus chu-na ka:m kara:n [Kashmiri]

home-foc who aux-neg work do

Intended: 'Who doesn't work even at home?'

In Chapter 3 I developed an account in which both interrogative and non-interrogative focused constituents appear in Spec, CP in Kashmiri.

It appears, then, that all focused constituents, whether interrogative or not, preferentially appear preverbally in Hindi-Urdu and Kashmiri. The account which will be presented here attributes the position of focused material in each language to the properties of the phase-defining heads C and *v*, respectively. In effect, what we see is the same clausal topology in Kashmiri and Hindi-Urdu, lower in Hindi-Urdu (on the *v* head) and higher in Kashmiri (on the C head).

3.2 Adverbs

Adverbs that are typically analyzed as adjoining to *v*P, such as 'always', can appear before the verb and the direct object in an unmarked Hindi-Urdu sentence as in (33) (Schmidt 1999). They can also appear before the verb and a wh-phrase, as in (34).

(33) Vo mujhe hamesha cai pilata hai [Hindi-Urdu]

He me always tea drink-cause aux

He always has me drink tea. (Schmidt 1999)

(34) Vo aap-ko hamesha kya pilata hai [Hindi-Urdu]

He you always what drink-cause aux

What does he always have you drink? (Schmidt 1999)

At first glance, these facts are not very informative, because both the direct object and the interrogative pronoun could be in their base positions in (33)-(34).

However, it is also possible for the adverb to follow the question word, as in (35). This suggests that in (35) the wh-word cannot be *in-situ.*

(35) Vo aap-ko kya hamesha pilata hai [Hindi-Urdu]

He you what always drink-cause aux

What does he always have you drink?

These observations fall into place naturally on our proposal. If the positioning requirement for this class of adverbs is that they attach to a complete *v*P, that requirement can be met at two derivational points on our assumptions: either (i) after both specifiers have been introduced, or (ii) after the external argument has been

introduced but before the second specifier (the wh-element) has been merged. The first alternative yields (34), the second (35).

It is unclear how we might account for (35) on the assumption that the wh-phrase is *in-situ*. However, some might object that (35) is simply evidence of scrambling of the wh-phrase out of its (*in-situ*) position as the complement to V, and into some higher position. In essence, this is not substantively different from what I propose here – that the wh-phrase arrives at its surface position in (35) through A-bar movement out of the VP. However, there is another, related fact which indicates that the position to which *kya* 'what' has moved in (35) has some special status.

Chandra (2005), citing Mahajan (1990), observes that an adverb like *jaldise* 'quickly' can appear following the direct object just in case the object receives focal stress.

(36) sita *kam* jaldise karthii thii [Hindi-Urdu]

sita work quickly did aux

'Sita worked quickly.' (Chandra)

If *kam* 'work' does not receive focal stress, (36) is not well-formed, although (37) is possible whether the object is stressed or not.

(37) sita jaldise kam karthii thii. [Hindi-Urdu]

Sita quickly work did aux

'Sita worked quickly.' (Chandra)

Under our account, *kam* 'work' in (36) is a focused phrase appearing in Spec, *v*P. The adverb can be adjoined either before (resulting in (36)) or after (resulting in (37)) the

object has shifted here for the purposes of focus. This suggests that when a wh-phrase or direct object appears to the immediate left of the *v*P-adjoined adverb it has not been simply 'scrambled', but it has been displaced to a specific A-bar position at which it checks focus-related features (interrogative or non-interrogative). I will take this evidence to support the claim that focused constituents (and in particular wh-phrases), appear in an A-bar position on the edge of *v*P (Spec, *v*P) in Hindi-Urdu.

This data contrasts sharply with the facts in Kashmiri, in which these adverbs must follow the second-position verb.

(38) akh´kis sÆ:t' chi hame:shÆ lada:n [Kashmiri]

one another with aux always fight

'(They) always fight with each other.' (Wali and Koul: 133)

(39) tÇmis nishi Çs hameshÆ no:kar [Kashmiri]

He near have always servants

'He always has servants.' (Wali and Koul: 140)

(40) *?tÇmis nishi hameshÆ Çs no:kar [Kashmiri]

he near always have servants

Intended: 'He always has servants'. (JC 2/16/06)

In Kashmiri, the verb is clearly in a position higher than the edge of *v*P (analyzed here as C). Because wh-phrases and focused constituents always appear immediately before the second-position verb (in Spec, CP), they are never found immediately

adjacent to *v*P-adjoined adverbs. No other permutations of these orders are possible in Kashmiri.

3.3 Sluicing in Hindi-Urdu and Kashmiri

Whether or not we would expect to find sluicing at all in Hindi-Urdu depends on the account of sluicing we subscribe to. If sluicing is actually an operation that deletes an IP or TP, and the wh-phrase which remains is located in the specifier of CP, then we would not expect to see sluicing in Hindi-Urdu. Since under every analysis the language has no wh-movement to Spec, CP, sluicing would be the only instance requiring such movement in Hindi-Urdu. If, on the other hand, we view sluicing as simply the deletion of the sister of a wh-phrase, regardless of the position of that phrase, then it would not be surprising to find sluicing in Hindi-Urdu (Grebenyova 2006).

In fact, sluicing is generally possible in Hindi-Urdu, a fact which is widely known and cited in the literature (see for instance Merchant 2001).

(41) Aisha-ne ek ciiz khariidii lekin mujhe nahii pataa (ki) kyaa. [Hindi-Urdu]

Aisha-erg a thing bought but I not know (that) what

Aisha bought something, but I don't know what. (Mahajan 2005)

(42) Khadija jaldii calii jaaegii lekin mujhe nahii pataa (ki) kyoN [Hindi-Urdu]

Khadija early leave go-fut but I not know (that) why

Khadija will leave early, but I don't know why. (Mahajan 2005)

However, in recent work on sluicing, Mahajan (2005) has given us reason to question whether the sluicing we see in Hindi-Urdu is precisely like the sluicing we see in a language like English. It appears that Hindi-Urdu sluicing, unlike sluicing in many other languages, does not have the effect of repairing island violations.

For example, (43a) would be a case of extraction out of a complex NP island. Normally we would expect to see the sluice in the lower clause repair the supposed extraction of *kis-ko* 'who' from the island. However, the construction remains ungrammatical, despite the sluice. Compare the ungrammaticality of (43a-b) to their English translations, or to an equivalent English sluice from a complex NP island in (44). Here the extraction of *who* from the complex NP is grammatical due to the sluice in the lower clause.

<u>Complex NP island[5]:</u>

(43) a. *mE jaanta huN ki Ravii-ne [Salmaa-ki ek larke-ko dii hui kitaab] phar dii
I know aux that Ravi-erg Salma-dat a boy-dat gave aux book tore gave
lekin mujhe nahiiN pataa (ki) kis-ko. [Hindi-Urdu]
but I not know (that) who.
'I know Ravi tore up the book Salma gave to a boy, but I don't know to whom.' (Mahajan 2005)

[5] (43a), (45a), and (47a) are provided by Mahajan, and the grammaticality judgements are confirmed by native-speaker informants, who report a significant contrast between sentences like (41)-(42) and the sentences in (43), (45), and (47). I have also provided in (43b) and (45b) some simpler examples which better mirror the Kashmiri sluicing data below and have more clearly grammatical English translations. Thanks especially to Uzma Rizvi and Shehrose Rehmani for their help with this section.

b. * mE is kahani-ko maantaa hun ki Ram-ne kisi-ko mara, lekin

I this story-acc believe aux that Ram-erg someone-acc killed but

un-ko nahiN pataa (ki) kis-ko. [Hindi-Urdu]

they-erg not know who-acc.

I believe the story that Ram killed someone, but they don't know who.

(44) I believe the claim that he bit someone, but they don't know who. (Ross 1969)

The results are similar when we attempt to repair sluices out of adjunct islands, as in (45)-(46), or out of coordinate structure islands, as in (47)-(48). In each case the Hindi-Urdu extraction is ungrammatical, despite the sluice, while in English the island violation is apparently amnestied by the sluice.

<u>Adjunct island:</u>

(45) a. *Salmaa Raam-ke ek larkii-se milne par bahut naaraaz hogii lekin mujhe

salma ram-dat a girl-with meet upon very angry become but I

nahiiN pataa (ki) kis-se. [Hindi-Urdu]

not know (that) who-with.

'Salma will be angry upon Ram's meeting with a girl, but I don't know with whom.' (Mahajan 2005)

b. ?*Ali qalam laayaa kuch likne ke liye lekin mujhe nahiiN

Ali-erg pen brought something write for but I not

pataa kyaa.

know what.

Ali brought a pen in order to write something, but I don't know what.

(46) Ben will be mad if Abby talks to one of the teachers, but I don't remember which. (Merchant 2001)

<u>Coordinate Structure island:</u>

(47) *Ali-ne Aisha or ek larke-ko saath saath dekhaa, lekin mujhe nahiiN pataa

Ali-erg Aisha and a boy-acc together saw but I not know

(ki) kis-ko. [Hindi-Urdu]

(that) who.

'Ali saw Aisha and a boy together, but I don't know who'. (Mahajan 2005)

(48) Bob ate dinner and saw a movie that night, but he didn't say which.

(Merchant 2001)

These preliminary results indicate that although Hindi-Urdu does permit sluicing, it does not necessarily pattern with English sluicing in repairing apparent island violations. Hindi-Urdu sluicing instead patterns with English VP-ellipsis (VPE), which also does not generally amnesty island violations. Chung, Ladusaw, and McCloskey (1995) offer the following ungrammatical examples of VPE

(49) *What did you leave before they started playing t?

(50) We left before they started playing party games. *What did you leave before they did?

Merchant (2001) also provides the following example:

(51) *They want to hire someone who speaks a Balkan language, but I don't know which they do.

Compare this with the sluiced structure in (52), which is grammatical.

(52) They want to hire someone who speaks a Balkan language, but I don't know which.

Though there are grammatical examples of VPE repairing island violations (see, for instance, Kennedy and Merchant (2000), Fox and Lasnik (2003)), in most VPE cases islands are not repaired, while the analogous sluicing constructions are quite natural. Merchant (2002, to appear) and Fox and Lasnik have offered various forms of the claim that it is the size of the ellipsis site that is crucial in the ability of the ellipsis to amnesty island violations. That is, ellipsis sites of VP size or smaller tend not to repair island violations, while larger, TP-sized ellipsis sites do.

I will not attempt at this point to apply the specific details of either of these proposals to Hindi-Urdu (but see footnote 8), but I want to point to the correlation between ellipsis site size and island repair. Though the results are preliminary, it appears that in Hindi-Urdu, the deletion of the sister of the wh-phrase does not systematically amnesty island violations. The proposal in this chapter has analyzed Hindi-Urdu wh-material as located in the specifier of *v*P. If this were the case, deletion of the sister of the wh-phrase in Hindi-Urdu would leave an ellipsis site of a

size significantly smaller than a full clause (TP/IP). Instead, the ellipsis site size would be similar to that left by English VPE. It would seem, then, that what appears to be sluicing in Hindi-Urdu might actually be quite different from sluicing in English, which would both support the syntactic position posited for wh-material in this account, as well as explain the apparent island violations pointed out by Mahajan for sentences like (43), (45), and (47).

We would expect to see the contrasts between Hindi-Urdu and Kashmiri that have been explored throughout this section borne out in the domain of sluicing as well. Kashmiri is a language that exhibits systematic wh-movement to Spec, CP. If this is the case, the size of the ellipsis site in a Kashmiri sluice should be large enough to amnesty island violations.

Not surprisingly, simple sluices are grammatical in Kashmiri, just as they are in Hindi-Urdu.[6]

(53) Ra:jI rani ka:Nh magar me chu-nI pata: k'a. [Kashmiri]

Raj cook-fut something, but I aux-not know what

'Raj will cook something, but I don't know what.' (JC 2/16/06)

---

[6] Mahajan reports that the complementizer *ki* 'that' can precede the wh-phrase in the grammatical Hindi-Urdu sluices (though this claim is controversial for Hindi-Urdu – see Merchant 2001, whose judgements are attributed to Rajesh Bhatt (p.c.)). It seems that the particle *ki* in Kashmiri cannot appear preceding a sluice (as in English).

(i) Ra:jI rani kaNh, magar me chu-nI pata: *ki k'a.
Raj cook-fut something, but I aux-no know ki what
Intended: 'Raj will cook a dish, but I don't know what.'

(54) kamitani khyav batI, magar me chu-nI pata: k´m. [Kashmiri]

Someone ate food but I aux-not know who

'Someone ate the food, but I don't know who.' (JC 2/16/06)

To the best of my knowledge, sluicing data from Kashmiri has not previously been reported in the literature. The data below was produced and verified by native-speaker Kashmiri informants. Note that multiple informants were in agreement. I attribute all judgements below to JC (2/16/06) and PK (3/31/06). Unlike in Hindi-Urdu, sluices in Kashmiri do appear to repair island violations. For instance, the extraction from what would be a Complex NP island in (55) is completely grammatical, just as in the English example above.

<u>Complex NP island[7]:</u>

(55) Tim ´:s' yi kath za:na:n ki Cheneyan mor kustani, magar me

They aux this story know that Cheney-erg killed someone but I

chu-nI pata: kus. [Kashmiri]

aux-not know who

'They believe the story that Cheney killed someone, but I don't know who.'

In the example in (55), what would seem to be an illicit extraction of *k´m* 'who' out of a complex NP is rendered grammatical by the ellipsis. Note that (53)-(55) provide support for a view of sluicing in which it is not TP which is elided, but instead the

[7] I am indebted to JC for these jugdements and discussion of these and other sluicing data.

sister to the wh-phrase (whatever that may be). This is because Kashmiri is a language exhibiting verb-second order in all finite clauses. If the contents of both the specifier of CP and the contents of the C head were permitted to remain, we would expect to see the second position verb also remaining behind following the sluice. However, just as in English, it is only the wh-phrase which can grammatically remain.

Sluicing also seems to repair what would be violations of a variety of other islands. I offer evidence for repair for adjunct and coordinate structure islands below.

<u>Adjunct Islands</u>

(56) Savitai a:si shararat, a:gar ts´ vuch-a-kh akh sha:gird-av manzI, magar
Savita be-fut angry if you see-2nd one student-abl from, but
me chu-nI pata: kus huh. [Kashmiri]
I aux-not know which one.
Savita will be angry if you see one of the students, but I don't know which one.

(57) Shi:lan on ka:kaz kalam akh ci:z le:khnI kh´trI, magar me chu-nI
Sheila brought paper pen a thing write for but I aux-not
pata: k'a. [Kashmiri]
know what
Sheila brought paper and pen in order to write something, but I don't what.

Coordinate Structure island:

(58) Savita-an on cake beyI onun akh syuun, magar me chu-nI ya:d

Savita-erg brought cake and brought some dish, but I aux-not remember

k'a

what

Savita brought cake and brought some dish, but I don't remember what.

(59) Shirini kar kenh ka:m, beyI vuch-in kenh rishtadar ti, magar ´:s' cha-na

Shirin did some work and saw some relatives part, but we aux-no

pata: kam rishtadar [Kashmiri]

know which relatives.

Shirin did some work and saw some relatives, but we don't know which relatives.

While I will leave the detailed analysis of these facts for another time, I do want to point out that this contrast between Hindi-Urdu and Kashmiri is what we expect to find given the hypothesis suggested above and the idea (Merchant (2002, to appear) and Fox and Lasnik (2003)) that what makes the difference between VPE and sluicing with respect to island violations is the size of the ellipsis site. That is, ellipsis

at the level of the clause produces the amnestying effect, while ellipsis at the level of *v*P does not. [8]

More specifically, if in Hindi-Urdu the wh-phrase moves to (and remains in) the specifier of *v*P, we would expect Hindi-Urdu sluicing to more closely parallel like English VPE. That is, we would expect that island violations will not be regularly repaired by the process of sluicing in Hindi-Urdu. However, if it is the case that Kashmiri, like English, moves wh-phrases to the specifier of CP, sluicing should amnesty island violations in the familiar way. In this view, if the ability of the deletion to repair island violations is a diagnostic of the size of the ellipsis, then the sluicing data from Hindi-Urdu and Kashmiri serve to inform us as to the position of the wh-phrase. If deletion of the sister of the wh-phrase is too small, in some sense, the island violation will not be repaired, as in the case of Hindi-Urdu. If the sister of the wh-phrase is of a larger size (TP-sized or greater), then island violations will be amnestied.

---

[8] It should be noted that the facts we have observed are not compatible with the specifics of Merchant's (to appear) proposal. Merchant suggests that the reason the smaller ellipsis sites (VPE-sized) do not permit island violation repair is that an illicit intermediate trace of wh-movment remains following deletion. In the case of an English sluice, Merchant claims that wh-movement to Spec, CP moves through VP and TP/IP, and leaves a trace in Spec, VP and Spec, IP. If either of the traces of movement remain, Merchant claims the result will be ungrammaticality. IP (clause-sized) ellipsis (sluicing) would remove both of these offending traces, but smaller VPE would leave the trace in Spec, IP.

It is unclear how this approach could be applied to the Hindi-Urdu case. In this chapter, we have proposed that the wh-phrase in Hindi-Urdu appears in Spec, *v*P. If a VP-sized site is elided, under Merchant's view, all offending traces of movement to this position should be eliminated. I do not see how the ungrammaticality of Hindi-Urdu sluicing could be attributed to a trace of wh-movment remaining following ellipsis. However, I do think that the basic intuition expressed in the account of Merchant and others does go some way toward describing the contrast we see between Kashmiri and Hindi-Urdu. That is, the smaller ellipsis-site (VPE-sized) does not systematically repair island violations, while the larger ellipsis-site (clause-sized) does.

We have now seen three types of empirical support for the proposed contrast between Hindi-Urdu and Kashmiri. The data on the interrogative and non-interrogative focus positions in the languages, on *v*P adjoined adverbials, and on sluicing all are consistent with the understanding offered here of the position to which wh-material moves in these two languages. First, the facts in 3.1 and 3.2 would be very surprising indeed if Hindi-Urdu were a true wh-*in-situ* language. Each data set suggests that there is a distinguished position for wh-material in Hindi-Urdu, and that this position is the specifier of *v*P. Second, the comparative facts in 3.1-3.3 highlight the difference in the relative positions of wh-material in Hindi-Urdu and Kashmiri, and strongly suggest that the same clausal topology exists in both languages, with the difference being that the specifier of CP in Kashmiri plays the role which in Hindi-Urdu is played by the specifier of *v*P. The following sections will continue to develop this line of research.

4. A-bar Movement in Hindi-Urdu: Extending the Account of Kashmiri

In this section, I will examine the account of Kashmiri developed in Chapter 3 with an eye toward adapting it to wh-dependencies in Hindi-Urdu. Specifically, we will ask if the view of Kashmiri wh-expletives may properly be extended to Hindi-Urdu. We must also ask whether Hindi-Urdu, often assumed to be a wh-*in-situ* language, actually exhibits wh-movement between clauses as Kashmiri is claimed to do.

4.1 Reviewing the Account of Kashmiri

In the account described in Chapter 3, the role of the wh-expletive is crucial. The distinction between full extraction from subordinate clauses and partial wh-movement in Kashmiri can be analyzed as the distinction between the operations Move and static Agree to satisfy uninterpretable features, just as in the A-system. Heads and wh-phrases will possess sets of interpretable and uninterpretable features. If a wh-expletive happens to be in the numeration, it can merge to satisfy the EPP on a head. Again, just as in the case of the DP-expletive, the wh-expletive is defective in that it has no interpretable features of its own, so the head into whose specifier it has merged must value its uninterpretable features by interacting with an accessible wh-phrase via static Agree over some distance.

Recall the three features at work controlling movement and agreement in the interrogative A-bar system: the EPP (common to the A and A-bar systems), the [Q] feature, and the [wh] feature. Importantly, it is the interpretable [Q] feature that is introduces an unselective binder of (choice) function variables (Reinhart 1998).

The operations Agree and Move are limited to the phase in the A-bar system, just as they are in the A-system. In Chapter 3, we considered only the CP-phase in our account of wh-movement in Kashmiri. However, in this chapter we will adjust this account slightly in order to consider the role that the *v*P-phase plays in wh-movement and wh-expletive constructions. In the phase theory of movement (Chomsky 2000, building on Fox (2000) and Nissenbaum (2000)) the *v* head defines a phase for the purposes of A-movement. That is, an argument within the *v*P may shift to the edge of

the *v*P-phase (Spec, *v*P), and may subsequently interact with higher heads, such as T. Current work is exploring the role of the *v*P-phase in A-bar movement (Rackowski and Richards 2005), and this strand of research will ultimately help to provide an understanding of Hindi-Urdu. For this reason, the analysis of Kashmiri given in this chapter will be more detailed than that in Chapter 3, in that it will address the nature of both the C and *v* heads. From a comparative perspective, we will see that certain properties of the two languages can be attributed specifically to the properties of these two phase-defining heads.

No intermediate probes, whether they are subordinate C heads, or *v* heads, will bear an interpretable [Q] feature. Instead, they will have an uninterpretable [wh] feature and the EPP. This will cause them to be active probes that require the wh-material to appear in their respective specifiers.

In a clause that is embedded, the wh-phrase may raise into the root clause, as in (3), repeated here.

(3) tse kem' chu-y ba:sa:n ki mohn-as dits kita:b [Kashmiri]

you who aux think that Mohan gave book

'Who do you think gave Mohan the book?'

In a sentence like (3), the subordinate C and *v* head will lack an interpretable [Q] feature. A wh-phrase that has raised to the specifier of the subordinate CP will still have uninterpretable features that require valuing. The matrix *v* will be much like the subordinate C, and will cause the wh-phrase to Move into its specifier in the matrix clause. The matrix C head will have an uninterpretable [wh] feature, the EPP, and the

interpretable [Q] feature. As a Probe, it will find the wh-phrase in the specifier of the matrix *v* and will enter into an Agree relation with this wh-phrase and attract it to its specifier. The wh-phrase will raise to the specifier of the matrix CP, and all features will be valued on the wh-phrase and all heads. The result will be full extraction. The particle *ki* is inserted post-syntactically (for more on this, see Chapter 2). This process and the features involved are shown in the stepwise derivation in (60). Beneath each head are the features present on that head.

(60)

a. [*v* . . . wh-XP . . .]
wh-*u* Q-*u*
EPP wh-*i*

b. [wh-XP [*v* . . . ___ . . .]]
Q-*u* wh-*u*
wh-*i* EPP

c. [C . . . [wh-XP [*v* . . . . .]]]
wh-*u* *Q*-*u* wh-*u*
EPP wh-*i* EPP

d. [wh-XP [C . . . [___ [*v* . . . . . .]]]
Q-*u* wh-*u* wh-*u*
wh-*i* EPP EPP

e. [*v* . . . [wh-XP [C . . . [___ [*v* . . . . . .]]]]
wh-*u* *Q*-*u* wh-*u* wh-*u*
EPP wh-*i* EPP EPP

f. [wh-XP [*v* . . . [____ [C . . . [___ [*v* . . . . . .]]]]]
Q-*u* wh-*u* wh-*u* wh-*u*
wh-*i* EPP EPP EPP

g. [C . . . [wh-XP [*v* . . . [____ [C . . . [___ [*v* . . . . . .]]]]]]
wh-*u* *Q*-*u* wh-*u* wh-*u* wh-*u*
Q-*i* wh-*i* EPP EPP EPP
EPP

h. [wh-XP [C . . . [____ [*v* . . . [____ [C . . . [___ [*v* . . . . . .]]]]]]
Q-*u* wh-*u* wh-*u* wh-*u* wh-*u*
wh-*i* Q-*i* EPP EPP EPP
EPP

(60) represents the extraction of the wh-XP and its passage through the specifier positions of each phase-defining head until reaching the matrix CP.

The derivation will not be licit if the uninterpretable feature of the wh-phrase is simply valued by the matrix C head via Agree while the full wh-phrase remains in the specifier of the subordinate CP. This is because, in the absence of a wh-expletive in the numeration, the EPP property of higher probes (*v* and C) will not be satisfied.

Let us now turn to a grammatical instance of partial wh-movement in Kashmiri as in (4), repeated here.

(4) tse k'a: chu-y ba:sa:n ki mohn-as kem' dits kita:b [Kahsmiri]

you expl aux think that Mohan who gave book

'Who do you think gave Mohan the book?'

In a partial movement construction, the subordinate C head will have only uninterpretable features: an uninterpretable [wh] feature, an uninterpretable [Q] feature, and the [EPP] property. The numeration happens to contain the wh-expletive *k'a:*. This expletive differs from a full wh-phrase in its feature content. As an expletive, it consists entirely of uninterpretable features, and contributes to the syntactic computation only an (uninterpretable) interrogative feature (written Q-*u* here). In Chapter 3 we offered evidence supporting the claim that the wh-expletive originates in a position inside the *v*P in which it receives case (Dayal 1996, Simpson 2000). It can then move through the specifiers of the phase-defining heads until its uninterpretable features are valued. Of course, though the EPP is satisfied on both the matrix *v* and C heads by the wh-expletive in the derivation schematized in (61), their

wh-features (uninterpretable) must still be valued via Agree. In this case the untinterpretable [wh] feature on *v* interacts with the interpretable [wh] feature on the full wh-phrase at the edge of the embedded CP. In turn, the uninterpretable [wh] feature on the matrix C is valued by interacting with the (now valued) [wh] feature on *v*, which is available because it has not yet transitioned to the interface for deletion.[9] This is much like the analysis given to three-clause wh-expletive constructions in Kashmiri featured in Chapter 3. In this way, all uninterpretable features are valued, and wh-expletives in the A-bar system serve the same purpose as expletives in the A-system – to satisfy the EPP and permit the head's features to be valued statically by some other accessible element (Simpson 2000). This process is diagramed in (61).

(61)

a. [*v* . . . wh-XP . . .]
wh-*u* Q-*u*
EPP wh-*i*

b. [wh-XP [*v* . . . ___ . . .]]
Q-*u* wh-*u*
wh-*i* EPP

c. [C . . . [wh-XP [*v* . . . . .]]]
wh-*u* *Q-u* wh-*u*
Q-*u* wh-*i* EPP
EPP

d. [wh-XP [C . . . [___ [*v* . . . . . .]]]
Q-u wh-*u* wh-*u*
wh-*i* Q-*u* EPP
EPP

e. [*v* . . . [wh-XP [C . . . [___ [*v* . . . . . .]]]]
wh-*u* *Q-u* wh-*u* wh-*u*
EPP wh-*i* Q-*u* EPP
EPP

[9] Permitting a Probe to interact with a Goal whose features have already been valued requires that we assume that Agree can take place when the Goal is inactive (though not Move). For further perspectives on this see Pesetsky and Torrego (2004) and Boeckx (2003).

f. [wh-expl [*v* . . . [wh-XP [C . . . [___ [*v* . . . . . .]]]]]
Q-u wh-*u* Q-*u* wh-*u* wh-*u*
EPP wh-*i* Q-*u* EPP
EPP

g. [C . . . [wh-expl [*v* . . . [wh-XP [C . . . [___ [*v* . . . . . .]]]]]]
wh-*u* *Q-u* wh-*u* Q-*u* wh-*u* wh-*u*
Q-*i* EPP wh-*i* Q-*u* EPP
EPP EPP

h. [wh-expl [C . . . [____ [*v* . . .[wh-XP [C . . . [___ [*v* . . . . . .]]]]]]
Q-u wh-*u* wh-*u* Q-*u* wh-*u* wh-*u*
Q-*i* EPP wh-*i* Q-*u* EPP
EPP EPP

It may seem, in (60) and (61) above, that the *v*P phases are irrelevant to the discussion here, and that they are doing little to no work in the analysis. For instance, we never see wh-material remain in the specificer of *v*P in Kashmiri. This is due to two factors. First, the C head in Kashmiri always has the EPP (wh-material always appears preceding the second position verb). In a construction with no wh-expletive, the wh-phrase which has moved into matrix Spec, vP must then interact with the matrix C-head to have its uninterpretable [Q]-feature valued. This interaction will trigger an application of of Move, since C always has the EPP in Kashmiri.

The second factor that prevents wh-material from appearing in Spec, *v*P in Kashmiri is our understanding that the wh-expletive originates in a position below v to which it can be assigned case. In the matrix clause of a two-clause wh-expletive construction in Kashmiri, the wh-expletive will first move into Spec, *v*P to value the EPP on *v*, and must then move to Spec, CP for reasons discussed above. Because this is so, the partially-moved full wh-phrase on the edge of the embedded clause will

never have the opportunity to move into Spec, vP (the wh-expl has already valued the EPP on v). This derives the anti-locality facts discussed in Ch 3.

So it is the fact that the C-head always has the EPP in Kashmiri and the position at which the the wh-expletive is first merged into the derivaiton that consipire to ensure that wh-material never remains in Spec, *v*P in Kashmiri, though it does move through this position**.** Interestingly, the primary thrust of the analysis to follow is that in Hindi-Urdu it is in the intermediate *v*Ps that we see partially-moved wh-phrases.

In summary, both full wh-movement and wh-expletive constructions arise due to featural interactions between phase-defining heads and wh-material. These interactions take the form either of Move or of Agree. Because C in Kashmiri must always have the EPP property, there must always be wh-material (either a full wh-phrase or a wh-expletive) in the specifier of CP in a constituent question in Kashmiri. Since in wh-expletive constructions there is partial movement of the full wh-phrase to an embedded Spec, CP the features of that full wh-phrase are then available to probes in the matrix clause. And more generally, the twin possibilities of partial and full wh-movement reflect the presence or absence of wh-expletives in the array of lexical resources out of which a given question is constructed. Let us keep these aspects of this account in mind as we turn to apply it to similar constructions in Hindi-Urdu.

4.2 Extending this Account to Hindi-Urdu

Hindi-Urdu has traditionally been viewed as a wh-*in-situ* language. That is, wh-material was not understood to occupy a distinguished position in the overt syntax. Previous approaches to Hindi-Urdu and similar languages (Dayal 1996, Mahajan 1990) proposed that wh-material be licensed and raise to a position of interpretation at the level of logical form, via movement to the specifier of CP. This movement supposedly occurs covertly, following the operations in the syntax, and is therefore never visible in the surface representation. LF movement is one solution to the problem of wh-*in-situ*, in that the same licensing that takes place in some languages by overt movement must take place in another way in Hindi-Urdu.

An alternative to this view presents itself in more recent work in the form of the operation Agree (Chomsky 2000, see also Bhatt 2003a). In this view, the Spec-head relation is no longer a central one, nor is it the way in which features are checked or valued. There is instead a mechanism in narrow syntax that can permit heads to interact with wh-material over a distance. This would mean that linguistic material that previously was required by the theory to move covertly to some position can now interact with higher probes in a different way, via the Agree operation. Agree permits the features of a wh-phrase to be valued under local command, instead of forcing movement to Spec, CP. If it is universally the case that C bears the interpretable [Q] feature in a main clause question, this head need not necessarily possess the EPP property. Unlike in Kashmiri, in which the wh-word must move to a C-peripheral position, in Hindi-Urdu it need not. In this case we could assume that the

uninterpretable [wh] feature on the C head in a single-clause question is valued in the local command configuration. The uninterpretable [Q] feature on the wh-phrase is simultaneously valued by its interpretable counterpart on C in this interaction. The end result is that the wh-phrase need not and must not move into the specifier of CP since all features are valued via Agree. This operation for the sentence in (9), repeated here, is represented by the schematic in (62).

(9) Hamid-ne kyaa pa<u>r</u>a? [Hindi-Urdu]

Hamid-erg what read

'What did Hamid read?'

(62) $[_{CP}$ C $[_{TP}$ Hamid-ne kyaa para]]
[Q-*i*] [Q-*u*]
[wh-*u*] [wh-*i*]

Further, if wh-phrases have the semantics of indefinites unselectively bound by interrogative operators, there is no reason to think that they must raise, for interpretive reasons, to a "scope position". I will continue to use that term here (since it is convenient and well-established) but this is purely for exposition and convenience.

If we can accept that there is no need, in general, for wh-phrases to move into interrogative Spec, CP in Hindi-Urdu, we have come some significant way toward understanding how question formation functions in the language. We have also identified a primary source of variation that distinguishes Hindi-Urdu and Kashmiri. Interrogative C lacks the EPP property in Hindi-Urdu, but bears the EPP property in

Kashmiri. Therefore Kashmiri will exhibit wh-movement to the left edge of the clause but Hindi-Urdu will not.

However this apparent advance immediately creates a puzzle for the case of wh-phrases taking scope over more than one clause. Consider again the Hindi-Urdu wh-expletive construction in example (2), repeated here.

(2) Sita-ne kyaa socaa ki Ravii-ne kis-ko dekhaa? [Hindi-Urdu]

Sita-erg expl thought that Ravi-erg who-acc saw

'Who did Sita think that Ravi saw?'

In the case of (2), how is it that the features on the matrix C get valued? It is clear that the wh-phrase in the lower clause remains clause-internal, just as the wh-phrase does in (9), and does not move to a left-peripheral position in the clause. If this is the case, the features of the wh-phrase *kis-ko* are inaccessible to the matrix C Probe, because they are not contained within its phase, or even on the edge of the immediately preceding phase. Recall the Kashmiri example in which it is the interpretable [wh] feature on the wh-phrase in the embedded Spec, CP which values the uninterpretable [wh] features in the matrix clause. The full wh-phrase in (2) is too deeply embedded for this to be possible, because two phase boundaries intervene (that defined by the matrix *v* head and that defined by the subordinate C head). The account of wh-*in-situ* as Agree immediately encounters a challenge here.

If the wh-expletive *kyaa* found in Hindi-Urdu is assumed to have the same characteristics as the wh-expletive found in Kashmiri, it also cannot be the wh-expletive in the matrix clause in (2) which values the features on the matrix C. The

wh-expletive has only a single, uninterpretable [Q] feature, and no [wh] feature at all. It can only function to satisfy the EPP on a head. So the question remains: how do the features on the full wh-phrase and the features on the matrix C head get valued?

There is a deeper question here as well, which concerns the very nature of the wh-expletive. In the case of Kashmiri, the wh-expletive serves the role of satisfying the EPP on the phase defining, [Q]-bearing head. However, it can hardly be the case that the wh-expletive in (2) satisfies the EPP on the matrix C head – it seems instead to be occupying the same preverbal position that full wh-phrases in Hindi-Urdu occupy. If it can neither value the features on the C head, nor satisfy the EPP on that head, the wh-expletive seems to have no purpose whatsoever.

Finally, there seems to be no wh-material of any kind in the intermediate Spec, CP. We have already determined that the full wh-phrase in the lower clause is not as high as the clause edge. The wh-expletive certainly does not occupy this position either, as it is clearly located in the matrix clause. We can conclude that there must be no wh-related EPP in the subordinate C head, otherwise there would have to either be wh-material in Spec, CP in the final representation, or wh-material would have to have moved through this position. A close examination of (2) reveals that neither of these has occurred. The Kashmiri wh-expletive construction has also been termed "partial movement", because the full wh-phrase in the lower clause has clearly moved from its base position to the left periphery of that clause. However, it is unclear in what sense (2) is a case of partial movement in Hindi-Urdu. Certainly the wh-phrase in the embedded clause has not moved to the clause edge.

This accumulation of questions suggests that the analysis reviewed in section 4.1 for Kashmiri cannot be extended to Hindi-Urdu, or certainly not without adjustment. That analysis was developed for a language that requires wh-phrases and wh-expletives to move into Spec, CP, so that they are always located at the phase-edge and satisfy the EPP on the C head. Kashmiri is just such a language, but Hindi-Urdu is clearly not. In what follows, we will combine the observations from this section with the evidence examined in section 3, which indicated that Hindi-Urdu wh-material can be understood as located in the specifier position of *v*P.

5. Wh-dependencies in Hindi-Urdu: the *v*P phase

Each of the questions in section 4.2 points to the same conclusion. The approach presented in 4.1 is designed to account for a language in which the position Spec, CP is crucial for wh-material. In order to understand Hindi-Urdu wh-dependencies, it seems that we need to design an account in which it is some other position in which wh-phrases and wh-expletives preferentially appear. As suggested in sections 2-3 above, I will pursue an account of Hindi-Urdu according to which wh-material is not actually *in-situ*, but instead in a designated position in the overt syntax which is lower than Spec, CP, specifically, Spec, *v*P.

5.1 Wh-Movement in Tagalog: A Case for [Q]-bearing *v*

Among those who have argued that the *v*P phase is a possible stopping-off point for long-distance wh-movement are Rackowski and Richards (2005). They offer

supporting evidence for the phase theory of movement by claiming that the *v*P is actually the only phase relevant to successive-cyclic wh-movement cross-linguistically.

Rackowski and Richards show that specific arguments in Tagalog must move to the specificer of *v*P to receive the appropriate semantic interpretation. Evidence for this claim includes overt morphology on the verb indicating agreement in case with the shifted argument. Arguments which must shift include wh-phrases. Rackowski and Richards compare this shift to object shift in Germanic languages.

On this interpretation, in the Tagalog sentence in (63) the verb shows agreement in case with the shifted wh-word (all Tagalog data from Rackowski and Richards).

(63) Sino [ang binigy-**an** ng lalaki ng bulaklak __]? [Tagalog]

who ANG gave-**Dat** CS man CS flower

'Who did the man give the flower to?'

The verb can also agree with a CP complement.

(64) Sasabih-**in** ng kalabaw na masarap ang bulaklak [Tagalog]

will.say-**Acc** CSwater-buffalo that delicious ANG flower

'A/the water buffalo will say that the flower is delicious'

In the case of long-distance wh-dependencies, in order for extraction to be possible from the embedded clause, the verb must agree with that clause.

(65) Kailan [sasabih-**in** ng sundalo [na uuwi ang Pangulo e]]

[Tagalog]

when will.say-**Acc** CS soldier that nom-willgohome ANG President

'When will the soldier say that the President will go home?'

(66) *Kailan [**m**agsasabi ang sundalo [na uuwi ang Pangulo e]]

[Tagalog]

when **nom**will.say ANG soldier that Nom-willgohome ANG President

'When will the soldier say that the President will go home?'

This fact leads Rackowski and Richards to propose that both *v* and interrogative C have features that must be valued in the process of wh-movement, and may also possess the EPP property, causing Goals to appear in their specifiers. We have seen this account at work in the re-examination of the analysis of Kashmiri wh-movement and wh-expletive constructions in 4.1 above. However, in a significant break with previous approaches, Rackowski and Richards claim that it is unnecessary for non-interrogative C to Agree with any wh-material at all. For example, for a long-distance wh-extraction in English such as in (67), Rackowski and Richards propose the derivation in (68).[10]

(67) Who did you say Bush appointed __?

(68) a. [$C_{[Q]}$ [v [C [v who]]]]

[10] Rackowski and Richards (2005) use different names for the features involved in wh-movement. I will be consistent and use the feature names presented in section 4.1 and Chapters 2-3. The outcomes are equivalent.

b. the matrix *v* agrees with the embedded C

c. [$C_{[Q]}$ [v [C [who v ~~who~~]]]]

d. [$C_{[Q]}$ [who v [C [~~who~~ v ~~who~~]]]]

In (68) *who* moves first to the specifier of the lower *v*. After the matrix *v* has agreed with the embedded C, the embedded C phase is transparent to the matrix *v* probe (we will consider this in more detail below). This probe can then find and interact with *who*, which moves into the matrix clause. In the final step of the derivation, *who* moves into the specifier of the interrogative matrix C.

It is crucial for Rackowski and Richards to assume that once a probe P has Agreed with a goal G, P can ignore G for the rest of the derivation. Though they point us to other resources (see Richards 1998, Hiraiwa 2001), for detailed explanation, they claim that once the Copy Theory of movement is assumed, such a notion will be required. This is because a probe in a multiple wh-question in a language like English will need to look beyond the copy of the (now) moved wh-phrase to the unmoved one. I will not examine this "transparent goal" claim in more detail at this juncture, though there is further discussion of it below.

Crucially for our discussion here, Rackowski and Richards acknowledge that languages which show evidence for wh-related morphology on intermediate C, or wh-expletives in the specifier of intermediate C, will prove challenging for this view. If non-interrogative CPs are not phases which force movement through their specifiers, we require alternative explanations of these facts. For instance, the morphology of

Irish complementizers has been analyzed as indicative of successive cyclic wh-movement through Spec, CP (McCloskey 1990, 2001). Complementizers exhibit a distinguished form if an A-bar binding relation (actually a movement relation) holds between a position within the CP they head and a position external to that CP. Rackowski and Richards claim this morphology simply indicates that the C in question has entered into an Agree relation with a [Q]-bearing *v*, which has in turn agreed with a wh-phrase.

An approach which depends solely on the specifier of *v*P as a stopping off point for long-distance wh-movement will also encounter difficulty in accounting for wh-expletives and instances of partial movement such as those found in languages like German and Kashmiri. Rackowski and Richards acknowledge that in order to understand why wh-material would seem to appear overtly in the specifier of a non-interrogative CP in these languages, they would need to assume a version of the indirect dependency approach. That is, they would need to claim that any CP in which a full wh-phrase or wh-expletive appears is in fact an interrogative CP, and that the scope properties of that wh-phrase are a result of a complex process of coindexation and/or covert clausal pied-piping. We have already seen here (in Chapter 2 and section 4), as well as in detail in Manetta (2005a), Beck and Berman (2000), Fanselow and Mahajan (2000), and Bayer (1996), reasons to wonder whether the indirect dependency approach is tenable for languages like Kashmiri and Hindi-Urdu. For instance, there seems to be clear evidence from phenomena that depend on command that the embedded CP in partial wh-movement constructions is a

complement rather than an adjunct (Bayer 1996, Simpson and Battarcharya 2003). If the indirect dependency approach does not afford an understanding of wh-expletive constructions or partial movement, then Rackowski and Richards' approach becomes problematic for languages like Kashmiri.

However, there is room for an alternative position that could admit Rackowski and Richards approach to Tagalog without sacrificing our current understanding of languages like Irish and Kashmiri. As I will discuss below, a view in which non-interrogative CPs are relevant for wh-movement in some languages and non-interrogative *v*Ps are relevant in others seems more supportable than either extreme in light of the evidence. In what follows, I will propose that for some languages (Kashmiri, Irish), the specifiers of non-interrogative CPs are wh-positions and are stopping points for wh-phrases and wh-expletives. For other languages, the specifier of *v*P is the only stopping off point for long wh-movement. If this is the case, we might expect to see some languages in which wh-phrases and wh-expletives can remain in Spec, *v*P. I will claim here that Hindi-Urdu is just such a language. Insofar as this line of reasoning is correct, we should be able to understand the contrasts between Hindi-Urdu and Kashmiri.

5.2 Hindi-Urdu Revisited

At this point we will bring together two strands of research. The first is the account of partial movement developed here, built on the intuition that these systems are shaped by the same mechanisms that shape the A-movement system. The second

is Rackowski and Richards' account of Tagalog, built on the intuition that the specifier position of the *v*P may play the role traditionally ascribed to the specifier position of CP. Let us now pursue a detailed analysis of the Hindi-Urdu wh-expletive construction in (2) combining these two approaches.

The first step is to assume that a full wh-phrase and a wh-expletive in Hindi-Urdu will possess exactly the same features as they do in Kashmiri. Specifically, as before, we assume that a wh-phrase will have an uninterpretable [Q] feature and an interpretable [wh] feature (which provides its interpretation as a choice function variable). A wh-expletive will have only an uninterpretable [Q] feature, but no interpretable features at all.

The sequence of heads which determines a long-distance wh-dependency in Hindi-Urdu will have features identical to those in Kashmiri. The primary difference will be that the heads themselves are not C heads but *v* heads. This will cause the wh-material to appear not in the specifier positions of CP, but in the specifier positions of *v*P. In the matrix clause the interrogative C head will bear an interpretable [Q] feature and an uninterpretable [wh] feature, just as it does in single-clause Hindi-Urdu sentences, and in Kashmiri. In this way, the [Q-*i*] will be interpreted by the semantics as an unselective binder of choice functions. This matrix C head, however, will not have the EPP property. As a consequence, no wh-material is found in the specifier of CP in Hindi-Urdu. Such elements are found, rather, in lower positions. This means that the features on the C head will then be valued by virtue of its relation with a wh-phrase in some accessible position within its domain (its own phase or at the edge of

the immediately lower phase). In this case, this phase will be the *v*P phase in the matrix clause. The schema in (69) illustrates the featural content of each of the relevant heads in the wh-expletive construction in (2), as well as the features of the wh-material. A detailed explanation of the derivation follows.

(69)

a. [*v* . . . wh-XP . . .]
wh-*u* Q-*u*
EPP wh-*i*

b. [wh-XP [*v* . . . ___ . . .]]
Q-*u* wh-*u*
wh-*i* EPP

c. [C . . . [wh-XP [*v* . . . ___. . .]]]
Q-*u* wh-*u*
wh-*i* EPP

d. [*v* . . wh-expl .. [C . . . [wh-XP [*v* . . . ___ . . .]]]]
wh-*u* Q-*u* Q-*u* wh-*u*
Q-*u* wh-*i* EPP
EPP

e. [wh-expl [*v* . . ____ ... [C . . . [wh-XP [*v* . . . ___ . . .]]]]]
Q-*u* wh-*u* wh-*i* wh-*u*
Q-*u* Q-*u* EPP
EPP

f. matrix *v* agrees with the embedded C

g. [C . . . [wh-expl [*v* . . . ____ [C . . . [wh-XP [*v* . . . ___ . . .]]]]]]
wh-*u* Q-*u* wh-*u* wh-*i* wh-*u*
Q-*i* Q-*u* Q-*u* EPP
EPP

h. [C . . . [wh-expl [*v* . . . ____ [C . . . [wh-XP [*v* . . . ___ . . .]]]]]]]
wh-*u* Q-*u* wh-*u* wh-*i* wh-*u*
Q-*i* Q-*u* *Q*-*u* EPP
EPP

There are two different types of heads that bear features relevant to wh-movement in (69). The first is *v*, which bears wh-features in both the subordinate and matrix clauses. The second is C, which only bears features relevant to wh-movement in the

main clause. Non-interrogative C heads do not bear a [wh] feature in Hindi-Urdu under this analysis.

Beginning in the lower clause, the wh-phrase *kis-ko* originates in object position. The *v* head in the lower clause is a probe possessing the EPP property, and interacts with and raises *kis-ko* into its second specifier position. In this interaction the [wh-*u*] feature on *v* is valued, as is the EPP. Note that this *v* head has no relevant interpretable features – the wh-phrase in its specifier will not be interpreted in this position.

If this account of Hindi-Urdu were precisely like the account of Kashmiri, we would expect the subordinate C head to have an uninterpretable [wh] feature and the EPP. That is, it would be one in the sequence of C and *v* probes interacting with the full wh-phrase. However, under the approach of Rackowski and Richards (2005), non-interrogative C heads have no wh-related features at all and they do not participate in wh-movement. In the hybrid view adopted here, though Kashmiri is a language in which non-interrogative C heads do have wh-features, Hindi-Urdu is a language in which they do not. This point will be discussed in more detail below.

Moving up in the structure, the next head with relevant features is the *v* head in the matrix clause. We have not yet determined the position at which the wh-expletive *kyaa* originates in the tree; that is, whether it is first merged into Spec, *v*P, or originates somewhere in the domain of *v*. This question will be addressed further below. Suffice it to say at this point that the matrix *v* first interacts with the wh-expletive. This interaction values the [Q-*u*] on both the Probe and Goal, as well as the

EPP property on *v*. However the [wh-*u*] on the *v* head remains unvalued and so the head continues to act as a probe.

Following the approach of Rackowski and Richards (2005), the matrix *v* head must have some features that require 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 morpho-phonological 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.

The matrix *v* must then probe to the edge of the lower *v* phase, finding the wh-phrase *kis-ko* in its specifier. In interacting with the wh-phrase, the *v* head values its [wh-*u*] and simultaneously the [Q-*u*] on the wh-phrase is valued. At this point all of the features on the matrix *v* have been valued. In the interpretive component, as before, the C bearing interpretable [Q] triggers the introduction of the unselective binding operator which binds the choice-function variable of the wh-phrase.

The only remaining head with unvalued features is the matrix C, which probes its domain up to the edge of the lower phase, or the matrix *v*. It values its [wh-*u*] feature with the [wh] feature on the matrix *v*. Now the derivation is complete and licit, with all features valued and no unvalued features remaining on any wh-material or in any head.

Let us take a moment here to recall the basic intuition that this account is attempting to capture. We have seen a body of empirical evidence that suggests that the surface position for wh-material in Kashmiri is the specifier position of CP, but

that in Hindi-Urdu such material occupies the specifier position of *v*P. We have also seen that in Hindi-Urdu, unlike in Kashmiri, wh-material can never appear in the specifier position of any intermediate CP, and we have no specific morphological evidence that it has ever appeared in this position. Fundamentally, what we wish to claim is that the same clausal topology surfaces in both Kashmiri and Hindi-Urdu, but at the CP layer in Kashmiri, and the *v*P layer in Hindi-Urdu.

As I see it, there are a number of ways of capturing this in the framework of the Minimalist Program. One possible view is that in Hindi-Urdu, CP is not a phase at all. This view would claim that phases are parametrized across languages, and defined in some language-specific manner. Of course, this approach would require a major rethinking of our understanding of phases. To this point phases (CP, *v*P, DP), have been identified as universal processing units of the derivation. It is unclear what it would mean to say that in a certain language, a certain projection did not constitute a phase, or possibly only constituted a phase for one type of movement but not another.

A second possible way to capture the basic intuition described above is to assume that wh-material moves through the specifier of every phase-defining head (C and *v*) on its way to its ultimate position, regardless of language. In this view, the fundamental distinction between Hindi-Urdu and Kashmiri is that the surface position of wh-material in Hindi-Urdu is in the specifier position of *v*P, and in Kashmiri is in the specifier position of CP. Regardless of whether there is any overt evidence that wh-material moves through the specifier of a head, we would assume that it must do

so as part of the fundamental design of phases. The disadvantage of this approach is, for those languages that have no evidence of wh-material ever appearing in the specifier position of CP, we must posit that it does pass through these positions. On the other hand, the advantage of this view is a certain uniformity in the understanding of A-bar movement across different language types. Every C and *v* head in every language possesses features relevant to wh-movement. What primarily varies by language is which of these heads possess the EPP property.

A third possible way to capture the *v*/C distinction between Hindi-Urdu and Kashmiri is to adopt the view advocated in Rackowski and Richards (2005). Specifically, when a Probe agrees with a phase head, it can probe beyond that phase boundary to the edge of the next lower phase. This approach requires us to assume that completed phases are not immediately exported to the interfaces (Chomsky 2004), but instead remain in the workspace throughout the derivation. Features in lower phases must be available for interaction with Probes in phases higher than the next immediate phase, in the special case that these probes have agreed with intervening phase-defining heads. However, this approach does appear to account for the facts of wh-movement in Tagalog and Hindi-Urdu. It also offers a way of understanding why no wh-material ever appears in the specifier position of intermediate CPs in Hindi-Urdu. This is because the C head never bears any features relevant to wh-movment in Hindi-Urdu under this account, and therefore no wh-material ever moves to its specifier.

At this point, I see no clear empirical test for distinguishing among these proposals. For the sake of familiarity I will adopt the Rackowski and Richards view. In addition, the basic claim Rackowski and Richards make (that the *v*P-phase can play a crucial role in wh-processes) correlates well with our empirical observations about Hindi-Urdu. Our analysis could be expressed under any of the sets of assumptions above (and possibly others). What is crucial is that the basic intuition about the relative clausal organization of A-bar movement in Hindi-Urdu and Kashmiri is captured.

Returning specifically to the schematic in (69), one can observe that for this account to work, the two *v* heads (matrix and embedded) have different sets of features. The matrix *v* has a [Q] feature, while the lower *v* does not.[11] The matrix *v* Probe finds a matching Goal, the wh-expletive, with a [Q] feature. The wh-expletive raises into the specifier position of matrix *v* and is rendered inactive. An additional effect of the presence of a [Q] feature on matrix *v* is to value the uninterpretable [Q] feature on the full wh-phrase in the specifier of the lower *v*P. This possibility is expected, given that we have adopted Rackowski and Richards' position that once a probe has agreed with a lower goal which is a head (the intermediate C head, in this case), then it may penetrate any phase defined by that head. The analysis in (69)

[11] We should note here, as is mentioned in Chapter 3, that there is evidence that intermediate heads in a wh-movement sequence should have interrogative features of some kind. In particular Henry (1995) observes that in Belfast English subject-auxiliary inversion takes place not only in highest C head in a wh-movement sequence, but in intermediate heads as well. She takes this as an indication that at least in this context these intermediate heads share the interrogative status of the matrix C-head. In the present account, this notion is reflected in the [Q] which appears on intermediate C and *v* heads in Kashmiri and on *v* heads in Hindi-Urdu.

ensures that the wh-phrase in the lower clause is 'frozen' in the specifier of the embedded *v*P.

This account of wh-expletive constructions in Hindi-Urdu offers answers to a number of questions that up until now remained mysterious. First and foremost, the combined strands of research have suggested that the syntactic position of wh-phrases and wh-expletives which preferentially appear preverbally is the specifier of *v*P. In section 3 I presented a range of empirical support for this claim, and we have seen here that the claim is well-integrated into an account of long-distance wh-dependencies in the language.

Second, we have now determined how exactly the features on probes in the matrix clause of a Hindi-Urdu wh-expletive construction get valued. The features of the full wh-phrase which remains in the lower clause in a sentence like (2) are in fact accessible to these probes, because the wh-phrase is located on the edge of the relevant phase: the *v*P phase. Since the matrix *v* head may ignore the intermediate CP phase following agreement with it, it can probe and interact with the full wh-phrase without difficulty. The matrix C head can in turn interact with the matrix *v* head. In this way the features of the full wh-phrase are accessible and ultimately result in the valuing of features on the interrogative C.

Third, we can address why no wh-phrases or wh-expletives are ever found in the specifiers of non-interrogative CPs in Hindi-Urdu, nor is there evidence that wh-material has moved through these positions. Under the view we have adopted from Rackowski and Richards, this is due to the fact that non-interrogative CPs in Hindi-

Urdu are irrelevant for wh-movement. Non-interrogative C heads possess no wh-related features, and therefore do not interact with wh-material at all (though see above for other ways to implement this view).

Fourth, and perhaps most importantly, this account provides a way of understanding the role of the wh-expletive in Hindi-Urdu. The wh-expletive serves to value the EPP on the attracting head for long wh-movement, and in Hindi-Urdu this happens to be the *v* head. The features on that head may then be valued via static Agree with some other goal in its domain. Thus, wh-expletives in Hindi-Urdu and Kashmiri can be viewed as being governed by exactly the same set of mechanisms.

Finally, Hindi-Urdu can now be considered a language that exhibits a true case of 'partial' wh-movement, in that the full wh-phrase in the lower clause moves from its base position into the specifier of the embedded *v*P. This resolves a puzzling asymmetry between languages that have wh-expletives and those languages that exhibit the phenomenon called 'partial' movement. Now Hindi-Urdu can be understood to fall under both of these categories.

This approach, as well as the data in (1)-(2), raises the following question: what would happen if that wh-phrase in the lower clause were not frozen in place, but instead was forced to continue raising to have its features satisfied? What if no wh-expletive was available in the numeration to satisfy the EPP property on the higher *v*? The result would be displacement, an additional strategy for forming a long-distance wh-dependency in Hindi-Urdu.

Let us reconsider (1), repeated here.

(1) Sita-ne kis-ko socaa ki Ravii-ne ____ dekhaa? [Hindi-Urdu]

Sita-erg who-acc thought that Ravi-erg saw

'Who did Sita think that Ravi saw?'

Recall that (1) is nearly identical to the wh-expletive construction in (2), except that the wh-word that originates in the lower clause appears in the matrix clause. The interpretations of (1) and (2) are generally viewed to be the same. There is significant debate in the literature about the nature of (and even the grammaticality of) the displacement in (1). Some researchers contend that this is just another example of wh-movement (Simpson 2000, Simpson and Battarcharya 2003), familiar from a wide range of languages. Others (Dayal, 1996) contend that this is a form of scrambling and entirely unrelated to either the wh-expletive construction in (2), or to the usual form of wh-movement visible in other languages.

Given the current theoretical framework, there is a way of viewing (1) in which it is less relevant what name is given to the movement. If, as is claimed by Chomsky (2004), all movement is driven by the interaction of features and/or by the EPP property, then movement in both a full wh-movement language like Kashmiri and in a so-called wh-scrambling language like Hindi-Urdu must be driven by the same basic mechanisms. The upshot of this notion is that the same features that induce wh-expletive constructions might also induce displacement of the wh-word in (1). Whether this displacement is in fact termed scrambling or movement matters little.

The tree in (70) depicts the feature bundles that would appear on each head in the long-distance wh-displacement in (1). Of course, the wh-phrase *kis-ko* has the same features it has in a sentence like (2)/(69). Again, a detailed explanation of the derivation follows.

(70)

a. [*v* . . . wh-XP . . .]
wh-*u* Q-*u*
EPP wh-*i*

b. [wh-XP [*v* . . . ___ . . .]]
Q-*u* wh-*u*
wh-*i* EPP

c. [C . . . [wh-XP [*v* . . . ___. . .]]]
Q-*u* wh-*u*
wh-*i* EPP

d. [*v* . . . [C . . . [wh-XP [*v* . . . ___ . . .]]]]
wh-*u* Q-*u* wh-*u*
EPP wh-*i* EPP

e. matrix *v* agrees with the embedded C

f. [wh-XP [*v* . . . [C . . . [_____ [*v* . . . ___ . . .]]]]]
Q-*u* wh-*u* wh-*u*
wh-*i* EPP EPP

i. [C . . . [wh-XP [*v* . . . [C . . . [____ [*v* . . . ___ . . .]]]]]]
wh-*u* Q-*u* wh-*u* wh-*u*
Q-*i* wh-*i* EPP EPP

Just as in (69), in (70) the full wh-phrase *kis-ko* has an uninterpretable [Q] feature and an interpretable [wh] feature. The probe *v* in the lower clause interacts with *kis-ko* and raises it to an outer specifier position. This interaction values the [wh-*u*] on *v* as well as the EPP. At this point *kis-ko* is at a phase edge. The embedded C head has no features relevant to wh-movement, and therefore does not interact with it.

According to Rackowski and Richards' approach, the *v* in the matrix clause must probe and Agree with the C head. When this occurs, the embedded CP phase becomes transparent to the *v* head, and it can probe for material beneath it.

If there is no wh-expletive in the numeration, the matrix *v* head must probe to the edge of the embedded *v*P-phase, and it will find and interact with the wh-phrase *kis-ko*. The wh-phrase will raise into an outer specifier of the matrix *v*P, valuing the [wh-*u*] on the matrix *v* and satisfying the EPP property on that head. Note that the wh-phrase does not yet have all of its features valued (namely [Q-*u*]).

As in (69), the C head in (70) has an uninterpretable [wh] feature and an interpretable [Q] feature. The wh-phrase is on the edge of the immediately lower phase (the matrix *v*P), and so the C Probe can interact with it, and all features are mutually valued. That is, the [wh-*u*] on the C head and the [Q-*u*] on the wh-phrase. At this point all features in the derivation have been valued, and the derivation is licit. Whether the wh-displacement in (70) is termed wh-movement or scrambling, the wh-phrase raises into a higher clause, attracted by features of a higher, phase-defining head.[12]

---

[12] A noticeable difference between (69) and (70) is the lack of a Q-*u* feature on the matrix *v* in (70). If there were a Q-*u* on this *v*, the derivation would arrive at the same result. Both Q-*u* features (on the matrix *v* and on the full wh-phrase) would be valued at the point at which the wh-phrase moves into the specifier of the matrix *v*P. The matrix C head would still probe its domain and interact with the wh-phrase. All features will have been valued without difficulty. In other words, its effect is not a critical one in this configuration. Whether or not the*v* that appears here is of the Q or non-Q variety is immaterial in the way this derivation would proceed (for an analogous process, see Chapter 3, Appendix). More accurately, whether it is the*v* bearing a Q feature or the *v* without it that is present in the numeration will have no effect on the surface outcome nor on the grammaticality of the resulting structure.

Let us briefly address at what point the wh-expletive is first introduced in the clause. I argued in Chapter 3 that in Kashmiri the wh-expletive is base generated in a position in which it can be assigned case, and here I will suggest that the same holds true for Hindi-Urdu. This follows claims by Simpson (2000) that wh-expletives are base generated in the specifier of the agreement projection AgrO, and when other DPs are present which need to check case in this position, it will not be possible to generate the wh-expletive. As in Kashmiri, this correctly predicts that sentences in which another wh-DP or the expletive object *yeh* co-occurs in a single clause with the wh-expletive are ungrammatical in Hindi-Urdu.

(71) *Sita-ne yeh kyaa socaa ki ravii-ne kis-ko dekhaa? [Hindi-Urdu]

Sita this expl thought that ravi who saw

Intended: 'Who did Sita think that Ravi saw?' (Simpson 2000)

(72) *Sita-ne kis-ko kyaa dekhaa? [Hindi-Urdu]

Sita-erg who-acc expl saw

Intended: 'Who did Sita see? ' (Simpson 2000)

Updating this view to reflect the framework of this chapter, we will make the same claims concerning the wh-expletive in Hindi-Urdu as we have done for Kashmiri. The wh-expletive is base generated in a position in which it can have accusative case valued; that is, within the c-command domain of the accusative case licenser transitive *v*. Since the wh-expletive cannot be introduced by semantic selection, we know that it must be merged into the specifier of a head that has the EPP property. Aspect is a functional head in the c-command domain of *v* on which it would be

reasonable to posit the EPP property. The wh-expletive introduced into the specifier of AspP has three features: an uninterpretable D feature, an uninterpretable case feature (accusative), and the uninterpretable Q feature. Note that it has no interpretable features at all. When the transitive *v* is introduced, it will interact with the wh-expletive and the expletive's uninterpretable accusative case feature and [D] feature will be valued.

In this view, Hindi-Urdu sentences like (71) and (72) are impossible because only one goal can interact with the *v* head and have its uninterpretable case feature valued. If there is more than one potential goal, such as an additional wh-phrase or a clausal expletive, the uninterpretable case feature on one or the other will go unvalued, and the derivation will fail to converge.

5.3 Phases and the Comparison of Hindi-Urdu and Kashmiri

It is relatively easy to demonstrate that Hindi-Urdu wh-dependencies cannot be approached in precisely the same way as those in Kashmiri. Evidence above indicates that all tensed verbs in Kashmiri appear in C. Since wh-phrases immediately precede verbs in Kashmiri, wh-material must be higher than C, in at least Spec, CP. Hindi-Urdu, however, is different. The verb in Hindi-Urdu appears to be located in a head lower than C, and there is no evidence that the wh-phrase moves as high as Spec, CP. This syntactic difference between the two languages seems to have no semantic correlate and, on the account developed here, reflects only a difference in how wh-features on C receive their value. This distinction, based in our empirical

understanding of Hindi-Urdu and Kashmiri phrase structure, prevents accounts of wh-movement in the two languages from being identical.

We have also suggested above that there is a possible view of A-bar movement in which all C-heads and all *v*-heads have wh-features in every language. In Kashmiri, we have direct morphological evidence that Spec, CP is a stopping-off point for long wh-movement. This is because partial movement takes place to exactly this position, and wh-expletives in the language, whether in intermediate or matrix clauses, also inhabit Spec, CP (Manetta 2005, Bhatt 1999). For other languages, such as Irish, morphology on the C heads themselves has been interpreted to reflect wh-movement through this position (see McCloskey 2001). I have argued here that Hindi-Urdu provides support for the same claims involving Spec, *v*P. I have analyzed partially moved wh-phrases in Hindi-Urdu as coming to rest in this position, and wh-expletives as fulfilling an EPP requirement related to wh-movement on *v*. For other languages, such as Tagalog, morphology on the verb has been interpreted to reflect wh-movement to Spec, *v*P (Rackowski and Richards 2005). I have claimed here that Hindi-Urdu and Kashmiri each choose a distinct head to be the attracting head for successive-cyclic wh-movement.

There is an asymmetry between the account of Kashmiri in Chapter 3 and the present account of Hindi-Urdu. In Kashmiri, the wh-phrase must pass through intermediate specifiers of *v*P, as well as non-interrogative CPs, in long wh-movement. On the other hand, we have claimed here, following Rackowski and Richards (2005), that in Hindi-Urdu, the wh-phrase never passes through the specifier of non-

interrogative CPs. At the root of this difference is the notion that *v* can enter into an Agree relation with the CP which is its complement. Under the assumptions made by Richards (2005) and Rackowski and Richards, this means that once this interaction has occurred the *v* can continue to probe beyond the CP-phase to wh-material at the edge of the *v*P below.[13] On the other hand, C is understood to undertake no Agree interaction with *v* under normal circumstances. Therefore a C is never free to probe through or beyond a *v*P-phase edge.

It does not seem impossible that a language could exist such that morphology reflecting long wh-movement could appear on both verbs and complementizers. This would be a language that reflected every potential stopping point of successive-cyclic movement with some agreement reflex. However, I know of no language so morphologically rich. Instead, it seems that languages choose a way in which to mark the long displacement of wh-material. Either the language requires that wh-material transition through non-interrogative Spec, CP, in which case any morphology which appears does so in the C-domain, or the language allows wh-material to skip these intermediate CPs and target Spec, *v*P, in which case any morphology appears in the verbal domain. Why the phase-defining heads should interact with wh-movement precisely in this way is unknown.

As discussed above, we have couched our empirical observations in terms of a Rackowski and Richards'-style approach. However, it is possible that there are

[13] Of course, this view is incompatible with an understanding of phases in which each phase is dispatched to the interfaces upon completion.

alternative ways of capturing our central intuition, and this is clearly an avenue for future research.

As a final comment on this issue, I want to emphasize here that, contra Rackowski and Richards, this paper claims that adopting the *v*P-based approach to languages like Tagalog and Hindi-Urdu does not dictate the need to sacrifice our current understanding of languages like Irish and Kashmiri. In fact the core of our account of the contrasts between the two languages requires that C(P) is crucial in facilitating wh-movement in Kashmiri. We have a theory in which the two different processes of forming long-distance wh-dependencies coexist, and in fact reflect deep similarities between the phase-defining heads.

5.4 Further Discussion of Direct and Indirect Dependency Accounts

In this section, I will review several analyses previously proposed for wh-expletive constructions that are especially relevant to our discussion of Hindi-Urdu. In Chapter 3 (section 3.2) I introduced both direct and indirect dependency accounts of partial wh-movement and wh-expletive constructions. Here I discuss several more recent approaches to Hindi-Urdu (Mahajan 1990, 2000) and related languages such as Bangla (Simpson and Battarcharya 2003), and situate these approaches with respect to the current effort. I also return to the basic indirect dependency account (Dayal 1996) for further exploration.

The account of wh-dependencies in Hindi-Urdu, unlike that in Kashmiri, has historically been tied to its surface SOV word order and its supposed wh-*in-situ*

status. It is therefore helpful to discern which of these claims about the phrase structure of Hindi-Urdu are crucial to any approach to wh-displacement in the language. In the process of exploring this question, I will discuss how each of these proposals is distinct from the other and ultimately how they differ from the account presented here.

As described in Chapter 3, the representative direct dependency approach to wh-expletive constructions is McDaniel's (1989) account of German and Romani. Recall that direct dependency relies on a syntactic wh-chain with a specific set of well-formedness conditions that connect a wh-phrase and its trace with a wh-expletive. The direct dependency approach assumes that full and partial wh-movement are different manifestations of the same phenomenon, and as a result both create chains with similar (though not identical) well-formedness conditions.

Mahajan (1990) offers what is in essence a direct dependency account of wh-expletive constructions in Hindi-Urdu, though it shares some features with the indirect dependency account as well. Mahajan argues against any form of movement to Spec, CP, whether overtly or at LF. He proposes instead that Hindi-Urdu wh-phrases are in-situ at surface structure, and then undergo QR (adjunction to IP) at LF. In this view, apparent complement clauses are adjoined to the main IP. Extraction in full wh-movement cases occurs prior to the extraposition of the clause. In a wh-expletive construction, the expletive undergoes QR to adjoin to its own (matrix) IP, while the full wh-phrase undergoes QR to adjoin to the adjoined (lower) IP. According to Mahajan's account, this structure permits the formation of a chain

linking the wh-expletive *kyaa* with the full wh-phrase. This has the effect of projecting the scope of the full wh-phrase. In this respect, this represents a direct dependency account, because a syntactic chain is posited between the wh-expletive and the full wh-phrase.

Mahajan (2000) presents another version of this direct dependency approach for Hindi-Urdu that hinges on the operation clausal pied-piping. The difference between this approach and the typical direct dependency analysis is that the wh-expletive is not replaced by its associated wh-phrase at LF. Instead, the wh-expletive is associated with the entire complement clause, which is then raised into the specifier of the matrix CP at LF. In the case of example (2), repeated here, the final LF under Mahajan's (2000) approach is in (73).

(2) Sita-ne kyaa socaa ki Ravii-ne kis-ko dekhaa?

Sita-erg expl thought that Ravi-erg who-acc saw

'Who did Sita think that Ravi saw?'

(73) $[_{CP}\ [_{Cp_i}$ kis-ko$_j$ [ki ravii-ne t_j dekhaa]] kyaa$_i$ [Q [siitaa-ne t_i socaa t_{CP}]]]

This account has something in common with the indirect dependency view, particularly in its reliance upon coindexation and movement at LF. This is nevertheless a direct dependency account because the wh-expletive is ultimately replaced by the movement of the lower clause containing the full wh-word at LF, as opposed to the separate LF movement of wh-phrase in the lower clause and wh-expletive in the higher clause to their respective Spec, CPs.

An account which fits neither in the direct nor indirect dependency camp is Simpson and Battarcharya's (2003) analysis of wh-expletive constructions in another Indic language, Bangla. The major new assumption in this proposal is that Bangla, though typically analyzed as SOV and wh-*in-situ* just like Hindi-Urdu, is in fact SVO and has obligatory overt wh-movement. According to Simpson and Battarcharya, this word order, as well as overt wh-movement, is typically disguised by a number of factors, but is manifest in certain conttexts. If this is the case, then there is no mystery why verbs take finite clausal complements to their right. Simpson and Battarcharya point to the binding-theoretic arguments made by Bayer (1996) and Mahajan (1997). In (74), the matrix clause indirect object binds a pronoun in the postverbal CP. This bound variable reading should not be available if the CP is adjoined higher than the indirect object, since the indirect object could not c-command the pronoun inside CP.

(74) Tumi prottek-ˇa chele-ke$_i$ bolecho[$_{CP}$ ke ta-ke$_i$ durga pujo-y notun jama kapor

You each-cl boy-acc said who he-acc Durga Puja new shirt clothes

debe [Bangla]

give

'You told each boy who will give him new clothes at Durga Puja.'

(Simpson and Battarycharya 2003)

Under Simpson and Battarcharya's account, there is also no mystery in the cases of Bangla sentences like that in (75), in which full wh-extraction appears to have taken place out of the complement clause.

(75) jçn ke_i bollo [$_{CP}$ t_i cole gQche] [Bangla]

John who said Ram left gone

'Who did John say left?' (Simpson and Battarcharya 2003)

Simpson and Battarcharya also argue that wh-movement in Bangla is in fact to a specifier position in the C domain, but that this is disguised by a preference for "non-*wh* arguments and adjunctions in still higher positions in the clause" (Simpson and Battarcharya 2003; 139). If this is in fact the case, "topic-like" subjects precede wh-phrases, which precede verbs.

While the account of Simpson and Battarcharya is not concerned with the role of wh-expletives at all, its conclusions are certainly relevant to questions of Hindi-Urdu wh-dependencies. Specifically, this account reinforces the claim that in Indic languages, finite CPs appearing to the right of the selecting verb can in fact be viewed as complements, not adjuncts. Second, this view claims full, obligatory wh-movement for a language previously viewed as wh-*in-situ*, based on evidence from wh-scope and word order. Although the account I will have proposed in this chapter stops short of assuming underlying SVO word order for Hindi-Urdu, I do claim that finite CPs are complements in the language, and that the displacement of wh-phrases that we see in Hindi-Urdu is in fact wh-movement. In this way, the data presented for Bangla in Simpson and Battarcharya, as well as the account given for that data, does inform the view I ultimately take.

Because the indirect dependency approach (Dayal 1996) was in fact designed to account for Hindi-Urdu, and has been frequently assumed in the literature (as recently as Rackowski and Richards (2005), Breuning (2006)), it is worth taking another moment to compare it to the account proposed here. The major feature that these accounts share is that the wh-expletive or scope-marker is assumed to be generated not in an A-bar position such as the specifier of CP or *v*P, but instead in a position in which it acts as an argument of the verb and/or receives case. Although this position is not precisely the same in each account, it is a feature differentiating this view from the direct dependency approach.

Dayal writes that "Hindi shows quite clearly that the matrix wh is not an A-bar expletive but an argument since it occurs in argument position". (Dayal 1996; 66) There are two aspects of this statement that require elucidation. The first is the nature of wh-expletives. Following Simpson (2000), I have adopted an analysis in which wh-expletives are assigned DP case. Recall that this assumption helped us to understand both overt morphological case on wh-expletives in languages like Hungarian, as well as wh-expletive co-occurrence restrictions. Wh-expletives have no true wh-properties, hence no interpretable features – in this account this is precisely the definition of 'expletive' (whether wh- or DP-). We must also question whether or not the expletive in Hindi-Urdu really occurs in the same position as a direct object. In terms of linear order, the expletive does seem to be in the same relative position (before the verb) that an object would be in. However, this is not necessarily an indication of the structural position of the wh-expletive. I claim here that the expletive

is not occupying the position that a direct object of the verb would occupy, but instead moves from the specifier of AspectP to the specifier of *v*P. While this position is linearly pre-verbal, it is not the direct object position. I have offered empirical support for this claim in Section 3 of this chapter.

An important distinction between the indirect dependency view and the account presented here is the status of the finite subordinate clause in Hindi-Urdu. In an account of wh-dependencies in which it is the valuing of features in the narrow syntax, and not movement at LF, which licenses wh-material, one of the primary motivations for adjunction of finite clauses disappears. Recall that although wh-*in-situ* phrases in a single clause appear to take scope over the clause as a whole (that is, move to Spec, CP at LF), wh-*in-situ* phrases in subordinate clauses in Hindi-Urdu do not. This has been a longstanding puzzle concerning Hindi-Urdu wh-dependencies, and we are now in a position to solve it.

Dayal's (1996) reasoning was that the finite subordinate clause must then be an island and that subjacency must be operative at LF. The adjunction of the subordinate clause to the matrix CP/IP was one way of creating islandhood.

In the view presented here, in which it is the valuing of features that determines whether a sentence with wh-material will be licit, it is no surprise that a sentence containing wh-material in a subordinate clause in Hindi-Urdu cannot be grammatical unless a wh-expletive is present in the higher clause. This is because wh-features on the *v* head in Hindi-Urdu are always bundled with the EPP property. The EPP must be satisfied, whether by full wh-movement/scrambling of the wh-material

from the lower clause, or by the insertion of a wh-expletive. The absence of either of these would result in a derivation in which the matrix clause was interrogative (or had wh-related features) to crash, because there would be no way to value either the EPP on the matrix *v,* or the uninterpretable [wh] feature on the matrix C. The conclusion here is that in a simple Hindi-Urdu indirect question, like that in (76), there are no wh-related features in the matrix clause.

(76) tum jaante ho ki us-ne kyaa kiyaa

you know aux that he-erg what did

You know what he did.

≠ 'What did you know that he did.'

There is then no mystery as to why the interpretation "What did you know he did?" is not possible in these cases; a sentence of this form with wh-related features in the matrix clause would crash. The only grammatical sentences with a wh-phrase from the subordinate clause taking matrix scope are those with overt wh-movement/scrambling or with a wh-expletive in the matrix clause. This is a purely syntactic requirement that has the effect of limiting the licit derivations in Hindi-Urdu which have matrix question interpretations.

The longstanding puzzle that originally led us to maintain that the subordinate clause is a strong island (and that this is relevant at LF) has been resolved. In the view presented here, sentences like (76) in Hindi-Urdu tell us nothing about whether or not the subordinate clause is adjoined to the matrix CP/IP, and in fact give us no particular motivation to assume that it is.

5.5 Hindi-Urdu Phrase Structure Continued

We have spent a great deal of time in this chapter analyzing wh-questions comprised of multiple clauses in Hindi-Urdu. In this section I want to demonstrate how the account proposed here can successfully derive the basic word orders for single-clause constituent questions in the language.

As mentioned above, the unmarked position for the question word in Hindi-Urdu is immediately before the sentence-final verb, regardless of the grammatical role of the question word (Bhatt 2003b, Schmitt 1999).

(77) kitab-ko kis-ne parhi?

book-acc who read

'Who read the book?'

(78) Hamid-ne kyaa parha?

Hamid-erg what read

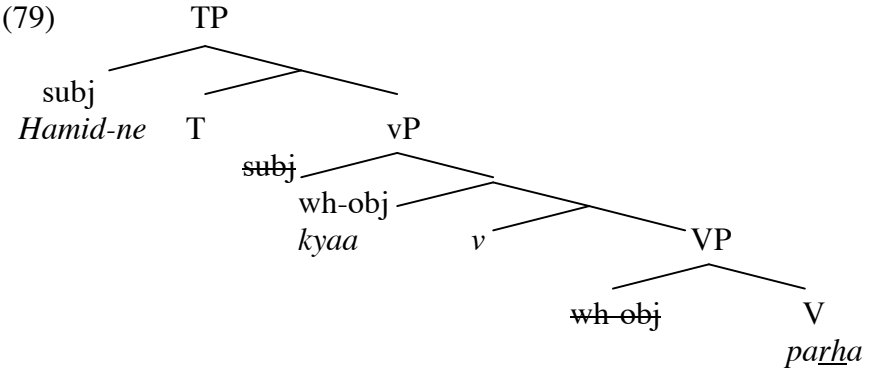
'What did Hamid read?'

,Mahajan (1990) term this tendency "verb-wh adjacency effects". Interestingly for our comparison with Kashmiri, he considers whether this tendency could be the result of verb-second – that is, whether the verb could have moved to C and the wh-phrase to the specificer of CP. Mahajan offers a number of arguments against this approach, including the fact that the wh-phrase is preferentially adjacent to the main verb, not to the tensed auxiliary. In the end he concludes that although this effect cannot be derived via verb-second, he will leave the issue of adjacency effects open. However, the approach put forward in this chapter does offer an understanding of the proximity

of the wh-phrase and the verb. We will examine how this account would arrive at each of the orders in (77)-(78).

In the case of (78), in which the wh-object immediately precedes the verb, the course of the derivation is relatively clear.

(79)



In the structure in (79), the lower (and therefore silent) occurrences of a moved constituent are represented with a strikethrough. The subject, originating in the external argument position in Spec, *v*P has raised into Spec, TP (where it has been assigned nominative case). The wh-object originated as the complement of V, and raised to Spec, *v*P due to its status as a wh-phrase. The process of valuing features proceeds, and we arrive at the constituent order in (78).

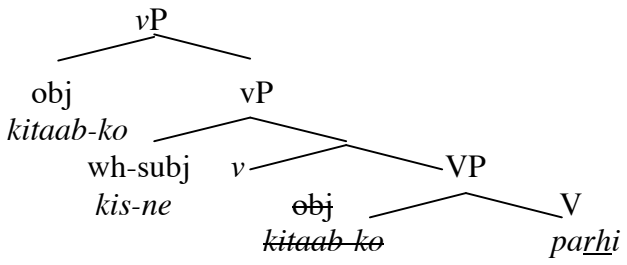
As in the case of any Hindi-Urdu sentence, there are a variety of other grammatical surface word orders. I will assume here that other possible orders of this sentence are achieved via scrambling. There is a significant amount of evidence to suggest (and most native-speaker linguists agree) that the 'preferred' order in constituent questions has the wh-phrase appearing immediately before the verb. In this account, we have taken this preference to be an indication that the wh-phrase

values its features by raising to the specifier position of *v*P. If an alternative word order is required in which this constituent or some other must move, I will assume that this displacement occurs after the wh-phrase values its wh-related features. Much literature has been devoted to the characteristics of scrambling in Hindi-Urdu (i.e. Mahajan 1990, Kidwai 2000, Mahajan 1994, Bhatt 2003b). While we do not have the space to devote to this important subject here, suffice it to say that this account is compatible with a range of previous accounts of scrambling.

In this chapter, I have claimed that wh-displacement from one clause to another, often called 'scrambling', should be understood as wh-movement to value uninterpretable wh-features as in Kashmiri. For this reason I have not referred to wh-movement in Hindi-Urdu as 'scrambling', although I have indicated that it is unclear what is meant by that term in a framework in which <u>all</u> movement is feature and EPP-driven. It is generally agreed that there seem to be different varieties of the phenomenon referred to as 'scrambling' (for instance, Mahajan (1994) and Bhatt (2003b) identify 'short', 'medium', and 'long'). However, in what follows I will differentiate between wh-movement in Hindi-Urdu, defined clearly above, which I will not refer to as 'scrambling', and the process of displacement which results in a variety of discourse-function-motivated word orders for a given clause. I will call this process 'scrambling' for lack of a better term, with the understanding that by all accounts this is actually an umbrella term for a set of processes that may have little in common.

When the subject of the sentence is a wh-phrase, the constituent order is typically as it is in (77), in which the wh-subject precedes the verb. Note that the subject originates in the external argument position, which is typically viewed as the specifier position of *v*P. If this is the case in Hindi-Urdu (and we have no reason to believe that it is not), then a subject question phrase will be originally merged into the derivation at the very position at which it must also value wh-related features. The external argument will never actually be in a position at which it is commanded by *v*. It will be necessary to assume here that upon first merge into a specifier position, a lexical item can potentially value features on the head into whose specifier it has merged. A proposal of this kind has been most clearly articulated for situations of ergative and quirky case assignment (see especially Ura 2000). In the sentence in (77), then, the wh-subject would remain in the specifier of *v*P. In this case, the object has shifted outside of the VP. The structure could potentially take the form in (80).

(80)



Another structure which would give rise to the word order in (77) is one in which the subject moves to Spec, TP, and the object has shifted via intermediate object scrambling (to an A-bar position) over the subject, following Bhatt (2003b) and

Mahajan (1994). The satisfaction of wh-features on *v* would need still to occur in the way suggested above.

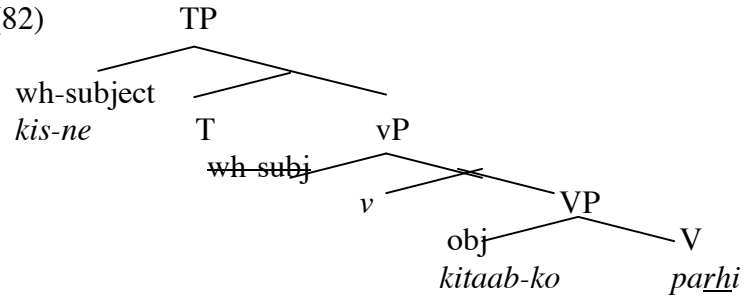
An alternative word order to that in (77), in which the wh-subject precedes the object, is also available, and is occasionally cited by some speakers as preferred (Dayal 1996). Accounting for this order, shown in (81), would mean that the subject will raise to Spec, TP after valuing the wh-features on *v*P upon its Merge into the derivation.

(81) Kis-ne kitaab-ko parhi?

Who-erg book-acc read

'Who read the book?'

(82)



Suffice it to say that the natural outcomes can be accounted for given our assumptions about Hindi-Urdu phrase structure and our analysis of wh-questions. These accounts are consistent with the empirical evidence supplied in section 3 of this chapter that wh-material in Hindi-Urdu is in the specifier position of *v*P.

5.6 Feature structure on *v*

It is clear that a wide range of features would need to appear on *v* if this account proves correct. For instance, a single *v* could need to have features that interact with the external argument (merged into Spec, *v*P), features that interact with the complement of the verb (assigning accusative case to and agreeing with that DP), and wh-features. In Chapter 1, I put forward an account of the left periphery that organizes bundles of features on the phase-defining head C. This account attempted to capture the insights of the so-called cartographic project without the need for a hierarchy of distinct functional projections. In my proposal, each feature bundle contains all the features required for a single probe-goal interaction. These bundles are stacked on the head such that they are accessed in the fixed (top down) order. In this way, the bundle of features on C that interact with a Topic, for instance, is distinct from the bundle that interacts with a wh-phrase.

If the account of Hindi-Urdu proposed in this section is correct, it appears, unsurprisingly, that features on the phase-defining head *v* will also require similar organization. This is an expected outcome in a world in which the featural organization of the phase-defining functional heads plays a fundamental role in determining the syntactic character of a language.

The extension of the proposal to the *v* head emphasizes the third theoretical strand that we have discussed in connection with this approach to Hindi-Urdu. That strand suggests that cross-linguistic variation may follow from properties of specifically the phase-defining functional heads (Chomsky 2005). If, for instance, the

distinct characteristics of wh-dependencies in Hindi-Urdu and Kashmiri are in fact the result of distinct features present on C and *v* respectively, it would be natural that the features on these two heads have similar mechanisms of organization. We can then understand the range of interactions *v* must participate in not as evidence for a series of pre-verbal functional projections, but instead attribute them to the characteristics of features on the single functional head *v*.