



University of  
Massachusetts  
Amherst

## Hindi Wh and Pleonastic Operators

Item Type	Article
Authors	Srivastav, Veneeta
Download date	2026-03-14 10:25:47
Link to Item	<a href="https://hdl.handle.net/20.500.14394/36623">https://hdl.handle.net/20.500.14394/36623</a>

## HINDI WH AND PLEONASTIC OPERATORS\*

Veneeta Srivastav

Cornell University

This paper is an attempt to characterize the properties of question formation in Hindi. Like many South Asian Languages, Hindi has wh in-situ but this wh behaves very differently from the Chinese wh in-situ, as noted by Davison (1984 and 1988), Gurtu (1985), Mahajan (1987) and Wali (1988). We present here an analysis which derives the interpretation of Hindi wh from properties of Hindi phrase structure.

In root clauses, Hindi wh remains in-situ:

- 1a. tum kahãã jaa rahe ho  
you where are going  
"Where are you going ?"
- b. tum kisko pasand karte ho  
you whom like  
"Who do you like ?"

Following Huang's (1982) account of Chinese we can assume that the Hindi wh in 1 (a)-(b) move to Spec of CP at LF, yielding direct questions.

\* I would like to thank Gennaro Chierchia, Wayne Harbert, Jim Huang, Yoshihisa Kitagawa and Kashi Wali for helpful discussions. All errors are my own.

In (2) and (3), however, we see that Hindi wh has properties different from the Chinese wh when it is embedded inside a complement. The wh in (2) are inside non-finite complements, and yield direct questions:

- 2a. tum [kyaa karnaa] jaante ho  
 you what to-do know  
 "What do you know to do?"  
 b. tum [kahãa jaanaa] caahtee ho  
 you where to-go want  
 "Where do you want to go?"

In (3) the wh expressions occur inside finite complements and yield indirect questions:

- 3a. tum jaante ho [ki usne kyaa kiyaa]  
 you know that he what did  
 "You know what he did."  
 b. tum jaante ho [ki vo kahãa gayaa]  
 you know that he where went  
 "You know where he went"

Comparing 3(a) with its Chinese counterpart (4) reveals the difference between the two languages:

4. ni zhidao ta zuo-le sheme  
 you know he did what  
 "What do you know he did?" &  
 "You know what he did."

As argued by Huang (1982), (4) is ambiguous between a direct and an indirect question reading since the matrix verb zhidao "know" can take a + or - wh complement. The wh is therefore free to move to the lower comp, yielding an indirect question, or to the higher comp, yielding a direct question. The Hindi sentence (3) being unambiguous clearly shows that something prevents movement of the wh to matrix comp in Hindi.

We now turn to the strategy Hindi employs to get direct question readings in structures like (3). The primary strategy for direct question formation is what can be called the Scope Marking Strategy:<sup>1</sup>

---

<sup>1</sup> There is also an extraction strategy in Hindi (see appendix for discussion).

## HINDI WH AND PLEONASTIC OPERATORS

- 5a. tum kyaa jaante ho [ki usne kyaa kiyaa]  
 you what know that he what did  
 "What do you know he did ?"
- b. tum kyaa jaante ho [ki kaun aayegaa]  
 you what know that who will come  
 "Who do you know will come ?"

The questions in (5) have a wh expression in the matrix clause which serves only to extend the scope of the wh in the lower clause; it is otherwise semantically vacuous. I therefore call the first wh a pleonastic.<sup>2</sup>

The data in (1) through (5) raise two questions which are addressed in sections 1 and 2: What makes a finite complement a scope island and what is the nature of the Scope Marking Strategy? Our analysis leads to the conclusion that Subjacency obtains at LF; this issue is dealt with in section 3. We end our discussion of Hindi wh by pointing out a problem that certain facts of Hindi raise for current theories of questions.

Section 1: If we look at (2) and (3) we notice that in addition to a difference in finiteness, the complements differ in position. Non-finite complements precede the verb, finite complements follow it. It has to be determined which is criterial for blocking scope in (3), finiteness or the postverbal position. The crucial example comes from the extraposed non-finite clause in (7):

- 7a. tum [PRO kyaa karnaa] caahte ho  
 you what to-do want
- b. \* tum t<sub>i</sub> caahte ho [PRO kyaa karnaa]<sub>i</sub>  
 you want what to-do  
 "What do you want to do?"
8. [CP [IP [IP [VP.. t<sub>i</sub>..] INFL] [CP<sub>i</sub>]]]  
 ↑ \_\_\_\_\_>< \_\_\_\_\_ |

As noted by Davison (1988) a postverbal non-finite clause does not yield a direct question. We can conclude that a clause is a scope island in Hindi when it occurs post verbally. Since finite clauses are always postverbal they are scope islands.

---

<sup>2</sup> The pleonastic is always kyaa "what", regardless of the wh in the embedded clause.

We will now attempt an explanation of why this should be so. Consider (9), a variant of (3), in which a lexical NP precedes the verb.

9. tum yeh jaante ho ki usne kyaa kiyaa  
 you this know that he what did  
 "You know what he did"

Intuitively, yeh "this" stands in place of the finite clause. In this sense it is a pleonastic. Hindi being an SOV language in which case and theta role are assigned to the left, we conclude that the pleonastic yeh is in argument position and absorbs case and theta role in (9). The actual complement is base generated in adjunct position and forms a chain with the pleonastic argument, as shown in (10):

10. [CP [IP [IP [VP.. this<sub>i</sub>..] INFL][CP<sub>i</sub>]]  
 ↑ \_\_\_\_\_ >< \_\_\_\_\_ |

If we consider (3), the version without an overt NP in preverbal position, there are two possible analyses, as shown in (10)':

- 10'. [CP [IP [IP [VP.. t/pro<sub>i</sub>..] INFL][CP<sub>i</sub>]]  
 ↑ \_\_\_\_\_ >< \_\_\_\_\_ |

The complement could be extraposed, in which case it would be coindexed with a trace in preverbal position. Or it could be a base generated adjunct coindexed with pro. The data leaves the choice between the two analyses underdetermined.<sup>3</sup> In either case, the explanation for the scope facts becomes clear. The embedded CP, being in adjunct position, is not L-marked by the verb. It is therefore a blocking category and a barrier. The IP dominating it inherits barrierhood so that movement of wh to matrix comp involves crossing two barriers, a Subjacency violation. The explanation for the Hindi facts follows the standard account for ruling out extraction out of adjuncts (Chomsky 1986).<sup>4</sup>

Leaving aside the implications of this account for the visibility of Subjacency effects at LF, this

---

<sup>3</sup> An extraposition analysis is needed to explain the "extraction" cases discussed in the appendix.

<sup>4</sup> Since we are dealing with an object, an ECP account is ruled out. The trace of kyaa, being lexically governed, would be properly governed.

account answers the question why finite complements are scope islands but it raises another question. Why can't finite complements occur in argument position? Stowell (1981) has proposed a Case Resistance Principle (CRP) which disallows finite clauses from appearing in cased positions. This seems to be at work in Hindi. If CPs cannot appear to the left of the verb, the only possibility for finite complementation is one in which the complement is a structural adjunct linked to a pleonastic in argument position, as in (10) or (10').

Invoking the CRP in Hindi might seem problematic, at first, for an analysis of non-finite complementation since CRP also disallows infinitives from appearing in cased positions. If CRP applies in Hindi, it would follow that the non-finite complements in (2) cannot be infinitives since they appear in cased positions. A closer look at these complements, however, shows this to be a welcome result.

It is easily confirmed that non-finite complements in Hindi are actually gerunds. They have the same distribution as NPs and take the full range of postpositions. Further, when non-finite complements have lexical subjects these subjects have genitive case. Non-finite complements being gerunds, it is expected that unlike finite complements, they may appear in the preverbal position. The verb L-marks such complements so that a wh inside a non-finite complement can move to matrix comp without violating Subjacency. Thus we get direct questions in (2).

Analysing non-finite complements as gerunds makes a further prediction with regard to wh interpretation. Gerunds are usually analysed as IPs or NPs but not CPs (Baker 1985, Milsark 1988). This means that a wh inside a gerund cannot take scope over the gerund, since there is no specCP position available. We see in 11(a) and (b) that this prediction is borne out.

- 11a. \*tum [kyaa karnaa] jaante ho  
       you what to-do know  
       "You know what to do."  
 b. \*tum [kyaa karnaa] puuch rahee ho  
       you what to-do are asking  
       "You are asking what to do."

11(a) shows the impossibility of indirect question readings with non-finite complements; 11(b)

shows the impossibility of non-finite complementation with verbs which must take a +wh complement.

Analysing the finite complement as being an adjunct linked to an argument, we see, allows us to explain a range of facts about wh interpretation in embedded contexts. In the following section we extend this account to cover the Scope Marking Strategy.

**Section 2:** In (12) we see that the pleonastic wh *kyaa* in the matrix must cooccur with a wh embedded in the postverbal clause, and further, that the scope of this wh is extended to the matrix clause:

- 12a. tum *kyaa* jaante ho ki usne *kyaa* kiyaa  
 you what know that he what did  
 "What do you know that he did ?"
- b. tum *kyaa* socte ho ki vo gaaii yaa nahii  
 you what think that she went or not  
 "Do you think she went or not ?"
- c. \*tum *kyaa* jaante ho ki usne yeh kiyaa  
 you what know that he this did  
 "What do you know that he did this ?"

A series of pleonastics can result in an unbounded extension of the scope of the embedded wh:

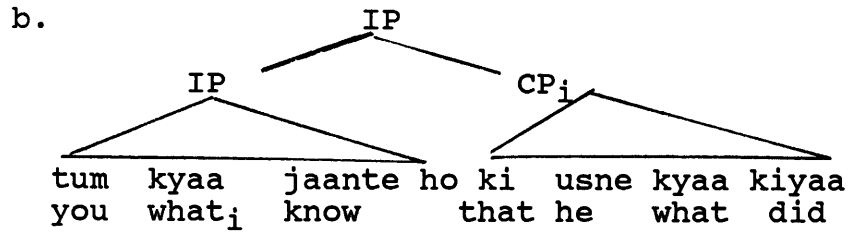
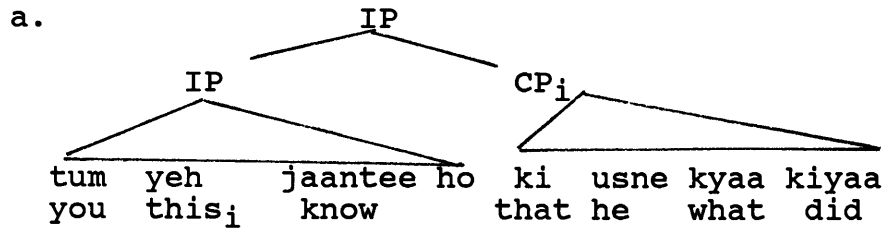
13. tum *kyaa* jaante ho ki vo *kyaa* soctaa hai  
 you what know that he what thinks  
 ki kaun aayegaa  
 that who will come  
 "Who do you know that he thinks will come ?"

Focusing on the simple case, though, we see a clear parallel between 12(a) and (9), repeated below:

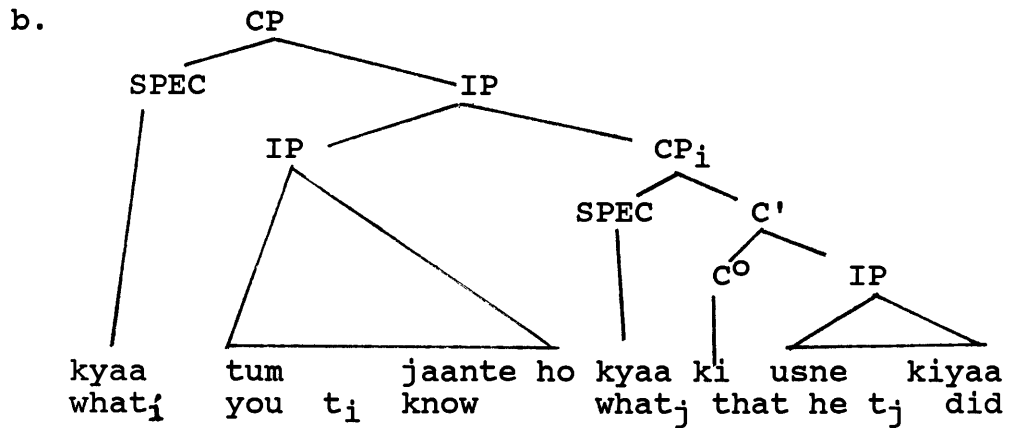
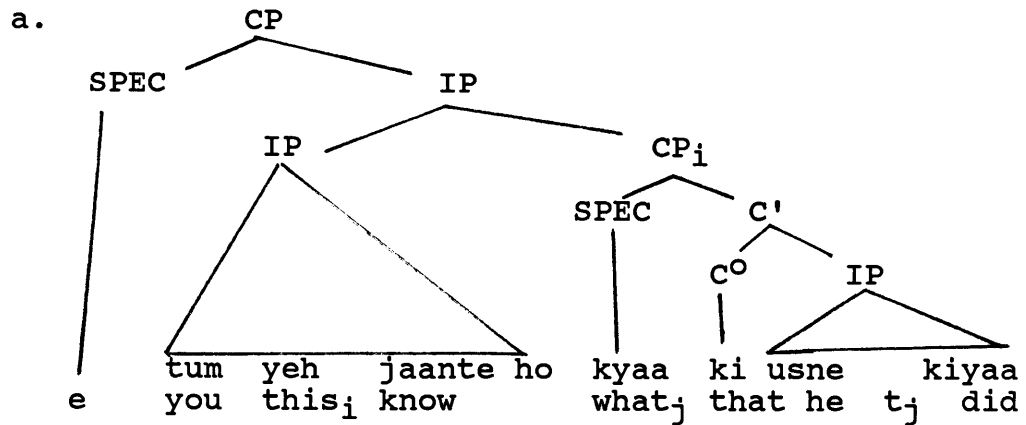
- 14a. tum *yeh* jaante ho ki usne *kyaa* kiyaa  
 you this know that he what did  
 "You know what he did."
- b. tum *kyaa* jaante ho ki usne *kyaa* kiyaa  
 you what know that he what did  
 "What do you know that he did ?"

*Kyaa* "what", like *yeh* "this", can be considered a pleonastic in argument position linked to the complement in adjunct position. Being a wh expression as opposed to a referential expression, however, *kyaa* differs from *yeh* in being an operator. 14(a) and (b), under this view, have similar S-structures (15(a)-(b)) but distinct LFs (16(a)-(b)).

15. S-Structure Representations of 14(a)-(b):



16. LF Representations of 14(a)-(b):



Let us consider the difference between 16(a) and (b). The wh in the lower clause moves to the lower

SpecCP in both structures. As discussed in section 1, further movement is blocked since the lower clause is in adjunct position. In addition, the interrogative pleonastic *kyaa* "what" moves to matrix comp in 16(b). The crucial difference, then, is that there is an operator in the matrix comp of 16(b) but not 16(a).

Let us assume that a *wh* expression in SpecCP can be interpreted as an operator binding a variable (its own trace) or a variable binding another variable (an analogue of the intermediate trace in long distance movement cases). In 16(a), the *wh* in the lower clause can only be interpreted as an operator. If it were to be interpreted as a variable, this variable would not be bound, there being no operator in the matrix comp. In 16(b), on the other hand, the *wh* in the lower comp can only be interpreted as a variable bound by the pleonastic operator. If it were to be interpreted as an operator binding the trace in the lower clause, the pleonastic operator would become vacuous. Its own trace is identified with the adjunct by coindexation, and via spec head agreement with the spec of the adjunct. The pleonastic operator therefore needs to be linked to a +*wh* complement, indirectly binding the embedded *wh* and extending its scope. Thus, the LFs in (16) yield unambiguous interpretations for (14).

This characterization of the pleonastic *wh*, however, needs to be extended. In (17) there are two *wh* expressions in the lower clause and both have matrix scope. This means that the pleonastic in (17) binds two variables. This can be accounted for if we take the pleonastic to be a polyadic operator which can bind more than one variable simultaneously.<sup>5 6</sup>

17. *tum kyaa soctee ho [ki kisne kyaa kharida]*  
 you what think that who what bought  
 "For which *x* and *y*, you think *x* bought *y*"

Before concluding this section I want to compare, very briefly, this analysis of the pleonastic operator with two others known to me.

---

<sup>5</sup> In Srivastav (forthcoming) this is worked out in the semantics.

<sup>6</sup> Hindi *socnaa* can take a + or a -*wh* complement. In the first case it would be translated as English "to think", in the second as "to wonder".

In Davison (1984) and (1988) the terms "pro S" and "interrogative expletive" are used to describe pleonastic *kyaa*. Its function is to re-establish "connectedness" in terms of Kayne (1983), thereby making it possible for the embedded *wh* to take wide scope. Crucially, her LFs for 14(a) and (b) are similar with respect to the pleonastic elements. She notes that the availability of the wide scope reading of the *wh* in 14(b) should entail the narrow scope reading. As we have seen, however, the sentence does not have a narrow scope reading. Thus, her account of the Scope Marking Strategy fails empirically.<sup>7</sup>

The second account of Scope Marking is van Riemsdijk's (1983) for certain dialects of German. As seen in 18(a) through (c), German employs a strategy similar to the one in Hindi. The difference is that German also allows extraction of *wh* so that there are three equivalent forms of the direct question in (18).

- 18a. *was glaubst du, was Peter meint,*  
 what think you what Peter believes  
*mit wem Maria gesprochen hat*  
 with whom Maria spoken has
- b. *was glaubst du, mit wem Peter meint dass*  
 Maria gesprochen hat
- c. *mit wem glaubst du, dass Peter meint dass*  
 Maria gesprochen hat  
 "With whom do you think that Peter believes  
 Maria has spoken?"

van Riemsdijk accounts for these facts by extending the ECP to include pleonastic *was* "what":

- 19a. Every scope marker (or its trace) must govern its scope index.
- b. Every scope index must be governed by a scope marker (or its trace).

As we can see, (19) requires a one-to-one correspondence between operators and variables. While this will account for much of the Hindi facts it will not account for (17) in which a single pleonastic must bind two variables. van Riemsdijk's account, which is otherwise compatible with mine, would have to be

---

<sup>7</sup> Davison also has no explanation for the fact that *wh* inside non-finite complements do not have narrow scope readings (cf. 11).

modified for Hindi.<sup>8</sup> It would have to separate pleonastic scope markers which can govern more than one index from ordinary wh which can govern only one. Such a modification, however, would amount to a notational variant of the one developed here.

We have shown in sections 1 and 2 that the behavior of Hindi wh can be explained with reference to the phrase structure of the language if pleonastic elements are properly characterized.

**Section 3:** The explanation for the absence of wide scope readings of wh in finite complements developed in section 1 turns on the assumption that Subjacency applies at LF. Chomsky (1986), following Huang's (1982) account of Chinese, assumes that Subjacency obtains at S-structure but not at LF. The contrast we have to explain is between Chinese, in which the interpretation of wh in-situ is not constrained by Subjacency and the Hindi phenomenon discussed in section 1 which clearly shows Subjacency effects at LF. One way of reconciling the two would be in terms of a parameter. Such an approach, however, is problematic from the point of view of learnability since evidence for the parameter could hardly be accessible to the child learning Hindi or Chinese.

A more plausible solution is provided by Fiengo et al (1988) who claim that Subjacency obtains at LF though its effect is not visible. Briefly, they contend that an adjunct may be debarrierized by adjunction. Since adjunction is freer at LF than at S-structure, adjuncts in in-situ languages are not scope islands. Because adjunction prevents antecedent government, however, LF extraction out of adjuncts is restricted to those expressions whose traces are lexically governed. Thus Chinese allows extraction of arguments but not adjuncts out of adjuncts. Since Subjacency applies only vacuously at LF, the interpretation of Chinese wh in-situ appears to be constrained only by ECP.

In Hindi we saw that arguments as well as adjuncts cannot be extracted out of complements which are in adjunct position. The line of reasoning suggested by Fiengo et al obviously cannot be applied

---

<sup>8</sup> I believe the same facts hold for the German translation of (17) as for Hindi.



adjunction prevents antecedent government. The adjunct therefore remains a scope island.<sup>11</sup>

If this approach to the phenomenon is on the right track, we predict that there would be a difference between postverbal complements in adjunct position and ordinary adjuncts in Hindi. While the former are absolute islands for extraction, the latter should display the argument-adjunct asymmetry familiar from studies of Chinese. We see that ordinary adjuncts indeed allow extraction of arguments but not adjuncts, suggesting that ECP and not subjacency is at issue:

- 21a. tum [kisko dekhne ke baad] ghar gayii  
 you who seeing after home went  
 "Who did you go home after seeing?"  
 b. \* tum [kaise jaatee samai] usko dekhii  
 you how going time her saw  
 "How did you see her while going?"

While much remains to be done in this area, we have demonstrated that Subjacency does apply at LF. We have also suggested how the visibility of Subjacency effects in the structure under consideration can be reconciled with its absence in languages like Chinese.

**Conclusion:** We would like to conclude our discussion of Hindi wh by pointing out a problem that certain facts of Hindi raise for current theories of questions. A recognised heuristic for understanding the meaning of questions is in terms of their answers. For example, the question in (22) is considered ambiguous since it can be answered in two ways, as noted by Baker (1970).

22. who knows [where Mary bought what]

---

<sup>11</sup> A rather intriguing fact in Hindi is that two finite CPs cannot be coordinated. Thus:

(i)\* usne kahaa ki anu aayii aur ki raam gayaa  
 he said that Anu came and that Ram went

(ii) usne kahaa ki anu aayii aur raam gayaa  
 he said that Anu came and Ram went

The ungrammaticality of (i) can be explained in our account since the two kīs are not properly governed by the inflected verb -- the CP dominating the two conjuncts blocks antecedent government.

Under standard assumptions, an individual answer such as "John" represents the reading in which only **who** is in matrix comp. A pair list answer such as "John knows where Mary bought the book and Bill knows where she bought the pen" represents the reading in which **what** has also moved to matrix comp.

If we consider the Hindi counterpart of (22) we find that it too allows for a pair list answer:

23. kaun janta hai ki merine kahãa kyaa kharida  
 who knows that Mary where what bought  
 "Who knows where Mary bought what?"

We would have to say then that (23) should be associated with two LFs, on a par with (22).

But we are now faced with a paradox. Examples like (3), which lack a direct question reading, show that Hindi wh cannot move out of finite complements; examples like (23), which allow for pair list answers, show that it must. Note that this is a problem for any account of Hindi wh, not just the one outlined here. Frameworks that do not involve syntactic wh movement and a level of LF (for example, Engdahl (1986) and Groenendijk and Stokhof (1984)) nevertheless have scope mechanisms which ensure that wh expressions can be interpreted at different nodes in the tree. As such, the interpretation of the Hindi wh in (3) and (23) is equally problematic in any of the current frameworks.

We suggest, given the robustness of the judgements in (3), that an alternative account for pair list answers is required, if not for all languages, at least for Hindi. This, however, is clearly beyond the scope of this paper (see Srivastav forthcoming).

We have argued in this paper that Hindi exemplifies a type of wh language which is different from in-situ languages such as Chinese. We have shown, however, that wh phenomenon in Hindi is amenable to a principled explanation based on the special properties of the phrase structure of the language and a modified view of the relevance of Subjacency at LF.

**Appendix:** Gurtu (1985) and Mahajan (1987) claim that Hindi has extraction at S-Structure but many speakers do not accept questions like (i). We believe that acceptability of such questions is sensitive to

intonation. (i), for example, becomes acceptable if the matrix subject is stressed.

- i. kyaa<sub>i</sub> tum jaante ho ki usne t<sub>i</sub> kiyaa  
 what you know that he did  
 "What do you know that he did?"

We propose here that the extraction strategy, though a legitimate construction, is not the normal strategy for forming questions but a topicalized structure, with a wh expression IP adjoined at S-Structure.<sup>12</sup> This adjunction has semantic consequences -- (i) has a contrastive reading and may be translated as "What do you, as opposed to others, know that he did?" We give below arguments to support our claim.

Mahajan (1987) notes that an extracted wh in Hindi follows, rather than precedes, the complementizer:

- ii. anune puchaa ki kaun tum socte ho ki t aayaa  
 Anu asked that who you think that came  
 "Anu asked who you think came."

He therefore argues that Hindi ki "that" is in pre specification. This proposal is untenable, however, in view of the fact that in relative clauses ki cannot precede the wh word when the two cooccur:

- iii. \*vo baat ki jisko tum samajhte ho  
 that matter that which you understand  
 "The matter that you understand."

An immediate advantage of our proposal that the "extraction strategy" actually involves IP adjunction rather than movement to SpecCP yields a straightforward account of the facts in (ii) and (iii). The wh in (ii) is IP adjoined and thus follows the complementizer. The wh in (iii) is in SpecCP and therefore precedes it.

In section 1 we claimed that the postverbal finite clause in Hindi is a scope island. We have to show that the topicalization in (i), which also involves extraction, does not contradict this claim. Consider (iv):

---

<sup>12</sup> At LF, of course, the wh would move to SpecCP for interpretation.

- iv. \* kyaa<sub>i</sub> tum yeh jaante ho ki usne t<sub>i</sub> kiyaa  
 what you this know that he did  
 "What do you know that he did?"

(iv) shows that topicalization never cooccurs with an overt pleonastic in preverbal position. Taking the presence of a pleonastic as a diagnostic for complements base generated in adjunct position, we conclude that topicalization never takes place out of such complements. We hypothesize that they involve extraction out of clauses which are in preverbal position at D-Structure. At this point the complements are L-marked and movement out of them does not violate Subjacency.

### References

- Baker, C.L. (1970) "Notes on the Description of English Questions: The Role of an Abstract Question Morpheme," *Foundations of Language* 6.
- Baker, M. (1985) "Syntactic Affixation and English Gerunds," *Proceedings of WCCFL* 4.
- Chomsky, N. (1986) *Barriers*, MIT Press.
- Davison, A. (1984) "Syntactic Constraints on WH-in situ: WH-questions in Hindi-Urdu," *LSA*.
- (1988) "Syntactic Structure in Semantic Representation," *LSA*.
- Engdahl, E. (1986) *Constituent Questions*, Reidel.
- Fiengo, R., C-T. J. Huang, H. Lasnik and T. Reinhart. (1988) "The Syntax of Wh-in-Situ," *WCCFL* 7.
- Gurtu, M. (1985) *Anaphoric Relations in Hindi and English*, Ph.D. dissertation, CIEFL (India).
- Groenendijk, J. and M. Stokhof. (1984) *Studies in the Semantics of Questions and the Pragmatics of Answers*.
- Huang, C-T. J. (1982) *Logical Relations in Chinese and the Theory of Grammar*, Ph.D. dissertation, MIT.
- Kayne, R. (1983) "Connectedness," *LI* 14.
- Mahajan, A.K. (1987) "Notes on Wh-questions in Hindi," *South Asian Languages Association Roundtable*.
- Milsark, G.L. (1988) "Singl-ing", *LI* 19.
- Srivastav, V. (forthcoming) *The Syntax and Semantics of Wh Dependencies in Hindi*, Ph.D. dissertation, Cornell University.
- Stowell, T. (1981) *Origins of Phrase Structure*, Ph.D. dissertation, MIT.
- van Riemsdijk, H. (1983 ?) "Correspondence Effects and the Empty Category Principle," *Tilburg Papers in Language and Literature*.
- Wali, K. (1988) "A Note on Wh Questions in Marathi and Kashmiri," *Cornell Working Papers in Linguistics* 8.