

MULTIPLE WH QUESTIONS (SynCom #66)
to appear in M. Everaert and H. van Riemsdijk (eds.)
The Blackwell Companion to Syntax, 2003

Veneeta Dayal
Rutgers University
dayal@rci.rutgers.edu

I. INTRODUCTION

1.1: Wh Expressions as Diagnostics of Scope	3
1.1.1. Fronting and Possible Answers as Indicators of Scope	3
1.1.2. Constraints on Scope: ECP and Subjacency	4
1.1.3. Alternatives to Covert Movement	7
1.1.4. Overview of the article	10
1.2 Cross-linguistic Variation in Multiple Wh Questions	11
1.2.1. Non-Fronting Languages	11
1.2.2. Multiple Fronting Languages	14
1.2.3. Languages Without Multiple Wh Questions	16
1.2.4. Optional Fronting Languages	17
1.2.5. Explanations for Typological Variation	18

II: SUPERIORITY EFFECTS

2.1: Amelioration under D-linking	19
2.2: Explanations for Superiority	23
2.2.1. Within the Principles and Parameters Framework	23
2.2.2. Economy-based Accounts	25
2.2.3. Accounts based on Functional WH	27
2.3: Typological Variation in Superiority Effects	31
2.3.1. Superiority in Multiple Fronting Languages	31
2.3.2. Languages without Superiority Effects	33

III: SUBJACENCY AND WH IN-SITU

3.1: The Diagnostic of Possible Answers	34
3.1.1: Single vs. Multiple Pair Answers	34
3.1.2: Deriving Single and Multiple Pair Answers	36
3.2. Pair Answers Across Islands	39
3.2.1: Single Pair Answers Across Islands	39

3.2.2: Multiple Pair Answers Across Islands	40
3.3. Re-assessing Subjacency at LF	42
3.3.1: The Evidence For and Against Subjacency	42
3.3.2: Specification, Scope and D-linking	45
<u>IV: CONCLUSION</u>	47
<u>V. REFERENCES</u>	47

I. INTRODUCTION*

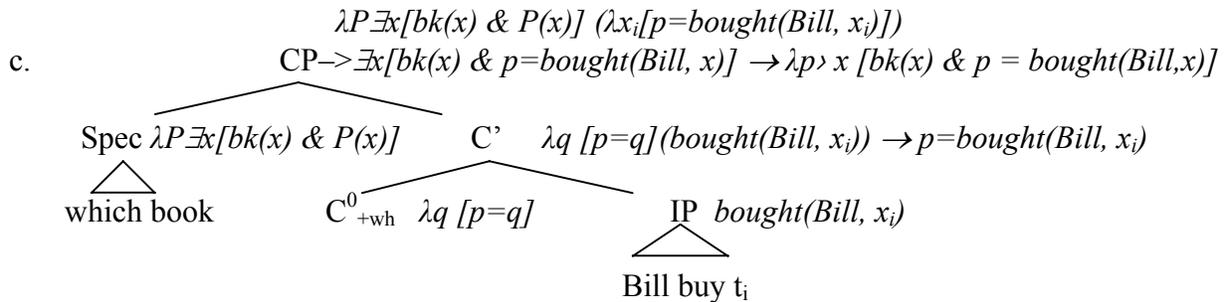
1.1: Wh Expressions as Diagnostics of Scope

1.1.1. Fronting and Possible Answers as Indicators of Scope

I will begin this survey of multiple wh questions by outlining those theoretical claims for which the study of wh dependencies has been particularly important. These include the claim that there is a level of syntactic representation intermediate between overt syntax and the interpretive module, known as Logical Form (LF), and that movement at this level is subject to the Empty Category Principle (ECP) but not Subjacency.

Wh expressions provide a special window into the nature of syntactic dependencies because in many languages they appear at the periphery of the clause, in the position at which they are assumed to be interpreted. In this they differ from quantified expressions, for example, which also may be argued to require clausal scope for interpretation but do not seem to occur in a specially designated scope position in any language. To appreciate the diagnostic power of wh expressions, consider a simple question in English, such as (1a) and one of its possible answers given in (1b):

- 1a. Which book did Bill buy?
 b. (Bill bought) W&P.
- 2a. $\lambda p \exists x [\text{book}(x) \ \& \ p = \text{bought}(\text{Bill}, x)]$
 b. {Bill bought W&P, Bill bought Aspects}



Intuitively, the fronted wh expression is linked to the object of the verb *buy*. Syntactic tests, using agreement and binding facts for example, can further establish the syntactic dependency between the fronted element and the gap in argument position. In logical form this dependency is understood as an operator-variable relation. Under a standard semantics for questions such as Hamblin (1973) or Karttunen (1977), (1a) is interpreted as in (2a), yielding the set of propositions obtained by varying assignments to the variable bound by the wh \exists operator.¹ In a context where the quantifier ranges

* I am grateful to Chris Barker and Satoshi Tomioka for very useful comments and to Heather Robinson and Scot Zola for help with the preparation of the ms.

¹ Karttunen differs from Hamblin in restricting the set of propositions denoted by the question to those that are true. This difference affects the precise way in which questions are

over the set {War and Peace, Aspects}, we get the set of propositions in (2b) as the denotation of the question. An appropriate answer is understood to be one that picks out the true proposition from this set, specifying in effect, a value for the fronted wh expression. (2c) makes explicit the mapping from syntactic structure to semantic interpretation (Bittner 1994, von Stechow 1996). The crucial operation that shifts a sentence from a declarative proposition-type meaning to an interrogative set-of-propositions type meaning hinges on the propositional variable p over which lambda abstraction can take place. Crucially for our purposes, the fronted wh expression must be interpreted outside the scope of this variable. If the +WH specification on C^0 is taken to be the position at which the propositional variable is introduced, and the fronted wh taken to be in Spec CP, the mapping from syntax to semantics is straightforwardly compositional. Thus, one might take fronting of wh expressions as direct evidence of their scope position.²

Multiple wh questions have special syntactic and semantic properties. In English they are formed by fronting one wh expression and leaving the other(s) in situ. Possible answers to such questions specify values not only for the fronted wh expression but also for those in situ:

- 3a. Which student bought which book?
 b. Bill bought W&P, Sue bought Aspects.

- 4a. $\lambda p \exists x \exists y [\text{student}(x) \ \& \ \text{book}(y) \ \& \ p = \text{bought}(x,y)]$
 b. {Bill bought W&P, Bill bought Aspects, Sue bought W&P, Sue bought Aspects}
 c. $[_{CP} \text{Which student}_i \ [_{IP} t_i \text{ bought which book}]] \rightarrow [_{CP} \text{which book}_j \ \text{Which student}_i \ [_{IP} t_i \text{ bought } t_j]]$

If we take possible answers as a diagnostic, and assume that specification of values is due to clause external existential quantification binding variable positions inside IP, as the Hamblin-Karttunen semantics of questions suggests, multiple wh questions force us to the conclusion that there must be a way of interpreting wh expressions in situ that allows them to form operator variable relations of the kind shown in (4a). Following the syntax-semantics mapping outlined above, it appears plausible to assume that the wh in situ moves covertly to a position outside C^0 . A semantic argument for the existence of LF can thus be made using possible answers to multiple wh questions as indirect evidence.

1.1.2. Constraints on Scope: ECP and Subjacency

If considerations of interpretability lead us to postulate an abstract level of syntax, it should be possible to characterize it in theoretical terms. In particular, one could ask if there are constraints

related to their answers but does not affect the general point being made here. In section 3, however, the importance of using Hamblin sets as question denotations and encoding Karttunen's truth requirement in an answerhood operator will be shown to be relevant in making more fine-grained distinctions in pair answers. I should also note that Hamblin leaves it open whether the common noun in the wh expression will be interpreted outside the question nucleus via presupposition projection or in its base position.

² I ignore intensional operators since they do not affect the issues under discussion.

on the operation of Move α at this level. In this subsection we will consider evidence that has been used to establish the status of the ECP and Subadjacency as constraints on movement at LF.

To see how multiple wh questions have shaped our view of LF consider the following questions, uttered out of the blue:

- 5a. Who saw what?
 b. ?* What did who see?

- 6a. How did Bill read what?
 b. *What did Bill read how?

The choice of which wh expression will be fronted in a multiple wh question is clearly sensitive to the status of the argument. In (5) a subject is chosen over the object for fronting, in (6) an adjunct is preferred over the object. This phenomenon, which goes by the name of Superiority in the literature, can be accounted for by imposing syntactic requirements on traces:

- 7a. [CP [Spec-i what_j [Spec-i who_i]] [IP t_i saw t_j]]
 b. [CP [Spec-j who_i [Spec-i what_j]] [IP t_i saw t_j]]

- 8a. [CP [Spec-i what_j [Spec-i how_i]] [IP Bill read t_j t_i]]
 b. [CP [Spec-j how_i [Spec-j what_j]] [IP Bill read t_j t_i]]

9. Empty Category Principle:

Traces must be properly governed.

A properly governs B iff A theta-governs B or A antecedent-governs B (Chomsky 1986)

The ECP states that traces must be properly governed and an account of the Superiority effects can be given if it is assumed that Spec CP carries the index of the first wh expression that moves into it (Lasnik and Saito 1992). As the derivations in (7a) and (8a) show, the ECP is respected when covert movement is from the theta-marked direct object position, leaving the subject or the adjunct trace created at S-structure to be antecedent governed by the fronted wh expression. In the unacceptable cases in (7b) and (8b), the situation is reversed. Though the trace created at S-structure satisfies the ECP (doubly), the trace created at LF does not.

Note that ECP as a constraint on S-structure movement can also be established on the basis of examples like (10). The explanation for the contrast between (10a) and (10b) draws on the same elements of explanation as the contrast between (7a)-(8a) and (7b)-(8b). (10b) is unacceptable because an intervening complementizer blocks antecedent government of a trace in non-theta position:

- 10a. [CP [Spec-i Which author_i] [C' do [IP you believe [CP [IP t_i wrote the book]]]]?
 b. * [CP [Spec-i Which author_i] [C' do [IP you believe [CP that [IP t_i wrote the book]]]]?

What makes the argument from multiple wh questions for the existence of ECP effects at LF significant is that it gives content to the idea that LF is a level of syntax, with properties in common

with S-structure. This result is important because it argues against an alternative view in which assignment of scope can be viewed as a purely semantically driven step in the computation.

Let us turn now to constraints on scope covered by the subjacency condition. As Ross 1967 showed, though wh fronting can occur across clauses, it is restricted in certain contexts:

- 11a. Which book did Mary tell John to buy t?
 - b. Which book does Mary think John will buy t?
- 12a. *Which book does Mary know a man who has read t?
 - b. Which book is such that Mary knows a man who has read it?
- 13a. *Which book does John know where Mary bought t?
 - b. Which book is such that John knows where Mary bought it?

Here we have evidence that wh dependencies created at S-structure are syntactically constrained. In (12a), the fronted wh expression associated with a position inside the complex noun phrase is clearly unacceptable (Ross's CNPC). It is easy to see that the problem is not semantic in character. (12b) shows that if the relevant dependency between the wh expression and the gap inside the complex noun phrase could be established in (12a), it would be interpretable as a proper request for information. The contrast between (13a) and (13b) makes the same point for wh dependencies involving embedded questions (Ross' Wh-Island Constraint). The two can be subsumed under the Subjacency Condition:

- 14. Subjacency Condition (Chomsky 1986): If (α_i, α_{i+1}) is a link of a chain, then α_{i+1} is subjacent to α . β is n-subjacent to α iff there are fewer than n+1 barriers for β that exclude α .

Given that wh in-situ in multiple wh questions move at LF, one can test whether the Subjacency Condition also holds at that level. The following data test the relevance of the two subcases of Subjacency:

- 15a. Who knows a man who wrote which book?
 - b. Bill knows a man who wrote Aspects.
- 16a. $\lambda p \exists x \exists y [\text{person}(x) \ \& \ \text{book}(y) \ \& \ p = \exists z [\text{man}(z) \ \& \ \text{wrote}(z,y) \ \& \ \text{knows}(x,z)]]$
 - b. {Bill knows a man who wrote Aspects, Sue knows a man who wrote Aspects, Bill knows a man who wrote LGB, Sue knows a man who wrote LGB}
- 17a. Which student knows where Mary bought which book?
 - b. Bill does / Bill knows where Mary bought which book.
 - c. Bill knows where Mary bought Aspects and Sue knows where Mary bought LGB.

- 18a. $\lambda p \exists x \exists y$ [student(x) & book(y) & p = know(x, $\lambda p' \exists z$ [place(z) & p' = buy(mary,y at z)]]
 b. {Bill knows where Mary bought Aspects, Sue knows where Mary bought Aspects,
 Bill knows where Mary bought LGB, Sue knows where Mary bought LGB}

Baker (1970) argued that the possibility of answering (17a) with (17b) or (17c) indicates a syntactic ambiguity in the scope of *wh* in situ. Answers like (17b), specifying values for the fronted *wh* only, can be derived if the *wh* in-situ takes scope over the embedded clause. List answers like (17c), on the other hand, which pair students and books indicate that the *wh* in situ takes matrix scope, in violation of the Wh-Island Constraint. A similar argument, based on possible answers has been made for matrix scope of the *wh* in situ in (15), in violation of CNPC. These examples establish a difference between operator-variable dependencies created at S-structure and those created at LF. Based on such examples, it has been claimed that the Subjacency Condition does not hold at LF.

As in the case of the ECP, the scope of *wh* in-situ is crucial in establishing the status of Subjacency at LF. Quantifiers, for example, do not help in this regard. (19) shows that a reading where for each book there is a possibly different student who thinks John will buy it is not available:

19. Some student thinks John will buy every book.

Since we know from (11b) that this is a context which allows extraction, the inability of the \forall quantifier to take scope over the \exists suggests that it is a property of Quantified NP's that their scope must be clause-bounded (but see case # 113 for discussion). The relevance of subjacency at LF, therefore, cannot be determined by examining the behavior of quantifiers, and our primary source for this claim remains the interpretive possibilities of *wh* in-situ.

1.1.3. Alternatives to Covert Movement

We have seen that the observable syntactic structure of English multiple *wh* questions reveals the scopal property of one *wh* expression directly by fronting it, while concealing that of the other(s), for which we must look to indirect evidence in the form of appropriate answers. We have also seen that these twin indicators of scope make multiple *wh* questions an extremely powerful diagnostic tool. Adopting the Principles and Parameters model of grammar, I have demonstrated the power of multiple *wh* questions to test the syntactic properties not only of S-structure but also of LF.

There are, of course, many alternative analyses of the phenomenon that have been proposed over the years and we will discuss some of them in subsequent sections. For now, I would like to briefly mention three mechanisms for interpreting *wh* expressions which do not depend on movement, Cooper-storage, Unselective Binding and Choice Functions. My aim here is to explicate differences between them with respect to the assignment of scope as well as to note points of contact. I will not try at this point to summarize the particular arguments made or specific conclusions drawn by the proponents of each approach.

Let us take as a starting point, the idea that covert movement is not essential to the interpretation of *wh* in-situ, or other scope bearing elements for that matter. An alternative proposed by Baker (1970) depends on coindexation between a Q-morpheme in complementizer position and a *wh* in-situ. His analysis of (17a), repeated below as (20a), shows the two scope possibilities discussed earlier:

20a. Who knows where Mary bought what?

- b. $[Q_i \text{ who}_i [t_i \text{ knows } [Q_{j,k} \text{ where}_j \text{ Mary bought } t_j \text{ what}_k]]]$
- c. $[Q_{i,k} \text{ who}_i [t_i \text{ knows } [Q_j, \text{ where}_j \text{ Mary bought } t_j \text{ what}_k]]]$

Note that given a proposal like this, the notion of LF as essential to the syntax-semantics interface can be dispensed with, but only if it is combined with an alternative way of interpreting the wh expression than the one given in section 1.1. The wh expression there is interpreted like an existential generalized quantifier which combines with a predicative expression built up out of the denotation of C' , the syntactic constituent it has scope over. The challenge posed by structures like (20b)-(20c), or any other structure encoding scope without covert movement, is to find a compositional way of assigning clausal scope to expressions within the clause at the interface level.

Engdahl (1986), working within a non-derivational syntactic framework, suggests that wh in situ should be interpreted on a par with other generalized quantifiers using the storage mechanism of Cooper (1983). Briefly put, this involves interpreting the wh in argument position as a variable and keeping its quantificational force, the existential along with the restriction, in store till an appropriate scopal position is reached, at which point existential binding of the variable is effected. Hamblin-Karttunen denotations, equivalent to the ones derived under the movement approach, could thus be derived for multiple wh questions without having to take recourse to covert movement.³

A second approach for interpreting wh in situ without movement is proposed by Pesetsky (1987) and Nishigauchi (1986, 1990). Recognizing the kinship between wh expressions and indefinites, they adopt the unselective binding approach of Lewis (1976), Kamp (1981) and Heim (1982) where the indefinite provides a variable that must be bound by sentential operators. Under this approach, we might take the C^0_{+WH} to provide not only the propositional variable p , but also the existential quantifier binding the wh variables. In the basic cases, the resulting set of propositions is the same as in the other proposals considered so far, but there is a crucial difference with respect to the status of the common noun inside the wh. In this case, it remains within the scope of the propositional variable, simply adding an additional condition on the variable:

21a. Which student read which book?

- b. $[Q_{i,j} \text{ which student}_i [t_i \text{ read which book}_j]]]$
- c. $\lambda p \exists x y [\text{student}(x) \ \& \ p = \text{read}(x, y) \ \& \ \text{book}(y)]$

The third approach, proposed by Reinhart (1997), is similar in that it takes the wh in situ to be an indefinite bound by existential closure. The key difference is in the interpretation of the common noun. Reinhart points to the fact that the unselective binding approach makes the wrong prediction in cases like the following:

³ Of course, it should be emphasized that frameworks with or without movement may be combined with other theories of questions, such as Groenendijk and Stokhof (1984) where questions denote propositions rather than sets thereof. The considerations that go into the choice between Hamblin-Karttunen and G&S do not depend on the syntactic properties of multiple wh questions. I therefore do not discuss them here.

- 22a. Who will be offended if we invite which philosopher?
 b. $\lambda p \exists x y [p = [[\text{invite}(\text{we}, y) \ \& \ \text{philosopher}(y)] \rightarrow \text{offended}(x)]]$
 c. Lucie will be offended if we invite Donald Duck.

- 23a. $\lambda p \exists x \exists f [p = [[\text{invite}(\text{we}, y) \ \& \ f(\text{philosopher})] \rightarrow \text{offended}(x)]]$
 b. $\lambda p \exists x \exists f [\text{student}(x) \ \& \ p = \text{read}(x, f(\text{book}))]$

As (22b) shows, the *wh* in-situ is in the antecedent of a conditional. This means that any individual who is not a philosopher suffices to make the antecedent false and hence the conditional proposition true. (22c) thus qualifies as a potential answer to the question. Reinhart therefore argues for a treatment of *wh* in situ in terms of choice functions. Under this view the quantification ranges over functions that take a (non-empty) set and pick out an individual from that set. *Wh* expressions, like other indefinites, do not contribute an individual variable but denote a set to which the existentially quantified function applies. The problem is resolved since Donald Duck would simply not be a possible value for $f(\text{philosopher})$ in (23a). For completeness I give the interpretation of the simpler case (21a) in (23b).

The availability of well-defined semantic mechanisms like Storage, Unselective Binding and Choice Functions makes it clear that *wh* scope phenomena are not inextricably linked to movement and successful analyses have been given which dispense with movement or combine movement with alternative scope assignment strategies to different degrees. On the two ends of the spectrum, then, lie frameworks like the Principles and Parameters model that encode all scope relations via movement and frameworks that eschew movement altogether, such as Head-driven Phrase Structure Grammar. For the first type of framework, alternative scope mechanisms are obviously superfluous while they are clearly essential for the second type. There are also a number of interesting possibilities that fall between these two opposing points. One, advocated by Pesetsky (1987), favors *wh* movement at S-structure and LF but allows for certain (discourse-linked) *wh* expressions to be subject to unselective binding. Pesetsky thus argues for a uniform characterization of Move α at both syntactic levels, explaining the non-standard behavior of some *wh* in-situ as a reflection of their being a distinct phenomenon subject to distinct principles. Reinhart (1997), on the other hand, takes the possibility of covert movement to be antithetical to the basic tenets of Minimalism, and argues for a movement account for fronted *wh* only, reserving an account in terms of choice functions for all *wh* in-situ. As is clear in even this cursory a list, a number of *a priori* defensible positions are possible and one might wonder to what extent the issues raised by multiple *wh* questions, discussed within a framework utilizing movement as the sole scoping mechanism, transfer over to frameworks that rely on other mechanisms. I will try to clarify below the sense in which I believe the central questions about the interface between form and meaning are maintained across these divergent views.

Consider the role of Subjacency in distinguishing between S-structure and LF in the Principles and Parameters model. As we noted, this is based on two types of evidence, evidence of Subjacency effects in fronting possibilities and absence of Subjacency effects in answers specifying values for *wh* expressions inside islands. The first kind of evidence is direct, the second indirect in that it relies on the premise that possible answers specify values for all and only those *wh*

expressions that have matrix scope. In subsequent sections we will have a chance to probe this assumption more fully. Taking it at face value for now, it can be easily shown that the descriptive generalizations revealed by multiple wh questions have equal significance in alternative models of grammar.

In a non-derivational model like Engdahl's, for example, something akin to Subjacency applies to operator-gap dependencies when wh operators are in clause initial position. The same does not apply to wh in-situ whose quantificational force can be held in store beyond the relevant bounding nodes. Approaches such as Pesetsky's, adopting unselective binding for D-linked wh in-situ, or those like Reinhart's adopting choice functions for all wh in-situ make similar distinctions. A dichotomy between two scope mechanisms is recognized, only one of which is considered sensitive to subjacency.⁴ While there are substantive differences between the various positions, the point I would like to emphasize is that they share fundamental assumptions about the relation between possible answers and scope. To the extent that the argument for a particular property of abstract movement turns on the diagnostic of possible answers, and specification of values is linked to binding of a variable in argument position by an existential quantifier with scope over the propositional variable, it will also be of relevance in models employing alternative scope mechanisms. Although I will continue to talk in terms of movement much of what I say is intended to apply more generally.

1.1.4. Overview of the Article

In the previous subsections I have tried to convey some idea of the central role that multiple wh questions play at the interface of syntax and semantics. In doing so, I have introduced the two empirical phenomena that have been central in the analysis of multiple wh questions, Superiority effects and apparent violations of Subjacency. In the next subsection, I will broaden the empirical base by surveying languages in which multiple wh questions are formed differently from English. Section 1.2, therefore, has a primarily typological goal, after which the article takes up the task of examining the theoretical import of descriptive generalizations more closely. In doing so, it also tries to refine those generalizations.

Section 2 begins by noting that Superiority effects are not manifested in D-linked contexts and goes on to a survey of various accounts that have been given for the presence or absence of Superiority. It concludes by addressing the possibility of cross-linguistic variation in Superiority. Section 3 takes a closer look at subjacency violations typically associated with wh in-situ. It begins by examining the nature of pair answers on which this view is based and goes on to elaborate the semantics required to derive them. It ends with a reassessment of the claims that have been made about subjacency.

It might be worthwhile, before proceeding further, to delineate the intended coverage of this survey. Using answers to questions as a guide, I confine my study of multiple wh questions to those whose appropriate answers require specification of values for more than one wh expression. By this criteria, questions with the following schemata do not qualify as proper objects of this case study, if

⁴ Note that in all approaches, a further distinction between the scope of wh in-situ and quantified NP's also has to be made.

their appropriate answers give values for only one wh expression:

- 24a. [...Wh₁ ...[...wh₂...]]
b. [...Wh₁...[...wh₂...t₁]]

25a. Who wonders when Mary will come?

- b. ravi kyaa soctaa hai, ki meri kab aayegii
Ravi what thinks that Mary when will come

Hindi (Dayal 1996)

26a. Despre care_i ştii [cine_j t_j i-a povestit t_i]
About which you-know who to-him has told

“Which one do you know who told him about?”

Romanian (Comorovski 1989)

- b. ni xiang-zhidao [shei mai-le sheme]
You wonder who bought what

Chinese (Huang 1982)

Examples of the schema in (24a) are given in (25a)-(25b). (25a) only requires information about the identity of wh₁. (25b) instantiates what is known as the Scope Marking or Partial Wh Movement construction, characterized by the fact that it appears, on surface at least, to seek for the identity of wh₂ only. The schema in (24b) is exemplified by the possibility, available in some languages, of scoping out of a wh-island, whether at S-structure or LF. (26a) is a question about the fronted wh and although (26b) is ambiguous with regard to which wh takes matrix scope, only one of them can do so. While I will mention such cases in describing the cross-linguistic possibilities in the next subsection, and in other sections whenever relevant, I will not delve into issues raised by them. Scope Marking and Wh extraction out of islands are discussed in other case studies (SYNCOM REF # 77 and # 41 & 121 respectively), to which I refer the reader. I should also mention that this case should be read in conjunction with case # 121 on Wh In Situ, since there is substantial overlap between them.

1.2: Cross-linguistic Variation in Multiple Wh Questions

1.2.1. Non-Fronting Languages

English, the language I have used so far to illustrate the theoretical import of multiple wh questions, represents only one of three documented strategies for the formation of such questions. This section introduces the two other strategies, namely, the absence of fronting in some languages and the obligatory fronting of all wh expressions in others. We will also briefly consider two other language types, those in which multiple wh questions are unattested and those in which wh expressions are said to be optionally fronted. The section ends with a summary of some explanations that have been given for the existence of variation in (multiple) question formation across languages.

Let us begin with wh in-situ languages which have played a crucial role in the development of syntactic theory. We will base our survey of such languages on Chinese and Japanese. The

following illustrate simple multiple questions in the two languages respectively:⁵

- 27a. shei mai-le sheme (ne)
Who buy-ASP what Q
- b. dare-ga nani-o katta ka
who-NOM what-ACC bought Q
“Who bought what?”

- 28a. [_{CP} WH_i WH_j [_{IP} ...t_i ...t_j ...]]
b. [_{CP} Q_{ij} [_{IP} ...WH_i ...WH_j ...]]

The wh expressions are in their base positions here, a fact that could be demonstrated more clearly if arguments other than the subject were being questioned.⁶ Additionally, there is a morpheme identifying the clause as interrogative, either optionally or obligatorily. Broadly speaking, there are two proposals that have been made about the interpretation of wh in these languages. One, schematized in (28a), treats wh in situ as moving at LF in order to be interpreted. Under this view, fronting and non-fronting languages differ only in the level at which movement takes place. The other view allows wh expressions to remain in situ but interprets them via unselective binding, treating the Q-morpheme as the relevant operator. This is schematized in (28b).

I will begin by summarizing the key arguments for the movement approach, the reference point for which is Huang’s (1982) influential work on Chinese. It is obvious that superiority effects cannot be discerned in simple questions, but Huang argued that ECP is nevertheless operative in the interpretation of wh in situ. In a context like (29a) where the matrix verb selects a +WH complement, it is possible for the object wh in-situ to be interpreted with matrix scope, but not the adjunct. The availability of a wide scope reading for the embedded object also reveals the other important property of LF movement that Huang argued for, namely, the insensitivity to syntactic islands. This is further shown by the possibility of a direct question interpretation for (29b) where the wh expression is contained inside a complex NP:⁷

- 29a. ni xiang-zhidao [Lisi zeme mai-le sheme]
you wonder Lisi how bought-ASP what
“For what object x, you wonder how Lisi bought x?”
* “For what manner x, you wonder what Lisi bought in manner x?”

⁵ There is a typologically distinct type of wh in situ language, exemplified by many South Asian languages. See Case # 121 and the discussion in section 3 below.

⁶ In Japanese, it is possible to scramble the wh’s into sentence initial positions, but this is an option that is open to all arguments and can be shown not to involve wh movement.

⁷ Chinese does not show a subject-object asymmetry in extraction. This is explained as a consequence of the subject position being properly governed in Chinese.

- b. ni zui xihuan [piping shei de shu]
 You most like criticize who REL book
 “For which x, you like the book that criticizes x?”

Huang’s characterization of the interpretive possibilities of *wh* in situ has been challenged on the basis of Japanese by Nishigauchi (1986, 1990) and Choe (1984). Nishigauchi argues that although the Japanese counterpart of (29b) does have a direct question interpretation, the Japanese translation for (29a) can only be understood with both embedded *wh*’s taking narrow scope:

- 30a. Tanaka-kun-wa[Mary-ga doko-de dono hono-o katta]-ka sitte-imasu ka
 Tanaka -TOP Mary-NOM where which book-ACC bought Q know Q
 “Does Tanako know where Mary bought which book?”
 *”For which book x, does Tanako know where Mary bought x?”
- b. Kimi-wa [dare-ga kai-ta] hono]-o yomi masi-ta ka
 You-TOP who-NOM write book-ACC read Q
 “For which x, you read the book that x wrote?”

31. [Q_i [DP who_i [book that t_i wrote]]]_j you read t_j

Nishigauchi takes (30a) as evidence that the scope of *wh* in situ is constrained by subjacency. This is captured in the unselective binding approach he adopts by requiring *wh* in situ to be bound by the closest Q-operator. The apparent wide scope interpretation of (30b), he suggests, is due to local movement of the *wh* in situ to a DP adjoined position. Binding of the *wh* by the Q operator becomes possible after large scale Pied Piping of the DP, as shown in (31).⁸

The difference between Chinese and Japanese *wh* in situ has been studied by Aoun and Li (1993) and Tsai (1994) who argue in favor of unselective binding also for Chinese but locate the difference in the domain within which such binding must take place in the two languages. A different account of the facts is given by Watanabe (1992) who keeps to the view that LF *wh* movement is not subject to subjacency, positing S-structure movement of an abstract *wh* element in Japanese to account for the *wh*-island effect noted by Nishigauchi.

As we see, the languages under consideration make a significant contribution to our understanding of the syntactic principles governing scope assignment for *wh* in situ. The issue of whether Subjacency applies to *wh* in situ, even in languages like Japanese, is not a settled question. Watanabe (1992) and Dayal (1996) both note that conclusions about subjacency have to take into account the fact that introducing a *wh* in the matrix clause in structures like (30a) has a dramatic effect. (32), like its English counterpart, readily allows for an answer pairing individuals with books:

32. Dono sensei-ga [Mary-ga doko-de dono hono-o katta]-ka sitte-imasu ka
 Which professor-NOM Mary-NOM where which book-ACC bought Q know Q

⁸ As shown by von Stechow (1996), the syntax proposed by Nishigauchi is not adequate and yields incorrect answers unless obligatory reconstruction of the non-*wh* material is enforced.

”For which professor *x* and book *y*, does *x* know where Mary bought *y*?”

In section 3 we will return to the question of subjacency effects with *wh* in situ, from the perspective of a more refined understanding of the role of possible answers in diagnosing scope. The brief discussion here lays out the basic facts that have been reported in the literature.

1.2.2. Multiple Fronting Languages

Let us turn now to the other end of the spectrum, to languages with multiple *wh* fronting. The primary question posed by the possibility of multiple fronting is whether all the *wh*'s are in their final scope position in these languages or whether they move to peripheral positions within the clause. In (33a) all *wh* expressions are in SpecCP position, in (33b) they are adjoined to IP. (33c) represents a position in between where one *wh* occurs in Spec while the others are adjoined to IP:

- 33a. [_{CP} WH_i WH_j WH_n ... [_{IP} ... t_i t_j t_n ...]]
b. [_{IP} WH_i WH_j WH_n ... [_{IP} ... t_i t_j t_n ...]]
c. [_{SPEC-CP} WH_i [_{IP} WH_j WH_n [_{IP} ... t_i t_j t_n ...]]]

Various authors have argued for these possibilities, but we will base our survey of multiple fronting languages on Rudin (1988) who argues that multiple fronting languages instantiate either (33a) or (33c). Her study focuses on Polish, Serbo-Croatian, Czech, Bulgarian (all Slavic languages) and Romanian (a Romance language) (see also Comorovski 1989 and 1996). She claims that in Bulgarian and Romanian, all *wh* expressions move into Spec while in Serbo-Croatian, Polish and Czech one *wh* moves into Spec while the others adjoin to IP. She classifies the first set as +MFS (Multiply Filled Spec) language and the second set as -MFS language. I give examples below from Bulgarian to represent +MFS languages and Serbo-Croatian to represent the -MFS languages. I abstract away from a discussion of internal differences among the languages of each group:⁹

- 34a. Koj kogo vižda
Who whom sees
“Who saw whom?”
b.* Koj vižda kogo

35. Ko koga vidi
Who whom sees
“Who sees whom?”

⁹ The status of sentences with *wh* in situ in Serbo-Croatian *Ko vidi koga?* and other -MFS languages remains open. There is general unclarity whether *wh* in situ is possible, whether they are possible but only under a D-linked interpretation, or ruled out altogether. It can safely be assumed that there is, at the very least, a strong tendency or preference for multiple fronting in these languages.

Rudin identifies four properties distinguishing the two types of languages, which she claims follow from the proposed distinction in landing sites. Only +MFS languages allow multiple wh extraction from a clause and countenance wh-island violations. Only -MFS languages allow the insertion of clitics, parentheticals, adverbs etc after the first wh expression and allow the wh expressions to be in any order.

Consider the possibility of multiple extractions:¹⁰

36. Koj_i kude_j misliš [če e ostišul t_i t_j]
 who where think-2s that has gone
 “Who do you think that went where?”

37a. Ko_i zelite [da vam šta kupi t_i]
 Who want-2p to you what buy-3s
 “Who do you want to buy you what?”

b. Šta_i zelite [da vam ko kupi t_i]
 What want-2p to you who buy-3s
 “Who do you want to buy you what?”

38. [_{CP} WH_i WH_j ... [_{CP} [_{SPEC-CP} t'_i [t'_j]] ... t_i t_j ...]]

Since the matrix verb selects a -WH complement, both wh expressions must take matrix scope. This happens obligatorily at S-structure in Bulgarian. In Serbo-Croatian, however, only one embedded wh can be fronted. The other wh must occur in the preverbal focus position. This is analogous to English though it is possible to front the object wh as well as the subject wh without any apparent superiority effect. Rudin argues that languages that allow multiple wh's to adjoin to Spec-CP at S-structure (option 33a) provide an escape hatch for multiple wh expressions to move through without violating subadjacency, as shown in (38). Languages that adopt IP adjunction for all but one wh instead (option 33c) encounter a subadjacency violation if multiple wh's are extracted.¹¹

This difference in adjunction sites, she notes, also accounts for the fact that nothing can intervene between the first wh and other wh's in Bulgarian but may do so in Serbo-Croatian. In (39) the auxiliary obligatorily follows all the wh expressions. In (40), on the other hand, it must follow the first wh expression. Taking the auxiliary to be in C⁰ accounts for the distribution. Similar arguments are made on the basis of adverbials and parentheticals:

39. Koj (*e) kakvo (*e) na kogo e dal
 Who what to whom has given

¹⁰ In Romanian it is also possible for multiple extraction to take place from different clauses. That is, questions with the schema [_{CP} WH_i WH_j [_{IP} ... [_{CP} ... t_i ... [_{CP} ... t_j ...]]]] are considered acceptable.

¹¹ It should be obvious how multiple adjunction to Spec would account for wh island violations of the kind seen in (26a).

40. Ko je što (*je) kome (*je) dao
 Who has what to whom given
 “Who gave what to whom?”

Another related difference has to do with the order of fronted wh expressions in the two languages. The order of fronted wh expressions in Bulgarian is fixed. (39), for example, instantiates the only permissible sequence for this sentence. In Serbo-Croatian, on the other hand, the sequence in (40) is one of six possible orders. Rudin points out that the restriction in fronting seen in Bulgarian mirrors superiority effects in English and suggests that this is due to the fact that Bulgarian wh’s occupy the same position at S-structure that English wh’s occupy at LF, namely Spec CP. The first wh moves to Spec while the others adjoin to it, with adjunction being to the right. It is therefore expected that restrictions on fronting should be correlated in the two languages even though they differ in the level at which wh movement takes place. It is worth pointing out that the absence of superiority effects in Serbo-Croatian does not follow from its being a -MFS language. We will defer further discussion of this phenomenon till section 2. The brief survey here indicates the relevance of these languages to a theory of syntactic movement.

1.2.3. *Languages Without Multiple Wh Questions*

The survey of typological variation would not be complete without noting that multiple wh questions are not universally available. The two best-known examples of languages that do not allow multiple wh questions are Irish, discussed in McCloskey 1979, and Italian, analyzed by Calabrese 1984.

McCloskey notes that (41a)-(41b) are unacceptable in Irish and Calabrese gives (42a)-(42c) for Italian:

- 41a. * Cé aL rinne caidé
 Who COMP did what
 “Who did what?”
 b. * Caidé aL thug sé do cé
 Who COMP gave he to who
 “What did he give to whom?”
- 42a. *Chi ha scritto che cosa?
 Who has written what?
 b. *Chi è partito quando?
 Who left when?
 c. *Quale ragazza ha dato un bacio a quale ragazzo?
 Which girl gave a kiss to which boy?
- 43a. *MARIO ha scritto una LETTERA, (FRANCO un ARTICOLO, CARLO un LIBRO)
 MARIO has written a LETTER
 b. *FRANCO è partito alle CINQUE
 FRANCO has left at FIVE

- c. *SANDRA ha dato un bacio a CARLO
SANDRA gave a kiss to CARLO

According to McCloskey the impossibility of multiple wh in Irish is due to the fact that Irish interrogatives are not generated in argument position. Instead they select for S' (CP in current terminology) later followed by deletion of a pronoun in argument position. As he notes, however, this does not necessarily rule out multiple wh questions. To derive the absence of multiple wh questions, it would have to be stated that the selected S'/CP be non-interrogative.

Calabrese's account for Italian is different. He correlates the absence of multiple wh questions with the fact that Italian also disallows sentences like (43) which have multiple focus, sentences that, intuitively, would form appropriate answers to the questions in (42).¹² Calabrese's view is that wh expressions are marked with a focus feature, a feature that can only be assigned in the postverbal position in Italian. Thus, multiple wh questions are ruled out because such marking is necessarily restricted to one per clause.

McCloskey's and Calabrese's proposals each locate the absence of the relevant structures to language specific properties (we will discuss Adams (1984) below who has a different approach). Unfortunately, these proposals have not been pursued in subsequent literature and the theoretical implications of languages without multiple wh question remain largely unexplored. Legendre, Smolensky and Wilson 1998 discuss, for example, consequences of such languages for OT syntax. The question posed by the existence of such languages is what the optimal (hence grammatical) structure in Italian would be for the LF "Who ate what?". They point out that the impossibility of any such structure shows that members of the candidate set are not determined on the basis of shared LF's, as in Grimshaw 1997.

1.2.4. Optional Fronting Languages

A question that might be worth addressing here is whether there are languages in which wh fronting is optional. Though such languages have been reported, they do not intersect with strategies for multiple wh formation in interesting ways. A familiar example of an optional fronting language is French but multiple wh questions in French are formed analogously to English. Furthermore, the option of having wh in situ in single wh questions is restricted to root clauses. French is therefore typically classified as a fronting language. Examples of languages considered to have true optional fronting are Bahasa Indonesian (Saddy 1991), Egyptian Arabic (Wahba 1991) and Palauan (Georgopoulos 1991). However, the possibility of fronting in these languages is restricted to one wh expression and there appear to be no superiority type effects, typical of movement to Spec CP. Cheng (1991) provides evidence to argue that the apparent movement is, in fact, an instance of clefting and classifies these languages as in-situ. At any rate, no language has been reported in which there is the option of leaving all wh expressions in situ or moving all to Spec CP, that is, an optional multiple fronting language. We assume, therefore, that optional movement cross-

¹² It is generally accepted that this follows from the semantics of focus and questions. See Rooth 1992 for an explicit statement of the correlation between question denotations and focus in answers.

linguistically is either scrambling (which can affect all wh expressions) or clefting (which is restricted to a single wh expression). In either case, the principles governing such movement would not be limited to wh expressions but would apply to non-interrogative arguments as well.

1.2.5. Explanations for Typological Variation

In concluding our introduction to the multiple wh question phenomenon, let us consider how the attested variation in question formation strategies fits into the picture of universal grammar. As in other studies of typological variation, the goal is to derive differences on the basis of a few parametric choices.

The most worked out proposal has been made by Cheng (1991) and I will discuss that first. According to Cheng, all languages are required to distinguish interrogative clauses from non-interrogative ones at S-structure but they may differ with respect to how they do so. The Clausal Typing Hypothesis draws on the idea that a language may choose to mark interrogative clauses by the presence of a Q-morpheme in C^0 , or by movement of a wh expression to a Spec CP from where C^0 can be marked +WH via Spec-Head agreement. Combined with principles of economy, this predicts that languages with Q-morphemes will be wh in situ languages since there would be no motivation for overt movement at S-structure. Languages without such morphemes would be forced to move one wh to SpecCP. Languages like Chinese and English instantiate the two possibilities admitted by this approach. Multiple fronting languages like Bulgarian and Serbo-Croatian, however, are not predicted by this approach since only one wh movement can be forced by the requirement for Clausal Typing. An auxiliary explanation is therefore called for to account for movement of the other wh expressions.¹³

Cheng points to the fact that in the languages known to have multiple fronting, indefinites are formed by adding an affix to wh expressions. From this she concludes that wh expressions have a null interrogative determiner. Given that these languages lack Q morphemes one of the wh expressions must move to Spec CP for clausal typing. The motivation for movement of the other wh's, however, is independent of requirements of clausal typing. It is forced instead by the need for the null determiner to be licensed at S-structure, for which it must establish a local relation with the C^0_{+WH} . Adjunction to Spec CP or to IP both qualify as governed positions, accounting for Rudin's distinction between +MFS and -MFS languages.

Another proposal for the existence of different strategies in question formation is presented by Adams (1984) and, following her, Rudin (1988). This approach combines two standard views about parametric variation in wh movement. One idea is that all languages require wh movement for interpretation but they differ in the level by which such movement must take place, an idea for which Cheng's Clausal Typing Hypothesis provides the motivation. This allows for the broad distinction between wh movement and in situ languages. The other is that languages may or may not have a filter against multiple wh's in Spec, originally formulated as the filter against Doubly-filled

¹³ Cheng argues that the presence or absence of Q morphemes can be independently established on the basis of how yes-no questions are formed in a given language. Note that optional fronting languages, as mentioned in section 1.2.4 are argued not to involve wh movement at all, and are therefore not problematic for the Clausal Typing Hypothesis.

Comp. Though typically languages are expected to vary on this filter only at S-structure, Adams extends it further, arguing for a more general version of it: $*[_{\text{COMP/SpecCP}} \alpha \text{COMP/SpecCP}]$ at level X of the grammar. By barring adjunction at LF as well as S-structure, she derives the fact that languages like Irish and Italian do not have multiple wh questions and by allowing adjunction at S-structure and LF, the fact that Bulgarian and Serbo-Croatian require multiple fronting.¹⁴

Rudin, adopting Adams' proposal, notes that although it works straightforwardly for +MFS languages, the status of fronting in -MFS languages requires further elaboration. She takes the fact that only one wh moves into Spec to be evidence that such languages are like English, and unlike Bulgarian and Romanian, in enforcing the filter at S-structure. To drive the distinction between English which leaves the other wh's in situ and -MFS languages that require obligatory adjunction to IP, Rudin suggests that some languages may not allow LF movement to take place from A positions. While this is needed to complete the account, it should be noted that this part of the proposal remains quite speculative.

In concluding this section, I would like to point out that explanations for cross-linguistic variation would have to be modified once alternative scope assignment possibilities are admitted. For example, under Reinhart's proposal, non multiple wh languages cannot be derived by the Doubly Filled Comp Filter applying at S-structure and LF alone. Unless it is also stated that the language cannot interpret wh expressions via choice functions, the possibility of wh in situ remains open. In the case of -MFS languages, Reinhart's proposal could open up an alternative explanation for obligatory IP adjunction. If such languages had the Doubly-Filled Comp Filter at S-structure but not LF, it would allow movement of multiple wh at LF. Rudin's proposal that LF movement could only be from A' positions, forcing partial movement at S-structure, could be motivated by positing a language specific ban against choice functions. No doubt, more work would have to be done to see whether such an approach would hold up against other considerations, such as Reinhart's theoretical objection to LF movement within a minimalist framework. And at the empirical level, it would have to be shown that regular indefinites in these languages behave differently from languages in which choice functions exist, a not too promising prospect. My goal here in making these suggestions is not to present them as serious proposals but merely to indicate how developments in the semantics of questions can impact on typological explanations.

II: SUPERIORITY EFFECTS

2.1: Amelioration under D-linking

In the previous section, I isolated two empirical properties of multiple wh questions, Superiority effects and Subjacency violations, and outlined in broad terms the possible cross-linguistic variation in the formation of multiple wh questions. I would now like to take a closer look

¹⁴ Adams' filter applies to COMP, Rudin's to SpecCP. The differences between the two are not of direct concern here. I should also mention that the discussion of this filter has been somewhat simplified by ignoring differences between relative pronouns and interrogative pronouns and by ignoring the possibility of the constraint applying at PF, a point that is not adequately established.

at these empirical properties and their theoretical import. The goal of this section is to probe the nature of Superiority effects, leaving Subjacency violations for section 3. We will begin our survey of Superiority by discussing the notion of D-linking, a condition under which such effects seem to be canceled. The impact of D-linking is shown for languages like English where one wh expression is fronted, as well as for languages like Bulgarian where all such expressions move.

An important observation about Superiority is that the effects do not survive if the wh expressions are changed from a monomorphemic *who*, *what* etc to a complex phrase of the form *which N*. These effects are noted by Katz and Postal (1964) and Bolinger (1978) but the first extensive discussion is due to Pesetsky 1987:

- 44a. Who read what?
 - b. *What did who read?
 - c. Which book did which man read?

- 45a. Who did you persuade to read what?
 - b.* What did you persuade who to read?
 - c. Which book did you persuade which man to read?

A priori, the difference in acceptability could be due to a syntactic or a semantic difference. A possible syntactic difference might be that *which* is a specifier while *who/what* are heads. Pesetsky, however, rejects this possibility, pointing to the fact that *how many* which would also be classified a specifier, displays superiority effects:

- 46a. I need to know how many people_i t_i voted for whom.
 - b. * I need to know who(m)_j how many people voted for t_j.

Pesetsky argues that the difference in superiority effects is due to a lexical dichotomy between wh words, depending on whether they refer to a given set of entities in the context of utterance. While *which* phrases appear to bear the relevant relation to discourse inherently, monomorphemic wh expressions can lend themselves to the relevant D-(discourse) reading in appropriate contexts. When they do, Superiority effects are not manifested, as shown in the following (47a is from Bolinger):

- 47a. I know what just about everybody was asked to do, but *what* did *who* (actually) do?
 - b. I know that we need to install transistor A, transistor B, and transistor C, and I know that these three holes are for transistors, but I'll be damned if I can figure out from the instructions *where* what goes.

The correct generalization, then, is that the Superiority condition applies to non-D-linked wh only or, alternatively, that D-linked wh are exempt from the Superiority Condition. Pesetsky uses these facts to argue that the lexical distinction between two types of wh expressions correlates with a difference in the method of scope assignment. Non-D-linked wh expressions are assigned scope via movement and are therefore subject to normal constraints on movement. D-linked wh get scope by being bound by an unselective Q-operator (a la Choe and Nishigauchi) and are therefore not

regulated by the same constraints.¹⁵

The contrast in (44) for example is explained as follows:

48a. [_i who_i what_j [t_i read t_j]]

b. [Q_i which book_j [which man_i read t_j]]

c. *Nested Dependency Condition/Path Containment Condition:*

If two wh-trace dependencies overlap, one must contain the other.

In (48a) both wh's move, being non-D-linked, violating (48c). In (48b), however, there is only one dependency at issue. The other wh being interpreted in situ via unselective binding. Monomorphemic wh's, in contexts like (47), are interpreted similarly. Note that under this view, it is the D-linking of the wh in situ that is critical in the amelioration of Superiority effects. As pointed out to me by Chris Barker, however, (48b) is degraded if *which man* is replaced by *who* in (48b), a fact unexplained in Pesetsky's account.

Although the relevant distinction for Pesetsky is a semantic/pragmatic one, the fact that *which* NP's are exclusively D-linked provides a diagnostic which he uses to argue for a third type of wh expression. The fact that *the hell* in (49) is not compatible with *which* but can occur with *what* is presented as initial justification for the claim that they are aggressively non D-linked:

49a. What the hell book did you read that in?

b. * Which the hell book did you read that in?

c. ? What the hell book did which man read?

Given his claim that only D-linked wh can be interpreted in situ, the fact that aggressively non D-linked wh expressions exist provides a new tool for distinguishing moved vs. non-moved wh expressions in languages. (49a), under this view, would be a question requiring obligatory movement. It should be noted, however, that this kind of expression does not readily occur in multiple wh questions. (49c), to the extent that it is acceptable, does not lend itself to a multiple pair answer and has the feel of an echo question. This, as far as I can see, does not follow from the account presented. Thus some care needs to be taken in using such expressions to diagnose obligatory movement.

Pesetsky's identification of the role of D-linking in redeeming Superiority violations has been extremely influential in further developments in the theory of wh scope phenomenon. It is worth noting, however, that the aspect of his account that has held up is the correlation between *which* phrases and non-canonical scope, not the claim that such expressions are interpreted in situ. A potentially important piece of evidence with regard to the status of D-linked phrases and (non)-movement comes from multiple fronting languages. Pesetsky cites Wachowicz (1974)'s discussion of Polish as corroborating the prediction made by his theory that D-linked wh expressions may be

¹⁵ As a consequence of this, Pesetsky argues that Subjacency can be taken to be operative at LF as well as S-structure. Apparent violations of Subjacency would be due to a D-linked wh being bound long-distance by a Q-operator. Note, though, that Pesetsky's approach goes counter to Nishigauchi's claims regarding the scope of Japanese wh, an issue we return to in section 3.

exempted from fronting. But these facts have not gone unchallenged (see Cheng 1991 for discussion).

Doubts about a non-movement analysis for D-linked wh expressions surface when one considers Comorovski's (1989, 1996) discussion of multiple fronting language. Recall Rudin's observation that +MFS languages impose a strict restriction on the order of wh words (cf. 34 above) akin to the Superiority condition in English. Comorovski shows that this restriction is relaxed if the wh words are D-linked. The following Romanian examples illustrate this point:

50a. [Pe care_i cine_j [t_j l_i a văzut t_i]]

Which-Acc who him has seen

“Which one was last seen by whom?”

b. *[Ce_i cine_j [t_j a văzut t_i]]

What who has seen

51a. [Despre care_i cine_j [t_j ți a vorbit t_i]]

About which who to-you has told

“Who told you about which one?”

b. *[Despre ce_i cine_j [t_j ți a vorbit t_i]]

About what who to-you has told

Comorovski also notes that in Romanian (and Bulgarian), extraction out of wh islands are only acceptable under D-linking:

52a. [Despre care_i [știi [cine_j [t_j i a povestit t_i]]

About which you-know who to-him has told

“Which one do you know who told him about?”

b. *[Despre ce_i [știi [cine_j [t_j i a povestit t_i]]

About what you-know who to-him has told

It is clear from Comorovski's description of the facts, that D-linking is a significant factor in over-riding restrictions on fronting but not in making fronting itself unnecessary. This is corroborated in subsequent work on multiple fronting by Richards 1997. Of course, it should be noted that Pesetsky's claim that D-linked wh's can be interpreted via unselective binding without movement does not commit him to a non-movement analysis. He addresses, for example, the question of why English single wh questions require fronting by appealing to a language-specific need for the Q-morpheme to be triggered by the presence of a wh in Spec. Similarly, multiple fronting of D-linked expressions could be motivated on the basis of their status as polarity items requiring a local relation with the Q-operator, as proposed by Cheng (1991). However, facts like (50)-(52) do strike a cautionary note against a simple correlation between D-linking and non-movement. Setting this aside, the role of D-linking as a factor in the formation and interpretation of multiple wh questions is generally accepted as having cross-linguistic validity.

2.2. Explanations for Superiority

2.2.1. *Within the Principles and Parameters Framework*

With the caveat about D-linking in place, let us turn to explanations of the basic Superiority phenomenon. The first explanation is found in Chomsky 1973 who explained it by positing the Superiority Condition. In combination with strict cyclicity (53) derives not only the asymmetries observed in fronting but also wh-island effects:

53. Superiority Condition: No rule can involve X,Y in the structure ...X...[α ...Z...-WYZ...], where the rule applies ambiguously to Z and Y and Z is superior to Y.

Though the condition was essentially a description of the phenomenon in structural terms, it served to bring Superiority into theoretical center stage. Syntactic theory has since then grappled with this problem at every stage of its development. Instead of trying to faithfully summarize the many explanations that have been offered over time, I will try to highlight those analyses that have been particularly influential. Taking the Principles and Parameters model as a landmark, we can identify three distinct approaches to Superiority. As already discussed in section 1.1.2, a leading view of the phenomenon takes ECP as regulating fronting possibilities. Another proposal, discussed above (due to Pesetsky 1987), takes the relation between two dependencies as critical to the explanation. A third approach advocated by Aoun, Hornstein, Lightfoot and Weinberg 1987, relates superiority to Binding Theory.

The ECP-based account has already been shown to deliver the basic Superiority effect, namely that given a multiple wh question with one lexically governed argument and one not so governed, it is the latter that will be fronted. Here, I would like to indicate some further issues that arise as a consequence. The approach predicts, for example, that a multiple wh question with two wh expressions that are not lexically governed will be ungrammatical. The COMP-indexing mechanism will ensure that the overtly fronted wh will antecedent govern its trace but the other trace will necessarily remain ungoverned. This is shown in (54a)-(54b):

- 54a. *How did John go where?
- b. *Where did John go how?
- c. Who went where?

One problem that this approach encounters is the determination of the class of wh expressions that are lexically governed. In particular, the status of locative and temporal expressions, *where* and *when*, is an issue. The unacceptability of (54a) suggests that the trace of *where* could not be lexically governed, but there are other cases, such as (54c), where the explanation for acceptability turns on the trace being lexically governed. Attempts have been made to address this question but we will not discuss them at any length here, simply alerting the reader to the fact.

An interesting problem that surfaces under this approach has to do with the effect of additional wh expressions. They seem to improve the violation, even when that additional wh does not leave a lexically governed trace. This problem, noted in passing by Chomsky 1981, was discussed at some length by Kayne 1984. In fact, he discussed the effect of an additional wh in the

context of Superiority violations. (55a)-(55b) and (56a)-(56b) show normal constraints on fronting. (55c) and (56c) show the improvement due to the presence of an additional wh:

- 55a. I'd like to know who hid what there.
 b. *I'd like to know what who hid there.
 c. ?I'd like to know what who hid where.

- 56a. I'd like to know who hid it where.
 b. * I'd like to know where who hid it.
 c. ? I'd like to know where who hid what.

- 57a. Let $\beta_1 \dots \beta_j, \beta_{j+1} \dots \beta_n$ be a maximal set of empty categories in a tree T such that $\exists \alpha, \forall j, \beta_j$ is uniformly bound by α . Then $\{\alpha\} \cup (\cup_{1 \leq j \leq n} G_{\beta_j})$ must constitute a subtree of T.
 b. * [α where] [$_1$ who] [$_2$ [hid] [it]]
 c. [$_2$ [α where] [$_2$ [$_1$ who] [$_2$ [$_2$ hid] [$_2$ what]]]

According to Kayne, what is relevant in these cases is whether the fronted wh and the wh(s) in situ can satisfy connectedness, as defined in (57a). The unacceptable (56b), for example, has a representation as in (57b). The wh in situ, not being canonically governed, cannot form an appropriate sub-tree with the fronted wh. The additional wh in (57c) is governed by the verb. Being governed in the canonical direction for the language, its g-projection extends upwards, connecting with the fronted wh and properly including the subject wh in-situ. This is shown in (57c).

Kayne's augmentation of the ECP-based accounts remained till very recently the primary explanation for these facts. Note that this augmentation of the theory is distinct from the one argued for by Pesetsky (1987) in that it seeks to regulate ordinary wh expressions, not a special class of D-linked expressions. Similar augmentations would be required within theories adopting the Path Containment Condition (cf. 48c).

Aoun, Hornstein, Lightfoot and Weinberg (1987) propose an account that seeks to unify explanations for standard superiority effects and their absence with D-linked wh. It takes as its governing principle, the Generalized Binding Theory of Aoun (1986):

- 58a. An anaphor must be X-bound in its governing category.
 b. A pronominal must be X-free in its governing category.
 c. An R-expression must be A-free. (where X = A or A')

According to them, the critical difference between wh expressions that do and do not display Superiority is in their structural complexity. Arguing that QR moves only the wh quantifier, not its restriction, they posit the following representations:

- 59a. *What did who buy?
 b. [[$_i$ what $_i$ who $_j$] [t_j AGR $_j$ buy t_i]]
 60a. Which book did which woman buy?
 b. [[$_i$ which book $_i$ which $_j$] [[t_j woman] $_k$ AGR $_k$ buy t_i]]

Taking traces to be anaphoric elements subject to Principal A, they note an important difference between the two cases with respect to whether AGR counts as an accessible SUBJECT for the subject trace. In (59b) it clearly does, defining the governing domain within which the anaphor-trace must be A' bound. In (60b), however, AGR is indexed with the NP containing the anaphor-trace. The trace therefore does not have a governing domain and can remain free.

The Generalized Binding Account is meant to be a purely structural account of variation in superiority effects, potentially rendering Pesetsky's empirical generalization irrelevant to movement theory. As pointed out by Comorovski (1996), is it not clear whether this is so. She notes that cases claimed to show the irrelevance of non-D-linked Superiority violations, in fact, do carry presuppositions. Though the *wh*'s in (61) may not range over a set of familiar entities of the ordinary sort, they arguably range over a set of familiar kinds. She notes that a purely structural account of the phenomenon must meet the challenge of creating contexts showing this not to be the case. Recall that (46) above, where *how many N*, though complex, is shown to display superiority effects, makes the same point:

61. Which type of book does which type of man read?

2.2.2. *Economy-based Accounts*

A fundamental shift in the nature of the explanation for Superiority effects occurred with the advent of the Minimalist Program (Chomsky 1993, 1995) and Optimality Theory (Prince and Smolensky 1993). The guiding idea behind the Principles and Parameters model was that movement was always available unless such movement resulted in a structure that violated some principle of grammar. In contrast, the intuition informing Minimalism was to take movement as undesirable, unless forced by some principle. This conceptual shift is marked by the replacement of *Move α* by *Movement as Last Resort* as a central tenet of syntactic theory.

Consider the explanation for Superiority proposed in Chomsky 1993. Movement is forced only when the morphological features of an expression need to be discharged under a feature match with a functional category, as stated in the principle of *Greed*. Further, *Procrastinate* rates covert movement less costly than overt movement. Finally, *Shortest Move* requires movement to be to the closest potential landing site where *Greed* can be satisfied. This means that *wh* expressions, across and within languages, can differ in whether they have strong features that would force movement or weak features that could not trigger movement. Since only one *wh* expression is fronted in English multiple *wh* questions, it follows that only one *wh* expression carries strong features. The Superiority effect, then, reduces to a question about determining why subjects, rather than objects, in the basic case, must carry strong features.

Before we look at the specifics of the explanation, it may be worth noting that *wh* expressions that do not carry strong features will still need to be interpreted, either via covert movement or via choice functions, as argued by Reinhart 1997:¹⁶

¹⁶ *Wh* expressions can also be interpreted via skolem functions, which we discuss in section 2.2.3 under the term *functional wh*. The functional *wh* analysis is neutral with respect to whether it applies to fronted *wh* or to *wh* in-situ.

To return to the question of Superiority, one approach is to take *Shortest Move* to apply transderivationally. As elaborated in Kitahara 1994 and Oka 1995, the optimal candidate is chosen from a set of convergent derivations having the same numeration and the same LF. Under this view, (62a) and (63a) belong in the same reference set since they have identical LF's. It is obvious that *Shortest Move* is respected in the former but not in the latter, specially if the choice function analysis is adopted for interpreting wh in-situ. Note that this requires constraints to hold globally:

- 62a. [who_i [t_i saw what]]
 b. [what_j who_i [t_i saw t_j]]
 c. [Q_j who_i [t_i saw t_j]]

- 63a. *[what_j did [who see t_j]]
 b. [who_i what_j [t_i saw t_j]]
 c. [Q_i what_j [t_i saw t_j]]

An alternative possibility, mentioned in Chomsky 1993 and developed more fully in Chomsky 1995, is to put the burden for movement on the properties of the functional category which is the target of movement rather than on the moved expression. Given this view, wh expressions can be thought of as a morphologically unified class. We start, then, with a C⁰ that needs to attract a wh expression to its Spec in order to satisfy its requirements, in effect replacing the notion of *Move* with that of *Attract*. Now, there is a natural sense in which the subject can fulfil this requirement more economically than the object. The *Minimal Link Condition* is built into the definition of *Attract* in such a way as to capture this intuition. Note that this approach to Superiority does not rely on a global application of the constraints. This line of thought has been fully explored in Kitahara 1997.

Accounts of multiple questions within Optimality Theory also play on the tension between formal requirements forcing fronting, such as OP-SPEC, and a general preference against movement such as STAY (Grimshaw 1997). By ranking OP-SPEC > STAY, fronting of one wh expression is forced. She discusses the paradigm in (64):

- 64a. What will they put where?
 b. *What will where they put?

In (64a), OP-SPEC is violated by *where* while STAY is violated by *what*. In (64b), with *where* moving to an internal specifier position, *where* violates STAY in addition to OP-SPEC. (64a) is the optimal candidate since it involves one less violation of STAY. Although Superiority effects per se are not discussed there, it should be clear how optimal candidates can be identified on the basis of ranked constraints. Some interesting correlations between Minimalist accounts and OT accounts are made in essays contained in Barbosa et al 1998, in particular Ackema and Neeleman in addition to Legendre et al.

Ackema and Neeleman (1998), for example, derive the basic superiority effect through the interaction of three ranked constraints Q-marking (requiring a wh in Spec CP) » Stay (prohibition against movement) » Q-Scope (requiring all wh expressions to front). The following tableau shows how the explanation works:

65.

	Q-Marking	Stay	Q-Scope
☞ Who has seen what		*****	*
What has seen who		*****!	*
Who what has seen		*****!****	
Has who seen what	*!	***	**

It should be obvious how differences in ranking can result in different cross-linguistic possibilities for fronting.

There are further interesting developments within economy driven approaches that we will discuss in section 2.3 but before doing that I would like to turn to a somewhat different view of the meaning of wh expressions that has proved influential in recent literature and discuss its impact on studies of superiority.

2.2.3. Accounts Based on Functional WH

Another recent approach to Superiority effects takes its cue from the semantics of questions with quantifiers. An observation in the literature, dating back to Kiss 1993 and Comorovski 1989 is that multiple wh questions require a list answer in which every member of the subject term is paired but no such requirement is imposed on the object. Consider the following contexts given in Dayal 1996. In (66a) the fact that there is a woman who will not get paired does not seem to affect the felicity of the question, but in (67) the fact that a man will not get paired makes the question infelicitous:

66a. Speaker A: We're organizing singles tennis games between men and women. There are three men interested in playing against women, namely Bill, Mike and John. But there are four women interested in playing against men, namely Mary, Sue, Jane and Sarah.
Speaker B: So, which man is playing against which woman?

67a. Speaker A: We're organizing singles tennis games between men and women. There are four men interested in playing against women, namely Harry, Bill, Mike and John. But there are only three women interested in playing against men, namely Mary, Sue and Jane.
Speaker B: So, which man is playing against which woman?

Though the judgments are delicate there is a clear contrast which holds up against a wide range of data. This property of multiple wh questions has been explained by Hornstein (1995), Comorovski (1996) and Dayal (1996), working independently, in terms of Chierchia's (1993) explanation for the subject-object asymmetry in questions with quantifiers. A question with a universal quantifier in subject position allows functional and list answers while a question with the universal in object position allows neither:

- 68a. Which woman does every man like?
 b. (Every man likes) his mother.
 c. John likes Mary and Bill likes Sue.
- 69a. Which woman likes every man?
 b. *His mother (likes every man).
 c. *Mary likes John and Sue likes Bill.

Building on the view in Engdahl (1986) that wh expressions can quantify over functions from individuals to individuals, Chierchia argues that functional wh expressions leave a doubly indexed functional trace. The subscripted i-index identifying it with the wh operator is the functional variable. The superscripted a-index, which can be bound by a c-commanding argument, is an individual variable. Intuitively, the a-index corresponds to the pronoun in the functional answer and may be taken as having a pronominal character. The LF and interpretation for the functional reading of (68a), under this account, is given below:

- (70) a. [which woman_i [every man_j [t_j likes t_i^j]]]
 b. $\lambda p \exists f [\forall x (\text{woman}'(f(x))) \wedge p = \forall y [\text{man}'(y) \rightarrow \text{like}'(y, f(y))]]$
 $\Rightarrow \{\text{Every man likes his mother, Every man likes his wife, Every man likes his girlfriend}\}$

Wh quantification here is over variables whose possible values are functions to a set of women, such as *mother-of*, *wife-of* etc. The functional reading of the question denotes the set of propositions, each of which states that every man likes the individual he stands in some functional relation to. One of these propositions will be the true answer in a given context. The list answer, roughly speaking, is the graph of a function, listing for each member of the domain set the individual who stands in the relevant functional relation to that individual. We will defer further details of the implementation till section 3, except to note that the universal term which provides the domain set, has to take scope outside the propositional variable p. Turning to (69), Chierchia notes that for functional binding to take place an LF like (71a) in which the quantifier c-commands the wh trace would be needed. (71a) is ruled out as a case of weak crossover violation since the variable t_j is coindexed with a pronoun to its left, the a-index of the wh term:¹⁷

- (71) a. *[which woman_i [every man_j [t_i^j likes t_j]]]
 b. [which woman_i [t_i [every man_j [likes t_j]]]]

¹⁷ The a-index may be anaphoric instead of pronominal in some languages. It should be noted that Chierchia's analysis of functional/list answers is not tied to any particular account of WCO. Two proposals treating it in terms of leftness and relating it to functional answers are Williams (1994) and Jacobson (1994). The latter argues for a variable-free semantics to get the same results as Chierchia. Other references for WCO are Koopman and Sportiche (1982) and Safir (1986).

The well-formed LF in (71b) has the quantifier adjoined to VP from which no functional binding can take place. The absence of list answers to questions with quantifiers in object position follows since list answers under this approach are derivative on functional answers. The only available answer here is therefore the individual answer.

Hornstein, Dayal and Comorovski each point out that if multiple *wh* questions can encode the kind of functional dependencies that questions with quantifiers can, in a question with two *wh* expressions, the *a*-index of one *wh* can be bound by the other. Taking an example like (72a), there would be two potential LF's for it, only one of which would be syntactically well-formed:¹⁸

- (72) a. Which man saw which woman?
 b. $[_{CP}[_{wh} \text{woman}_j [_{CP}[_{wh} \text{man}_i [_{IP} t_i \text{ like } t_j^i]]]]]$
 c. $*[_{CP}[_{wh} \text{woman}_j [_{CP}[_{wh} \text{man}_i [_{IP} t_i^j \text{ like } t_j]]]]]$

In (72b) the subject term binds the *a*-index of the object term and there is no problem with this binding. In (72c) the object term binds the subject term but this involves a WCO violation since *which woman* crosses over the pronominal *a*-index of *which man* in order to bind it. This provides a straightforward account of why it is the subject *wh* has the “universal force” identified by Kiss and Comorovski. Dayal 1996 also provides evidence showing that in languages where scrambling overcomes WCO violations, it is possible to get the universal reading for the scrambled object.¹⁹

It is against this background that Hornstein and Comorovski develop their accounts of Superiority. Hornstein presents a number of cases showing the correlation between superiority effects and WCO as support for the general approach. The explanation can be demonstrated on the basis of a basic Superiority case like (73a). Given the Minimalist framework, the fronted *wh* is assumed to leave a copy in base position and one of them must be deleted before interpretation:

- 73a. What_i [who buy what_i]?
 b. What_i [who_j^i buy t_i]
 c. Who_i [t_i buy what_j^i]

¹⁸ Note that May 1985's account of this asymmetry does not work for multiple *wh* questions. For him, the (im)possibility for scope permutation turns on the landing site of the scopal expressions, rather than on their base positions. Two *wh*'s, unlike a *wh* and a quantifier, can form a local relation in Spec CP without violating the PCC. An object *wh* is therefore predicted to be able to take scope over the subject and (somehow) be interpreted universally. See Dayal 1996 for discussion.

¹⁹ See also Williams (1994) for similar ideas and related discussion.

- 74a. [What [who bought what]
 b. [[who bought what]
 c. [who [who bought what]
 d. [who [t bought what]

Hornstein assumes that wh's that bind syntactic variables can be interpreted as ranging over individuals, thereby providing the domain of the function. Full copies are interpreted functionally. Translating this into the current representation, this allows for two interpretive possibilities for the structure in (73a). The possibility in (73b), where we have a function from things to individuals who bought those things is ruled out as a WCO violation. The question that remains for Hornstein is why the possibility in (73c) is ruled out. This would require the steps in (74), deletion of the moved wh (74b), movement of the in situ wh (74c), followed by deletion of its copy (74d). But, as he notes, a structure respecting superiority would yield this result more economically since it would require only two steps (74c and 74d). Thus, (73a) is either ruled out by WCO (74b) or blocked by a more economical derivation. There are further possibilities within the minimalist assumptions for deriving this effect considered by Hornstein, which I will not go into here.

Now, consider what happens when the wh expressions are D-linked. (75b) shows that deleting the copy of the fronted wh and interpreting the expression in Spec leads to the familiar WCO violation, analogous to (73b) above. The key difference here, according to Hornstein, is that D-linked wh's do not have to be fronted in order to be interpreted as the domain term. Thus the derivation in (76) is viable. Crucially, there is no difference in the number of steps needed to interpret the structure regardless of whether the subject or the object is fronted:

- 75a. [which book [which man review which book]]
 b. *[which book_j [which man_i^j review t_j]]

- 76a. [which book [which man review which book]]
 b [[which man_i review which book_j¹]]

Comorovski's account starts from a different premise. She points to the fact that a list answer to a multiple wh requires the members of the domain set to be given. This says, in effect, that at least one wh expression must be D-linked. Two questions arise at this point. Why is there no discernible D-linking for subjects in multiple wh questions respecting Superiority? What role does D-linking play in Superiority violations? Comorovski draws on the idea in Kuno (1982) that a multiple question is a request for information about pairs, where the question provides a sorting key for arranging that information. In general, subjects tend to carry old information and are natural sorting keys.²⁰ Fronting of an object wh is a signal for a change in the structurally encoded criterion for sorting. In functional terms, this means that the object wh sets the domain of the function:

²⁰ See also Krifka 1999 for a similar explanation of the subject-object asymmetry in questions with quantifiers.

- 77a. Which book did which man recommend?
 b. [which man_j [which book_i [t_jⁱ recommend t_i]
 c. [Its_i author] recommended each book_i / Each book was recommended by its author.

It remains to be explained, then, why there is no perceived WCO effect. Comorovski notes that backward anaphora is generally possible with universals like *each N*. What *each*-universals and *which NP*'s share is the fact that they both carry existential presuppositions, the intuition driving the term D-linking. Taking the presuppositional nature of a *wh* term to be criterial in over-riding WCO, Superiority violations are claimed to be predictable on the basis of their possible functional answer.

While Hornstein and Comorovski both take the functional WH approach as a key element of the explanation, they end up with very different results. For Hornstein, the subject that has been crossed over by a D-linked *wh* remains the domain term but not for Comorovski. This difference is, in principle, testable. However, it should be noted that Hornstein's conclusion about the nature of the pairings permitted in a question with two *which N*'s does not accord with the intuitions reported in (66)-(67) above. For him, two D-linked phrases require the members of each term to be exhaustively paired so the direction of functional dependency is not predicted to make a difference. Comorovski, however, does predict that in acceptable cases of Superiority violation, it will be the object *wh* that is exhaustively paired while the subject *wh* may or may not be. In unpublished work Dayal 1994 reported that Superiority violations do result in a requirement of exhaustivity for the object *wh* but the judgments for the subject term were unclear. Similar effects seem to obtain for scrambled versions of multiple *wh* questions in languages like Hindi and Japanese, which is suggestive that the same points may be at issue (Dayal 1996). It would be fair to say, however, that the semantic consequences of Superiority violations remain an open issue in functional accounts of multiple *wh* questions as well as in other accounts.

2.3. Typological Variation in Superiority Effects

2.3.1 Superiority in Multiple Fronting Languages

Having surveyed a range of explanations for Superiority effects in English, let us return to the issue of ordering in multiple fronting languages. Recall that the basic generalization in Rudin was that in +MFS languages, where all movement was to Spec CP, there was a strict ordering among the fronted *wh* while in the -MFS languages the order was free. This issue is taken up in Richards 1997 who proposes an explanation within the Minimalist framework.

Richards takes Bulgarian multiple *wh* questions such as (34), repeated below in (78), as his starting point. As shown in (78a), the first step in the derivation involves movement of the highest *wh* as an instance of *Attract Closest*. This is analogous to the English Superiority case. The fact that the second *wh*, when it moves, is "tucked in" as in (79b) rather than extending the tree with an outer specifier as in (79c), is argued to be evidence that *Shortest Move* is in play:

- 78a. Koj kogo vižda
 Who whom sees
 "Who saw whom?"
 b.* Koj vižda kogo

- 79a. [_{Spec} koj_i [t_i vižda kogo]]
 b. [_{Spec-1} koj [_{Spec-2} kogo_j [t_i vižda t_j]]
 c.* [_{Spec-2} kogo_j [_{Spec-1} koj_i [t_i vižda t_j]]

Bosković 1997, however, extends Rudin’s paradigm by noting that although there are strict restrictions on the leftmost wh, the order of other wh expressions in Bulgarian is relatively free (see also Comorovski 1996). Thus putting the direct object before the indirect object does not lead to ungrammaticality in (80b). The order of the tucked in wh expressions appears to contradict *Attract Closest*:

- 80a. Koj na kogo kakvo dade?
 Who to whom what gave
 b.? Koj kakvo na kogo dade?

Richards explains this paradigm on the basis of a *Principle of Minimal Compliance* which he describes using the following metaphor. In a given structure, if one expression pays the taxes imposed on movement within that structure, other such expressions are exempted from paying those taxes. In this case, the second and the third wh expressions are exempted from *Attract Closest* since the dues are paid by the first wh. As Richards notes, multiple fronting languages provide strong empirical argument against the claim in Chomsky 1995 that each overt operation must extend the tree at its root.

Given that Richards argues for “tucking in” of wh expressions, the issue of Superiority violations under D-linking becomes relevant. He posits a Topic phrase above CP and allows D-linked wh expressions to optionally have matching morphological features. This makes it possible for a D-linked object to move to the left of a subject.

Picking up on these issues, Pesetsky 2000 proposes further refinements in the theory. According to him, wh expressions may undergo three distinct types of movement, overt phrasal movement, covert phrasal movement and feature movement. Adopting a single output syntax, Pesetsky locates the difference between overt and covert movement in pronunciation rules which target either the head or the tail of a movement chain. Feature movement, on the other hand, is the attraction of a morphological feature from within a phrase by some category. The full phrase, because it does not undergo any movement, is necessarily pronounced in situ. Crucial evidence from ACD phenomenon is presented for this three-way distinction but for reasons of space I will not go into those arguments here, confining myself to demonstrating how typological variations in multiple wh questions is derived under this approach. There are two other pieces that play a role in the explanation. Languages can differ in having multiple Spec Complementizers or single Spec Complementizers and only phrasal movement can satisfy multiple Spec requirements. In multiple interrogatives with D-linked wh’s this requirement is suspended or alternatively, D-linking is compatible with single Spec Complementizers. Pesetsky leaves this open.

Superiority effects in non D-linked questions and their absence in D-linked questions are explained as follows:

- 81a. Who C_{m-SPEC} [_ saw whom] *phrasal movement of who due to C_{m-SPEC}*
 b. Who whom C_{m-SPEC} [_ saw _] *phrasal movement of whom due to C_{m-SPEC}*
 c. Pronunciation in English: S V O
 d. Pronunciation in Bulgarian: S O V

- 82a. $C_{1-SPEC(f)}$ [f-which woman saw which man] *feature movement of which woman due to C_{1-SPEC}*
 b. Which man $C_{1-SPEC(f)}$ [f-which woman saw _] *phrasal movement of which man due to C_{1-SPEC}*
 c. Pronunciation in English: O S V
 d. Pronunciation in Bulgarian: O S V

In (81) phrasal movement of both wh's is forced by the multiple Spec requirement. This is suspended in the case of D-linking (82), making it possible for the subject to undergo feature movement. As would be obvious, the explanation for the absence of Superiority effects turns on the role of D-linking in allowing multiple wh interrogatives with single Spec Complementizers. As noted by Pesetsky himself, one weakness in the proposal is that the crucial connection between the semantic property of D-linking and syntactic scope mechanism remains as elusive as in earlier accounts of the phenomenon.

Pesetsky also accounts for the fact that Superiority Effects disappear in the presence of extra wh's even without the benefit of D-linking. The relevant steps for a representative case are as follows:

- 83a. $C_{m-SPEC(f)}$ [f-who put what where] *feature movement of who due to C_{m-SPEC}*
 b. What $C_{m-SPEC(f)}$ [f-who put _ where] *phrasal movement of what due to C_{m-SPEC}*
 c. What where $C_{m-SPEC(f)}$ [f-who put _ _] *phrasal movement of where due to C_{m-SPEC}*
 d. Pronunciation in English: What did who put where?

Crucially, there are two other wh expressions that can undergo phrasal movement to satisfy the C_{m-SPEC} requirement, leaving the subject wh to undergo feature movement.

More work has been done on Multiple Fronting languages, such as Grewendorf (2001), that are worth exploring but in the interests of space I will end my discussion here and turn briefly to a different problem having to do with the cross-linguistic applications of explanations for Superiority.

2.3.2. Languages without Superiority Effects

As would be obvious, the new typology of wh movement opens up new ways of explaining typological variation and in this concluding subsection I will outline how languages which do not show Superiority Effects are accounted for. One such case, mentioned in section 1, is the class of languages classified -MFS by Rudin. Another is German. For a -MFS language like Polish, Richards 1997 takes Rudin's general approach that they do not involve movement to Spec of CP. In Richards' terms, this means that multiple specifiers at CP level, with their characteristic "tucking in" effect, are not created in languages like Polish. Instead, apparent cases of multiple wh fronting are essentially movement to multiple specifiers of IP. Since each such position is equidistant to C,

Attract Closest is satisfied by movement of any wh to Spec CP position.

German displays a different paradigm in that both SVO and OVS order are acceptable, with S and O as wh expressions, without the benefit of D-linking. Pesetsky explains this by locating the difference between English and German in the possibility of C_{m-SPEC} . According to him, German lacks this possibility. Then it follows that in a multiple wh question, only one wh needs to move to fulfill the requirements of C_{1-SPEC} . The other wh necessarily undergoes feature movement, making a derivation analogous to (82) acceptable.

This account of the absence of Superiority Effects presents an alternative to an earlier proposal by Fanselow (1991, 1997). His basic idea was that the apparent Superiority Violation in German OSV is preceded by scrambling of O to pre-S position. Wh movement from a scrambled structure will respect economy only if the Object is moved. Fanselow draws on the following contrast as further evidence:

- 84a. Wer glaubte, dass der Peter ihr wen vorstellte
Who-ACC believed that the Peter-NOM her-DAT whom-ACC introduced
“Who believed that Peter introduced whom to her?”
b. ?* Wen glaubt wer, dass det Peter ihr vorstellte

In (84b) the embedded object cannot occur to the left of the matrix subject. This is because to do so would first require the object to scramble long-distance to the pre-subject position and then wh move to Spec C. Fanselow shows that such long distance scrambling is not available out of finite clauses, providing independent evidence for the proposed reliance of Superiority violations on scrambling possibilities. This proposal has been critiqued by Pesetsky, to which I refer the reader.

We have seen that fronting possibilities in multiple wh questions have played a critical role in the development of syntactic theory. Within the Principles and Parameters model they provided evidence for the ECP and within Minimalism and Optimality Theory they are shaping the view of economy/optimality in language. While the issues touched upon in this section raise further questions worthy of exploration, we will conclude our survey of the theoretical issues related to the Superiority phenomenon and turn to the Subjacency phenomenon in the next section.

III: SUBJACENCY AND WH IN-SITU

3.1. The Diagnostic of Possible Answers

3.1.1. *Single vs. Multiple Pair Answers*

In moving from a survey of issues surrounding Superiority to a survey of issues surrounding Subjacency, we shift from direct evidence in the form of grammaticality judgements about fronting to indirect evidence about intuitions regarding possible answers. It would, therefore, be appropriate to begin by taking a closer look at the diagnostic of possible answers. As mentioned in section 1, a standard assumption in the literature is that possible answers specify values for all and only those wh expressions that take matrix scope. In the case of simple single wh questions, the answer specifies values for one expression, in the case of multiple wh questions for two or more, as the case may be.

Given this assumption, the fact that questions like (15) and (17), repeated below, give values for more than one wh expression is indirect evidence that the wh expression inside the island is able to take matrix scope in violation of Subjacency:

85a. Which person knows the man who wrote which book?

b. Bill knows the man who wrote Aspects.

86a. Which student knows where Mary bought which book?

b. Bill knows where Mary bought Aspects and Sue knows where Mary bought LGB.

In Dayal (2002) I note that although both multiple wh questions allow a paired answer, the form of the pairing is different in the two cases. (85a) only lends itself to a single pair answer while (86a) can readily allow a multiple pair answer. A note of clarification about the nature of this contrast is in order. It may be possible to force a multiple pair answer to (85a) by constructing specific contexts. The point to note is that nothing particular needs to be done to get such answers for (86b). With this caveat in mind, we can see that similar scope taking mechanisms violating Subjacency could not be involved in both constructions. Before looking at the Subjacency cases, however, we will first examine the distinction between single pair and multiple pair answers in the basic cases.

Multiple wh questions have been said to have three different readings, a list reading, a REF-Q reading and an echo-Q reading. These readings were identified by Bolinger (1978), Wachowicz (1974, 1975) and Pope (1976), under various terms. The summary here is taken from Dayal (1996), which relies on their observations as well as on the discussion in Comorovski (1989). To see the difference between the three, consider three situations in which the multiple wh question *Who cooked what?* can occur and the answers it admits in those situations.

Let us demonstrate the list reading first. Take a context in which there are several dishes on the table and the questioner knows several people who have cooked dishes. She asks (87a) and gets the response in (87b):

87 a. Who cooked what?

b. John cooked the meat, Bill cooked the rice and Sue cooked the vegetables.

As Wachowicz puts it, the information the questioner is interested in is the proper pairing between two given sets. This reading crucially presupposes that there will be at least two pairs in the list. This is what we are calling the multiple pair reading.

Now, let us turn to the echo reading of the question. An appropriate context is one where the questioner only hears part of an utterance and wishes the utterance to be repeated. In the following, subscripting is supposed to indicate a string that is phonetically unclear:

88 a. _{john} cooked _{the meat}

b. Whó cooked whát?

c. John cooked the meat.

Echo questions can be distinguished from ordinary questions by their rising intonation.

Finally, consider REF-Questions, which asks for the identity of a unique pair of individuals, as opposed to a list of pairs. (88b) uttered in response to (88a) would have a REF-Q reading:

- 88a. He cooked something.
b. Who cooked what?
c. John cooked the meat.

REF-Questions were classified by Wachowicz as a type of echo question but Pope shows that though they are similar to echo questions they are intonationally distinct. The *wh* expressions do not need emphatic stress and are uttered with the same fall in intonation as ordinary questions. Further, they are not requests to repeat the information already provided in the discourse, but rather for elaboration.

That REF-Questions differ from echo questions can also be shown by cases that are clearly not triggered by previous utterances. (89a) and (89b) can be asked without a particular context and be answered with just a single pair of individuals. All that is required is that the questioner know that only a single pair can be named in the answer. (89a), for example, pragmatically rules out a multiple pair answer. Similarly, if there are only two individuals, only one of them can be the first to hit in (89b):

- 89a. Who killed Robert Kennedy when?
b. OK, who hit who first?

There are two approaches that have been taken to the distinction between single and multiple pair answers. One is to treat it as a pragmatic effect. Under this view, questions are able to yield multiple-pair answers but they are not restricted to them. Single-pair answers are simply a sub-case of the larger possibility. The other approach is to take this distinction as arising from the semantic representation of the question and in the next subsection we will look at the various proposals that have been made to derive this distinction. The reason for taking the semantic approach to this phenomenon seriously is that there are empirical effects that cannot be handled in a pragmatic approach. We have already seen evidence from two different domains about the nature of restrictions governing possible answers. In section 2.2.3 we saw that proponents of the functional *wh* approach have appealed to the intuition that in a multiple *wh* question only one *wh* expression, the superior one, is exhaustively listed and at the beginning of this section we saw that there are constructions where a pair answer is possible, but not a multiple-pair answer. Once we understand the semantic moves involved in deriving the two possibilities, we will re-examine the phenomenon of Subjacency violations with *wh* in-situ.

3.1.2. Deriving Single and Multiple Pair Answers

Higginbotham and May 1981 were the first to attempt an account of pair-list answers by proposing a semantic operation they termed *absorption*. The goal of this operation was to provide a transition from uniqueness effects in single *wh* questions like (90a) to multiple-pair answers in (90b). Although their proposal has been widely accepted in the syntactic literature, the details of the semantic operation are somewhat difficult to decipher:

- 90a. Which man came to the party?
 b. Which man knows which woman?

In this section I will present the account of the switch from uniqueness to list answers proposed in Dayal 1996, which has the advantage of being embedded in the propositional theory of questions introduced in section 1.

In Dayal 1996 I based my account on the paradigm below, where (91a) requires a unique individual to be named in the answer, (91b) a plurality of individuals while (91c) is neutral in this regard:

- 91a. Which man came to the party?
 b. Which men came to the party?
 c. Who came to the party?

- 92a. $\lambda p \exists x [\text{man/men}(x) \ \& \ p = x \text{ came to the party}]$
 b. {a came, b came, c came};
 c. {a came, b came, c came, a+b came, a+c came, b+c came, a+b+c came}
 d. $\text{Ans}(Q) = \{p \in Q \wedge \forall p' \in Q [p' \rightarrow p \subseteq p']\}$

Adopting the view that the domain of individuals includes atomic entities as well as their sums, I make the existential wh quantifier sensitive to number morphology. Now, in the singular case, the question denotes the set in (92b), in the plural case the set in (92c).²¹ Further, I define the Answerhood operator, as shown in (92d), to pick out the maximally informative true proposition in the set, namely the one that entails all the others. In the singular case, there can be no entailment relation between the various members of the set so the question can only be answered if a unique proposition is true. In the plural case, however, there can be more than one true proposition since their conjunction can satisfy (92d). Although the semantics allows there to be only one true proposition in the set, say *a came*, there is assumed to be an implicature that a plurality of individuals have the property in question. *Who* is taken to be like the plural case minus the plurality implicature.

Turning to multiple wh questions, it should be obvious that (92d) if applied to the Hamblin sets, presented in section 1 for example, will yield single pair answers. In section 2.2.3 we discussed the functional approach to multiple wh questions, advocated by Hornstein 1995, Dayal 1996 and Comorovski 1996. As argued by Dayal 1996, the functional approach only derives the intended effects if two conditions are added to the proposal of Chierchia 1993. One is the Answerhood operation introduced above, the other is the requirement that each proposition in the set be a full graph of the function. Let us consider this with the help of a concrete example:

²¹ See Link 1983 and the account of comparable uniqueness and maximality effects with definite determiners. This correlation between wh expressions and definites derives from an earlier version of the theory, presented in Srivastav 1991. It has also been proposed by Rullmann (1995).

- 93a. Which man_i likes which woman_jⁱ ?
- b. [Which woman_j [Which man_i [t_i likes t_jⁱ]]]
- c. $\lambda p \exists f_{\langle e, e \rangle} [\text{Dom}(f) = \text{man} \ \& \ \text{Range}(f) = \text{woman} \ \& \ p = \bigcap \lambda p' \exists x [p' = \text{like}(x, f(x))]]$
- d. {a likes c and b likes d,
 a likes c and b likes c,
 a likes d and b likes d,
 a likes d and b likes c}

The question is analyzed as involving a functional dependency respecting WCO, as in (93a). At LF we take the wh in-situ to undergo A' movement to Spec CP and the logical representation is as given in (93c). Without going into the details of the derivation, a functional dependency is posited whereby the subject term sets the domain of the function, the object term its range and the propositions relate individuals in the domain set with some functionally dependent entity in the range set. Crucially, each proposition in the denotation of the question is a conjunction of the atomic propositions obtained by varying the values of x and its dependent element f(x) in the IP denotation. If the set of men = {a,b} and the set of women = {c,d}, the question denotes the set of four propositions in (93d), where each proposition is the graph of a function relating the domain and the range sets. Although the particulars of this operation differ from that proposed by Higginbotham and May, this too can be thought of in terms of *absorption* of two wh operators.

The Answerhood operator in (92d), when applied to sets derived via *absorption*, picks out a unique maximally true proposition from the question denotation. As would be obvious, each possible answer ensures that the domain set is exhaustively paired and that the pairings are one-one or many-one. This is because although *a and b both like c* is acceptable as an answer to (93a), *a likes c and d* is not (see Engdahl 1986 and Dayal 1996 for further discussion of this issue). The ambiguity of multiple wh questions in allowing single pair vs. multiple pair answers, on this account, rests on the lexical ambiguity of wh expressions in being ordinary individual quantifiers, or expressions that can enter into functional dependencies.²²

Note that there are two key factors in deriving the desired effects, each proposition in the set must represent the full graph of the function and only one proposition must be picked out as a possible answer. As shown in Dayal 1996, Chierchia's analysis of questions with quantifiers is unable to derive appropriate answers because it allows atomic propositions in the set and takes the conjunction of any subset of the question as a possible answer. This problem is inherited by Comorovski 1996 who applies Chierchia's analysis to multiple wh questions:

- 94a. $\lambda p \exists x \exists f_{\langle e, e \rangle} [\text{man}(x) \ \& \ \text{Range}(f) = \text{woman} \ \& \ p = x \ \text{like} \ f(x)]$
- b. {a likes c, a likes d, b likes c, b likes d}

Comorovski's account, like Chierchia's as well as Engdahl's, yields sets of atomic propositions no different from ordinary Hamblin sets and does not impose restrictions on how many of them would

²² See also Sharvit 1997 and Bittner 1998 for further discussion of functional dependencies in questions.

enter into a possible answer. This predicts, in the case of (94) for example, that *a likes c and d* will be a possible answer, contrary to the stated goals of the analysis.

Recall that Hornstein's functional account differs from Comorovski's in not requiring covert movement of *wh* in-situ. It is somewhat hard to see what question denotations are at issue since he does not give a semantics for the structures. However, let me highlight a point that may be crucial in evaluating Hornstein's account. As we saw in (93), the relevant operation that delivers multiple pair answers is only possible if both *wh* expressions take scope outside C^0 , where the propositional variable is introduced. That is, the LF representation of the multiple *wh* question must be as given in (93b). In this respect, his syntax for multiple pair answers is no different from the syntax proposed for deriving standard Hamblin sets. Single pair answers arise when the Answerhood operator applies to question denotations derived from *wh* expressions interpreted as ranging over individual variables, multiple pair answers from its application to question denotations where *wh* expressions encode a functional dependency. For Hornstein's account to be complete, a semantics deriving the intended answers has to be given, which at this point has not been done.

3.2. Pair Answers Across Islands

3.2.1. Single Pair Answers Across Islands

In this section we will summarize the proposal in Dayal (2002), which is partially based on Dayal (1996), for deriving the difference between the types of pair answers possible when the *wh* expressions are separated by an island. The primary empirical generalization there is that multiple pair answers are restricted to *wh* islands that conform to a particular format dubbed *the wh triangle*, with single pair answers available in the other cases. We will begin by summarizing the account for single pair answers across islands before considering the role of the *wh triangle*.

It might be worth recalling that the possibility for *wh* in situ's to take scope outside islands has been explained in a number of ways, by taking Subjacency to be inapplicable at LF (Huang 1982), by unselective binding of D-linked *wh* expressions (Pesetsky 1987) and by existential closure involving a choice function interpretation for *wh* in-situ (Reinhart 1997). None of these proposals, however, have taken the distinction under discussion into account so it may be worth demonstrating this explicitly. While (95b) is a perfectly acceptable answer to (95a), (95c) is at least an awkward, if not a clearly unacceptable answer to it:

- 95a. Which linguist will be offended if we invite which philosopher?
b. Professor Smith will be offended if we invite Professor Brown.
c. #/*Professor Smith will be offended if we invite Professor Brown, and Professor King will be offended if we invite Professor Matthew.

This also holds of multiple *wh* questions such as the following, cited commonly in the literature as showing wide scope of *wh* in-situ:

- 96a. Which student read the book that which professor wrote?
b. Which student got a headache after she read which book?

In order to account for this fact in Dayal (2002) I note that the operation for forming multiple pair answers given in section 3.1.2 must be blocked from holding across islands. I constrain the *absorption* associated with functional dependencies from holding across islands by requiring movement at any level to obey standard constraints. Since the *wh* in-situ cannot move to matrix spec, it cannot fix the value for the range of the function and its trace cannot be functionally bound by the matrix *wh*.

The question then arises how the single pair answer can be derived without actual movement. Here I appeal to Reinhart's choice function analysis which is designed for this very purpose:

- 97a. [Which linguist_i [t_i will be offended if we invite which philosopher]]
 b. $\lambda p \exists x \exists f_{\langle\langle e, t \rangle, e \rangle} [\text{linguist}(x) \ \& \ p = x \text{ will be offended if we invite } f(\text{philosopher})]$
 c. {a will be offended if we invite c, b will be offended if we invite d, a will be offended if we invite d, b will be offended if we invite c}

As shown above, at the matrix level there is existential quantification over functions that yield an arbitrarily chosen entity out of the set they are applied to. Since the LF's of such questions do not have two *wh* operators that can undergo *absorption*, we get as the denotation of the question a set of atomic propositions. The Answerhood operator, when applied to such question denotations, yields single-pair answers. The choice function approach, then, derives what could be described as a non-quantificational or referential interpretation of the *wh* in-situ (see Saddy 1991 for a similar characterization of *wh* in-situ).

It may be worth noting that choice functions are not blocked from applying to *wh* in-situ in simple questions or in the *wh*-island cases to be discussed below. The claim is that they cannot be used to derive multiple pair answers. Since a Ref-Q reading is always possible, either choice functions or ordinary existential quantifiers can be used to derive sets of atomic propositions. The usefulness of choice functions surfaces when islands are involved by providing an alternative to covert movement. Covert movement, if allowed, would open up the way for *absorption* and incorrectly yield multiple pair answers across the board.

3.2.2. Multiple Pair Answers Across Islands

Turning to cases in which multiple pair answers are possible across islands, I note two properties that the structure must have. One, the embedded *wh* whose value is given must be in a multiple *wh* question. Two, this multiple *wh* question must be a complement of the matrix verb. These two conditions result in a *wh triangle* construction that is met in the well-known example in (98a). Sentences like (98b), attributed by Mahajan 1990 to David Pesetsky, with an intervening clause do not allow the relevant answers, even though the intervening clause does not introduce an island, as shown by (99). (98c) is somewhat marginal since the *wh* in the embedded clause is not fronted and English does not allow *wh* in-situ in single *wh* questions. Be that as it may, it still does not allow a multiple pair answer:

- 98a. Which student knows where Mary bought which book?
 b. Which student believes (that) John knows where Mary bought which book?

c. ? Which student knows that Mary bought which book?

99. Which book does Bill believe (that) John knows Mary bought?

In Dayal (1996) I treat the multiple-pair answer as relating the matrix wh and the multiple wh complement rather than the matrix wh and the embedded wh in-situ:

100a. [[where Mary bought which book]_i [which student_j [_{t_jⁱ knows t_i]]]}

b. Q = {where Mary bought A, where Mary bought B, where Mary bought C}

c. $\lambda p \exists f_{\langle\langle s, t \rangle, e \rangle} [\text{Dom}(f) = Q \ \& \ \text{Range}(f) = \text{student} \ \& \ p = \cap \lambda p' \exists q [p' = \text{know}(f(q), q)]]$

Here the embedded multiple wh question is interpreted as a second order question, that is, a set of questions rather than a set of propositions. Since the matrix verb is defined for questions, not sets of questions, the complement QR's to take clausal scope. Semantically, it denotes the set of questions in (100b) and the full question is interpreted as a relation between the members of this set and individual students who know the answer to them.

Given the local nature of QR, this explains why the intervening clause in (98b) effectively blocks the multiple pair answer. The explanation for the second property, the requirement that the complement question be a multiple wh question, shown by the contrast between (98a) and (98c) turns on the possibility of a second order interpretation for it. Independent evidence that multiple wh questions, but not single wh questions, can denote sets of questions can be seen by examining possible answers to the following echo questions:

101a. Mary bought whát at the store?

b. {Mary bought a book at the store, Mary bought a pen at the store...}

102a. Where did Mary buy whát?

b. {Where did Mary buy the book? Where did Mary buy the pen?...}

Extrapolating from echo questions to the multiple wh questions under issue, I offer a type-theoretic account of the multiple wh complement requirement for multiple pair answers across islands. Single wh complements can be interpreted in situ and do not involve semantically driven QR.

One point that should be noted is that the proposed functional dependency in (100a) appears to be in violation of WCO. I suggest that this is perhaps due to the D-linking of the embedded wh but the issue remains open. We will return briefly to the role of D-linking in multiple pair answers in section 3.3.2.

To sum up this section, regardless of whether the particular proposals for deriving single vs. multiple pair answers hold up, the general point that the diagnostic of paired answers cannot be used as a simple indicator of wide scope has to be recognized. The distinction in the types of paired answers admitted by questions with wh in-situ inside Complex Noun Phrases and those inside Wh-islands shows that a uniform account of their interpretation is difficult to maintain.

3.3. Re-assessing Subjacency at LF

3.3.1: *The Evidence For and Against Subjacency*

In the previous sub-sections the view that wh in-situ have exceptional scope properties has been subjected to close scrutiny by drawing a distinction between two types of paired answers and relating this distinction to explicit claims about the semantics of questions and answers. Let us return now to the empirical motivation behind claims about the scope of wh in-situ and re-examine the data which have provided the basis for those claims from the present perspective. As we will see, while some data provide definitive evidence in favor of one position, some can be shown to be consistent also with the opposite position. The three language groups which will be relevant in this process are East Asian languages, South Asian languages and the Slavic/Balkan languages. Though the last group does not have wh in-situ they provide crucial evidence on the issue under discussion.

Chinese, Japanese and Korean, all wh in-situ languages, have been key players in the debate about subjacency at LF. The earliest claims for the absence of subjacency at LF was based on the interpretive properties of embedded questions in Chinese (Huang 1982). And one of the earliest challenges to this position was made by Nishigauchi (1986, 1990), Choe (1984) and Pesetsky (1987) on the basis of Japanese and Korean data. The core facts were introduced already in section 1 but I will repeat some of them to make the discussion here self-contained.

Chinese wh in-situ shows neither CNPC nor Wh Island effects:

- 103a. ni zui xihuan [piping shei de shu]
You most like criticize who REL book
“For which x, you like the book that criticizes x?”
b. ni xiang-zhidao [Lisi zeme mai-le sheme]
you wonder Lisi how bought-ASP bought
“For what object x, you wonder how Lisi bought x?”

In contrast, Nishigauchi describes Japanese as immune to CNPC alone:

- 104a. Tanaka-kun-wa[Mary-ga doko-de dono hono-o katta]-ka sitte-imasu ka
Tanaka -TOP Mary-NOM where which book-ACC bought Q know Q
“Does Tanako know where Mary bought which book?”
*”For which book x, does Tanako know where Mary bought x?”
b. Kimi-wa [dare-ga kai-ta] hono]-o yomi masi-ta ka
You-TOP who-NOM write book-ACC read Q
“For which x, you read the book that x wrote?”

The judgement for (104a), however, has been challenged recently by Deguchi and Kitagawa (2002) who argue that the putatively unavailable reading does become available under an appropriate prosodic rendering of the question. If this is correct we would have to conclude that Japanese is like Chinese, as described by Huang.

Hindi seems to fall at the opposite end of the spectrum from Chinese in not displaying any

wide scope effect whatsoever, at least as far as standard cases of relativization and complementation are concerned (see Davison (1984), Gurtu (1985), Mahajan (1990), Srivastav (1991), Dayal (1996)). With non-finite complements, that come to the left, Hindi wh in-situ patterns with Chinese and Japanese:

- 105a. jaun jaantaa hai ki meri-ne kyaa khariidaa
 John knows that Mary what bought
 “John knows what Mary bought.”
 ”What does John know that Mary bought?” *unavailable*
- b. *jaun-ne vo kitaab khariidii jo kis-ne likhii
 John that book bought that who wrote
 “For which x, did John buy the book x wrote?”
- c. john-ne kis-kii likhii kitaab khariidii
 John whose written book bought
 “For which x, did John buy the book x wrote?”

Although the effects noted here are not specific to multiple wh questions, but apply more generally to the scope of wh expressions in these languages, the following data are critical in evaluation of specific claims. In all these languages multiple pair answers are readily available in the typical Baker configuration involving a wh triangle. In fact, in Dayal 1996 I noted that this configuration is sufficient to yield multiple pair answers even in languages like Bulgarian where both embedded wh expressions have undergone local overt movement, as in (106c).²³

- 106a. Dono sensei-ga [Mary-ga doko-de dono hono-o katta]-ka sitte-imasu ka
 Which professor-NOM Mary-NOM where which book-ACC bought Q know Q
 ”For which professor x and book y, does x know where Mary bought y?”
- b. kaun jaantaa hai ki meri-ne kahaan kyaa khariidaa
 who knows that Mary where what bought
 ”Who knows where Mary bought what?”
- c. Koj znae kakvo kade e kupila Mariya
 Who knows what where bought Maria
 “Who knows where Mary bought what?”

Turning now to claims about Subjacency at LF emanating from such facts, both positions have been successfully defended. One way of categorizing the analyses that have been proposed is to use assumptions about specification of values as a guide. Among those who take specification of values as evidence of matrix scope are Watanabe 1992, Richards 1997 and Deguchi and Kitagawa 2002 in addition, of course, to Huang. Those who take pair answers to under-determine scope are Kim 1991 and Mahajan 1990, in addition to Dayal 1996 and 2002. The challenge for the first group

²³ Watanabe 1992 and Richards 1997 discuss differences in acceptability judgements depending on the position of the matrix wh. Thanks to Satoshi Tomioka for pointing me to Deguchi and Kitagawa.

is to find a principled explanation for why a single *wh* may be blocked from taking wide scope in some languages while the addition of an extra *wh* makes that option available in those languages. The challenge for the second group is to find principled reasons for enforcing local scope and finding plausible ways of deriving paired answers without wide scope.

Starting with the first group, Watanabe (1992) reconciles the Japanese facts with Huang's view that *wh* in-situ are exempt from Subjacency by arguing that Japanese, contrary to appearances, is in fact a fronting language. The distinctive feature of Japanese is that fronting involves a phonetically invisible operator, not the full phrase. This derives the absence of any phonetic impact of movement while preserving the constraints on movement typical of fronting. In order to account for paired answers in (105), then, he appeals to the fact that overt movement is restricted to one *wh* operator, as in many languages with standard fronting. The invisible operator associated with the embedded indirect object undergoes local fronting to satisfy the Q-morpheme in embedded Comp, leaving the direct object to undergo full scale LF movement in violation of Subjacency. With regard to cases like (103b) he assumes pied piping, along the lines of Nishigauchi's proposal.

Recall from the previous section the proposal in Richards 1997 that constraints only have to be met once per cycle, encoded in the Principle of Minimal Compliance. Under this view, the contrast between (104a) and (106a) can be explained in the following way. In the case of (106a), the matrix *wh* moves into matrix Spec in accordance with Subjacency and by "paying the Subjacency tax" now allows the embedded *wh* to move to matrix spec in violation of Subjacency. In contrast, (104a) is a structure in which the matrix Spec does not have a *wh* which could pay the tax, with the effect of barring the embedded *wh* from moving into that position. Richards' proposal, then, differs crucially from Huang and Watanabe in maintaining the relevance of Subjacency for covert as well as overt movement. He agrees with them, however, in his assumption that when pair answers are available they are due to scope in violation of Subjacency.

For Deguchi and Kitagawa 2002, of course, there is no *additional wh effect*. The perceived contrast between (104a)-(104b) might be due to the fact that the relevant prosody is not as easily available with yes-no questions as with *wh* questions.

Although these approaches are extremely innovative in resolving the apparent paradox presented by these data, it might be worth noting that they are subject to the criticism leveled by Mahajan (1990) and Dayal (1996) against Huang's original account that they do not explain the loss of that reading when a clause intervenes between matrix and embedded *wh* (cf. (98b)). This is, of course, due to the fact that pair answers are tied to wide scope for the embedded *wh* in-situ and the intervening clauses under discussion do not introduce any barrier to movement. Further, they are also open to the criticism in Dayal 2002 that they cannot distinguish between single pair answers that hold for questions with one matrix *wh* and one *wh* inside Complex NP's or conditionals and multiple pair answers for questions where the embedded *wh* is in a *wh* island. Finally, they do not address the possibility of multiple pair answers to cases like (106c) where the operators are moved to embedded scope positions in overt syntax.

An alternative proposal, due to Kim 1991 and Mahajan 1990, aligns *wh* expressions with quantifiers and treats *wh* movement as an instance of Quantifier Raising. Taking QR to be a local operation, this effectively enforces locality in the interpretation of *wh* in-situ and the issue of Subjacency effects becomes moot. The question, of course, arises how wide scope effects for *wh* in-situ inside islands are to be explained. The natural move to derive such effects is to treat the containing clause as having scope, allowing pair answers to cases like (106) while still blocking

them in cases like (104)-(105). The QR approach to wh in-situ is very similar to the position advocated by Dayal 1996 and 2002 (discussed in sections 3.1 and 3.2) but there are differences between the two positions that may be worth bringing out here.

One problem pointed out in Dayal 1996 is that the syntactic domain of wh movement and quantifier scope, in Hindi at least, are not identical. In particular, a quantifier inside a gerund cannot take matrix scope while a wh expression in the same position is forced to take scope at the matrix level. Other, more substantive differences, have to do with semantic assumptions about wh expressions. If wh movement is like QR it would be to a position inside C^0 . It follows, then, that the relevant semantic interpretation for wh expressions would have to involve choice functions. As we saw in section 3.1 and 3.2 it is not a straightforward matter deriving the switch from uniqueness in ordinary single wh questions to multiple pair answers in multiple wh questions. And we also saw that the distinction between single pair and multiple pair answers in the more complex cases is intractable for any theory that takes a uniform approach to all wh expressions. These authors do not provide explicit discussion of the semantics of QR'd wh expressions, and it is possible that they do not intend the co-relation between quantifiers and wh expressions to be taken literally. If so, they might consider acceptable the idea proposed in Dayal 1996 and 2002 that wh movement is distinct from QR in terms of the landing site but akin to it in being a semantically-driven local operation.

Before concluding this discussion, I should note that the problem that remains recalcitrant for approaches positing locality at LF is the possibility of wide scope for embedded wh in Chinese, exemplified in (103). However, given our refined understanding of the diagnostic of possible answers the possibility of wide scope for Chinese and Japanese wh in-situ should be evaluated against further data. Some of the relevant cases would be the Chinese counterparts of the examples in (107):

107a. Who knows how Mary bought what?

b. Who believes that John knows where Mary bought what?

If the embedded wh can indeed move to matrix Spec, we expect questions like (107b) to allow multiple pair answers. If they do not allow such answers, it would indicate that non-local movement is not possible even in Chinese. As far as I am aware, these cases have not been discussed in the literature but the predictions of the various approaches summarized here are clear enough.

We see, then, that although positing absence of Subjacency at LF provides a simple explanation for some cases, it runs into serious problems when tested across a wider set of facts. In general, it predicts a large number of unavailable readings. As highlighted in section 3.1 and 3.2 positing mechanisms for assigning wide scope without movement, such as choice functions and deriving pair answers from them, also over-generates in the same way. The most restrictive approach consistent with the data surveyed seems to be one in which covert movement is restricted essentially to local positions, and a distinction made between questions with wh expressions in Spec positions and those with wh in situ at LF. As discussed here, there are problems in this approach as well that need further exploration.

3.3.2: Specification, Scope and D-linking

As would be clear from the discussion so far, broadly speaking there are two approaches one

might take in analyzing questions with *wh* in-situ in embedded clauses. One is to take specification of values in the answer for that *wh* as evidence of wide scope. In the kinds of examples we have been looking at, this has the consequence of making Subjacency potentially irrelevant at LF. The other is to deny that specification of values in the answer is an absolute indicator of scope. This allows for the preservation of Subjacency at LF, requiring of course an auxiliary explanation for the form of the possible answers.²⁴ Here I would like to mention of some facts that bear on this issue.

Kuno and Robinson (1972) challenged Baker's account of multiple-pair answers in cases like (108a) by pointing to the fact that the version in (108b) with no *wh* in-situ also allows such answers:

- 108a. Who knows where Mary bought what?
b. Who knows where Mary bought these books?

Since questions like (108b) could not involve LF wide scope of the embedded definite, an alternative pragmatic approach must be available for such cases. But if such an account exists, why would it not extend to the cases involving *wh* in-situ? This challenge was never answered and the status of (108b) and its implications was simply set it aside during the eighties when the theory of LF was being elaborated. In fact, even in current literature the importance of these examples remains underappreciated.

Dayal 1996 (see also Dayal 2003) notes that, in fact, Kuno and Robinson's point gains further weight when the following paradigm is taken into account. (109a) does not allow a multiple pair reading while (109b) does:

- 109a. Which woman knows where Mary bought these books?
b. Which woman knows where Mary bought which book?

Changing the matrix *wh* to a singular, which we have seen in section 3.1.2 to impose uniqueness requirements in single *wh* questions, leads to loss of multiple pair answers in questions with embedded plural definites (109a). This suggests that the list answer to (108b) derives from an individual answer like *John and Bill know where Mary bought these books*. It is just a more informative rendering of this statement, as claimed by Kuno and Robinson. However, one cannot conclude the same for questions with embedded *wh* expressions. As (109b) shows, when the *wh* is D-linked, inherently or in a given context, the possibility of multiple pair answers survives even with a matrix singular *wh*. This raises two questions. What is the structural and semantic properties at issue in such cases? Why should D-linking make a difference? The approaches surveyed above have all been attempts to provide an account of pair answers that is not based on pragmatics. With regard to the second question, beyond an intuitive co-relation between familiarity with the set of entities at issue, however, there is no formal characterization of the semantics of D-linking. The data presented here shows quite clearly that if the diagnostic of possible answers is to be used seriously, more work needs to be done on the topic.

²⁴ The correlation between specification of values and scope has played a pivotal role in the analysis of the Partial Wh-Movement/Scope Marking Construction surveyed in Case # 77.

IV: Conclusion

We have seen in this survey that questions in general, and multiple wh questions in particular, have been an extremely important phenomenon in the development of linguistic theory. Typologically, they are diverse enough to raise interesting questions about the nature of universal grammar and the ways in which individual languages can differ. Further, they lie at the interface between syntax and semantics and provide a unique window into the principles governing the assignment of scope. A particular feature of questions is that they can be analyzed on the basis of two types of evidence, direct evidence in the form of acceptability judgements about fronting and indirect evidence in the form of possible direct answers admitted by them. In doing this survey, I have tried to highlight the relevant conclusions drawn on the basis of multiple wh questions as well as to refine our understanding of the nature of the diagnostics used in arriving at those conclusions.

V: References

- Ackema, P. and A. Neeleman (1998). "WHOT?" in Barbosa et al (eds), *Is The Best Good Enough? Optimality and Competition in Syntax*. MIT Press, Cambridge MA.
- Adams, M. (1984). "Multiple Interrogation in Italian." *The Linguistic Review* 4, 1-27.
- Aoun, J. (1985). *Generalized Binding: The Syntax and Logical Form of Wh-interrogatives*. Foris, Dordrecht.
- Aoun, J. and Li, A. (1993). "On Some Differences Between Chinese and Japanese Wh-elements." *Linguistic Inquiry* 24, 365-372.
- Aoun, J., Hornstein, N., Lightfoot, D. and Weinberg, A. (1987). "Two Types of Locality." *Linguistic Inquiry* 18, 557-578.
- Baker, C. L. (1970). "Notes on the Description of English Questions: The Role of an Abstract Question Morpheme." *Foundations of Language* 6, 197-219.
- Barbosa, P., Fox, D., Hagstrom, P. and McGinnis, M. (eds.) (1998). *Is The Best Good Enough? Optimality and Competition in Syntax*, MIT Press, Cambridge MA.
- Bittner, M. (1994). "Cross-Linguistic Semantics." *Linguistics & Philosophy* 17, 53-108.
- Bittner, M. (1998). "Cross-Linguistic Semantics for Questions." *Linguistic and Philosophy* 21, 1-82.
- Bosković, Z. (1997). *Superiority Effects With Multiple Wh Fronting in Serbo-Croatian*, *Lingua* .
- Bolinger, D. (1978). "Asking More than One Thing at a Time." In H. Hiz (ed.), *Questions*. D.

- Reidel, Dordrecht, 107-150.
- Calabrese, A. (1984). "Multiple Questions and Focus in Italian." In W. de Geest and T. Putseys (eds.), *Sentential Complementation*. Foris, Dordrecht, 67-74.
- Cheng, L. (1991). *On the Typology of Questions*. Doctoral dissertation, Massachusetts Institute of Technology
- Chierchia, G. (1993). "Questions With Quantifiers." *Natural Language Semantics* 1.2, 181-234.
- Choe, J.W. (1984). "LF WH-Movement: A Case of Pied Piping?" Ms. University of Massachusetts, Amherst.
- Chomsky, N. (1973). "Conditions on Transformations." Stephen Anderson and Paul Kiparsky (eds.) in *A Festschrift for Morris Halle* Holt, Reinhart and Winston, New York.
- Chomsky, N. (1981). *Lectures in Government and Binding*. Mouton de Gruyter, Dordrecht.
- Chomsky, N. (1986). *Barriers*. MIT Press, Cambridge.
- Chomsky, N. (1993). "A Minimalist Program for Linguistic Theory." in *The View From Building 20* MIT Press, Cambridge, MA.
- Chomsky, N. (1995). *The Minimalist Program*. MIT Press, Cambridge.
- Comorovski, I. (1989). *Discourse and the Syntax of Multiple Constituent Questions*. Doctoral dissertation, Cornell University.
- Comorovski, I. (1996). *Interogative Phrases and the Syntax-Semantics Interface*. Kluwer: Dordrecht.
- Cooper, R. (1983). *Quantification and Syntactic Theory*. Reidel, Dordrecht.
- Davison, A. (1984). "Syntactic Constraints on Wh in-situ: Wh Questions in Hindi-Urdu." Paper presented at the Annual Meeting of the Linguistic Society of America.
- Dayal, V. (1994). "The Semantic Consequences of Superiority Violations", Paper presented at Semantics and Linguistic Theory IV.
- Dayal, V. (1996). *Locality in WH Quantification, Questions and Relative Clauses in Hindi*, Kluwer, Dordrecht.
- Dayal, V. (2002). *Single-pair vs. Multiple-pair Answers: Wh in-situ and Scope*. *Linguistic Inquiry* 33.3, 512-520.

- Dayal, V. (2003). "Two Types of Universal Terms in Questions", in Javier Gutiérrez-Rexach *Semantics: Critical Concepts*, Routledge.
- Deguchi, M. and Kitagawa, Y. 2002. "Prosody and Wh-questions", *Proceedings of NELS 32*, GLSA, University of Massachusetts, Amherst, 73-92.
- Engdahl, E. (1986). *Constituent Questions: With Special Reference to Swedish*. Reidel, Dordrecht.
- Fanselow (1991). *Minimale Syntax*. Habilitation Thesis, University of Passau.
- Fanselow (1997). "Minimal Link Effects in German." unpublished ms., University of Pottsdam.
- Georgopolous, C. (1991). *Syntactic Variables: Resumptive Pronouns and A' Binding in Palauan*, Kluwer, Dordrecht.
- Grewendorf, G. (2001). "Multiple Wh-Fronting." *Linguistic Inquiry* 32, 87-122.
- Grimshaw, J. (1997). "Projections, Heads and Optimality." *Linguistic Inquiry* 28, 373-422.
- Groenendijk, J. & M. Stokhof (1984). "On the Semantics of Questions and the Pragmatics of Answers", in F. Landman & F. Veltman (eds.), *Varieties of Formal Semantics*. Foris, Dordrecht, pp. 143-170.
- Gurtu, M. (1985). *Anaphoric Relations in Hindi and English*. Doctoral dissertation, CIEFL, Hyderabad.
- Hamblin, C. L. (1973). "Questions in Montague English." *Foundations of Language* 10, 41-53.
- Heim, I. (1982). *The Semantics of Definite and Indefinite Noun Phrases*. Doctoral dissertation, University of Massachusetts, Amherst.
- Higginbotham, J. and R. May. (1981). "Questions, Quantifiers and Crossing." *The Linguistic Review* 1, 41-80.
- Hornstein (1995). *Logical Form: From GB to Minimalism*. Blackwell, Cambridge, Ma.
- Huang, C.-T. J. (1982). *Logical Relations in Chinese and the Theory of Grammar*. Doctoral dissertation, Massachusetts Institute of Technology.
- Jacobson, P. (1984). "On Copular Connectivity", *Proceedings of Semantics and Linguistic Theory IV*, Cornell University.
- Kamp, H. (1981). "A Theory of Truth and Discourse Representation" in J. Groenendijk et al. (eds.), *Formal Methods in the Study of Language*. Institute Mathematical Center, Amsterdam.

- Reprinted in Groenendijk et al (eds.) (1984). *Truth, Interpretation and Information*, Foris, Dordrecht, 1-41.
- Karttunen, L. (1977). "Syntax and Semantics of Questions." *Linguistics and Philosophy* 1, 3-44.
- Katz and Postal (1964). *An Integrated Theory of Linguistic Description*. MIT Press, Cambridge, Ma.
- Kayne, R. (1984). *Connectedness and Binary Branching*. Foris Publications, Dordrecht.
- Kim, S. (1991). *Chain Scope and Quantification Structure*. Doctoral dissertation, Brandeis University.
- Kiss, K. (1993). "Wh-Movement and Specificity." *Natural Language and Linguistic Theory* 11.1, 85-120.
- Kitahara, H. (1994) *Target- α : A Unified Theory of Movement and Structure-building*, Doctoral Dissertation, Harvard University.
- Kitahara, H. (1997). *Elementary Operations and Optimal Derivations*. MIT Press, Cambridge, Ma.
- Koopman, H. and D. Sportiche. (1982). "Variables and the Bijection Principle", *Linguistic Review* 2, 139-60.
- Krifka, M. (1999). "Quantifying Into Questions" *Proceedings of Semantics and Linguistic Theory* 9, 181-198.
- Kuno, S. (1982). "The Focus of the Question and the Focus of the Answer." in *Papers from the Parasession on Nondeclaratives, Chicago Linguistic Society*. Dept. of Linguistics, University of Chicago.
- Kuno, S. and J. Robinson. (1972). "Multiple Wh-Questions." *Linguistic Inquiry* 3, 463-487.
- Lasnik, H. and Saito, M. (1992). *Move α : Conditions on its Application and Output*. MIT Press, Cambridge.
- Lewis, D. (1976). "Adverbs of Quantification." In E. Keenan (ed.), *Formal Semantics*. Cambridge University Press, Cambridge.
- Mahajan, A. (1990). *The A/A-bar Distinction and Movement Theory*. Doctoral dissertation, Massachusetts Institute of Technology.
- May, R. (1985). *Logical Form*, MIT Press, Cambridge.
- McCloskey, J. (1979). *Transformational Grammar and Model Theoretical Semantics*. Reidel, Dordrecht.

- Nishigauchi, T. (1986). *Quantification in Syntax*. Doctoral dissertation, University of Massachusetts, Amherst.
- Nishigauchi, T. (1990). *Quantification in the Theory of Grammar*. Kluwer, Dordrecht.
- Oka, T. (1995). "Fewest Steps and Island Sensitivity", *MITWL 27: Papers on Minimalist Syntax*, MITWPL, Cambridge.
- Pesetsky, D. (1987). "Wh-In-Situ: Movement and Unselective Binding." In E. Reuland and A. ter Meulen (eds.), *The Representation of (In)definiteness*. MIT Press, Cambridge.
- Pesetsky, D. (2000). *Phrasal Movement and Its Kin*. MIT Press, Cambridge.
- Pope, E. (1976). *Questions and Answers in English*. Mouton, The Hague.
- Prince, A. and Smolensky, P. (1993). *Optimality Theory: Constraint Interaction in Generative Grammar*. RuCCs Technical Report #2. Rutgers University Center for Cognitive Science, Piscataway, New Jersey.
- Reinhart, T. (1997). "Quantifier Scope: How Labor is Divided Between QR and Choice Functions" *Linguistics and Philosophy* 20, 335-397.
- Richards, N. W. (1997). *What Moves Where In Which Language?* Doctoral dissertation, Massachusetts Institute of Technology.
- Rooth, M. (1992). "A Theory of Focus Interpretation." *Natural Language Semantics*. 1.1, 75-116.
- Ross, J. (1967). *Constraints on Variables in Syntax*. Doctoral dissertation, Massachusetts Institute of Technology.
- Rudin, C. (1988). "On Multiple Questions and Multiple Wh Fronting." *Natural Language and Linguistic Theory* 6, 445-501.
- Rullmann, H. (1995). *Maximality in the Semantics of Wh-Constructions*. Doctoral dissertation, University of Massachusetts.
- Saddy, D. (1991). "Wh Scope Mechanisms in Bahasa Indonesia." Lisa Cheng and Hamida Demirdash, (eds.) in *MITWPL 15: More Papers on Wh Movement*.
- Safir, K. (1986). "Relative Clauses in a Theory of Binding and Levels", *Linguistic Inquiry* 17.4, 663-89.
- Sharvit, Y. (1997). *The Syntax and Semantics of Functional Relative Clauses*. Doctoral

- dissertation, Rutgers University.
- Srivastav, V. (1991). *WH Dependencies in Hindi and the Theory of Grammar*. Doctoral dissertation, Cornell University.
- von Stechow, A. (1996). "Against LF Pied-Piping." *Natural Language Semantics*.
- Tsai, W. (1994). "On Nominal Islands and LF Extraction in Chinese" *Natural Language and Linguistic Theory* 12: 121-175
- Wachowicz, K. (1974). *On the Syntax and Semantics of Multiple Questions*. Doctoral dissertation, University of Texas, Austin.
- Wachowicz, K. (1975). "Multiple Questions." *Linguistica Silesiana* 1, 145-151.
- Wahba, W. A.-F. B. (1991). "LF Movement in Iraqi Arabic." In C.-T. J. Huang and R. May (eds.), *Logical Structure and Linguistic Structure*. Kluwer Academic Publishers, Dordrecht.
- Watanabe, A. (1992). "Subjacency and S-Structure Movement of Wh-in-situ." *Journal of East Asian Linguistics* 1:255-291.
- Williams, E. (1994). *Thematic Structure in Syntax*, MIT Press, Cambridge.