Syntactic Inquiries into Discourse Restrictions on Multiple Interrogatives[*]

Kleanthes K. Grohmann
University of Maryland, College Park)

`http://www.wam.umd.edu/~grohmann`
`grohmann@wam.umd.edu`

Language is music. We always try to create that special note.
M'hamed Benkreira

Abstract

Multiple interrogatives in German involve highly thematic Wh-elements. This is captured by the notion of "Discourse-Restricted Quantification" (DRQ): discourse restrictions in German (and possibly other languages, though not English) force both Wh-elements in double questions (Multiple Wh) to move to a syntactically thematic position which is assumed to be a functional projection in an articulated CP-structure; this projection is Top(ic)P. This movement is motivated by an underlyingly complex structure assigned to the elements which in turn has overt realizations that can be found in (interrogative) partitives. The movement is instantiated similarly to pre-subject scrambling, independently taken to be topicalization. The higher Wh-topic then moves to CP in order to check the Wh-feature; the other Wh checks its Wh-feature covertly. Evidence from a number of languages shows interesting correlations of Wh-movement and Topicalization supporting this proposal. Semantically, the scope of multiple Wh-phrases in German is subject to DRQ: the sets of referents for both Wh-elements are limited in that they must be known to speaker and hearer. The differences between German and English will be addressed and there is also plenty of cross-linguistic discussion. A wide range of phenomena are scrutinized here, including (but not exclusively) Wh-adjuncts, Wh-islands and child language. The upshot is that a conception of Wh-topics is not only more natural than one would think at first but also superior to any other account presently made public in the minimalist literature.

1 Introduction

Preliminary inquiries into syntactic conditions on multiple interrogatives in the context of discourse restrictions constitute the main body of this study. They are "preliminary" in the sense that the leading idea—if on the right track—warrants extensive treatment in more depth and detail than the scope of this paper allows; and although (or: especially because) this paper is the culmination of work spanning almost two years (Grohmann 1997b, 1997c, 1998b), some issues remain to be addressed, while others have been added. The following sections are "inquiries" in the sense that while particular issues are investigated, some in more depth than others, much of it is still speculation (or at least open to discussion). One of the aims of this paper is to suggest a syntactic analysis of the phenomenon in question; however, as will become clear presently, the sensitivity of multiple interrogatives to discourse blurs a straightforward syntactic account to a high degree and thus the second goal of this study is to investigate some of these restrictions.

This paper investigates syntactic, semantic and ultimately pragmatic (discourse) properties of Wh-elements in multiple interrogatives.[1,2] In particular, I will look at Multiple Wh-constructions and consider an approach that involves movement operations prior to—or rather instead of—overt Wh-movement.[3] While English and German data constitute an important part of the approach, the attempt is made to tie in these observations with facts from other languages; an appropriate typological validation of the present proposal, however, is another project altogether. This project will hopefully lead to further fruitful studies along the lines suggested here. It is amazing, though, to see how much mileage we can gain from elaborating on a few basic observations.

The syntactic framework adopted here is the Minimalist Program in its (more or less) well understood current manifestation.[4] As such, knowledge of the basic theoretical background and underlying assumptions will be largely presupposed. As for the semantics, the reader will notice that no particular approach is endorsed here in order to keep the details as plain yet explicit as possible, yet the general direction instigated by Karttunen (1977) with respect to interrogatives is adopted where the interpretation of questions is the set of propositions that jointly constitutes a complete true answer. While the details of the semantics involved do not play an important role here, the role of "sets of propositions" does, especially the referent sets of the two Wh-elements involved in Multiple Wh and their contextual restrictions cross-linguistically.

What starts out as the standard paradigm, this study will be interested in how to explain the different grammaticality judgements for English vs. German, as illustrated in (1) and (2):[5]

(1)a.   Who kissed whom?
     b. * Whom did who kiss?
(2)a.   Wer hat wen geküßt?
     b.   Wen hat wer geküßt?

The leading idea is that context restricts the use and interpretation of Multiple Wh. In particular, (some) languages encode this restriction with a condition that these Wh-elements be syntactically marked as highly thematic; this results in overt topicalization of the two Wh-elements instead of movement to SpecCP (the projection containing the Wh-feature). I will call this property "Discourse-Restricted Quantification" (DRQ). It can be thought of as an overt movement-variant of D(iscourse)-Linking (Pesetsky 1987). In addition, this approach can also account for apparent asymmetries among languages with respect to observance of the Superiority Condition (Chomsky 1973). It has other properties, too, however which I try to disentangle somewhat in this study.

(In lieu of a more appropriate warning I should note that the approach endorsed here involves "massive topicalization," or more neutrally expressed, "generalized pied-piping" into the left periphery. I would like to suggest to the reader that we can find support for the most crucial claim—namely concerning German Multiple Wh—in the data presented here. In

---

[1]   Some notational remarks: I will often refer to Wh-elements simply as "Wh" throughout this paper for expository purposes. Likewise, I will use the term "Multiple Wh" to denote multiple Wh-constructions.

[2]   Although I will use the term "Multiple Wh" throughout, it should be noted that I am primarily concerned with Double Questions ("Binary Wh" as termed by Pesetsky (1998), which might in fact be a more appropriate term), i.e. interrogative constructions that contain exactly two Wh-elements. It is my belief that in order to understand the phenomenon of n-nary questions, we need to have a good grip on binary Wh-constructions. I hope to contribute towards such an endeavour.

[3]   As will become clear in the text, "prior to" does not refer to "pre- vs. post-Spell Out" but rather to movement that targets a projection below the Wh-projection.

[4]   The Minimalist Program is outlined in some depth in the collection of papers in Chomsky 1995 and much current work which can be found in volumes edited by Abraham et al. (1996), Abraham and Grohmann (1997), Wilder et al. (1997), Epstein and Hornstein (forthcoming), and Martin et al. (in press) as well as studies by Collins (1997), Hornstein (1997), Kitahara (1997), Zwart (1997a), Chomsky (in press), or Uriagereka (in press), to name but a few. Technical implementations—when relevant—will be introduced and clarified in the text.

[5]   Note that throughout this paper, I will present the German examples literally to the corresponding English ones; this will be marked either by referring the reader to the specific English example in the text or with the German example immediately following the English one. Only in cases where there are no corresponding English examples will proper glosses be provided.

how far some of these apparently "crazy"—if only non-traditional—ideas pertain to other (non-)related phenomena remains to be seen. I hope to show this in the near future.)

I will proceed as follows. The next section will provide some background on the present phenomenon: the Superiority Condition will be presented along with the crucial data from English and German before considering the influence A-movement may have on Multiple Wh; this approach will then be extended and applied to German. Section 3 contains the main proposal, after a brief discussion of D-linking: DRQ is a special kind of D-linking (an overt movement-variant) and operates strictly in Multiple Wh in German; syntactically, DRQ can be characterized as topicalization of Wh-elements. In section 4, the proposal will be refined, more data will be presented in support of it and a possible implementation of DRQ will be pursued for other languages (and for child language); the basic conjecture is that D-linking has properties of syntactic, semantic and pragmatic topicalization.

2 Superiority Effects and "Operations Prior to Wh-Movement"

This section will lay out the basic empirical coverage of the phenomenon of Multiple Wh. First, the Superiority Condition will be introduced to account for some minimal pairs in English, before showing some apparent Superiority violations from German. Some analyses of Superiority will be discussed briefly, both from a GB-perspective and in a minimalist framework. After presenting a possible minimalist adoption of the phenomenon, recent case studies will be presented which will be extended as a first step towards accounting for the German facts presented throughout the paper. As will become clear, a possible connection between (multiple) Wh-movement and scrambling will be drawn which will be the starting point for further considerations.

2.1 Multiple Wh and Superiority Issues

While the Superiority Condition holds well on a descriptive level for English, some data seem to suggest its absence in German. In order to investigate a possible asymmetry any further, let us establish a theoretical approach to the Superiority Condition and a good understanding of the data.

2.1.1 The Superiority Condition and Issues Regarding Multiple Wh
Multiple Wh-constructions of the kind we are interested here involve two Wh-elements. One of the two elements is displaced (namely, under standard assumptions, fronted to sentence-initial position—COMP or SpecCP), while the other Wh remains *in situ*, in its base-generated position.

The standard case for an English matrix question—and the same goes for embedded questions—involving two bare Wh-phrases is illustrated in (3):[6,7]

(3)a.   Who bought what?
     b. * What did who buy?

The contrast in (3) derives from the particular Wh that is moved. In the grammatical case, the Wh-subject moves to SpecCP, while in the ungrammatical case it is the Wh-object that is fronted over the Wh-subject. In other words, in Multiple Wh-constructions in English the lower Wh-element may not be fronted over the higher one. This generalization is captured by the Superiority Condition (Chomsky 1973:246):

(4)*Superiority Condition*
     a.   No rule can involve X, Y in the structure … X… [$_\alpha$… Z …—WYV]…
          where the rule applies ambiguously to Z and Y, and Z is superior to Y.
     b.   The category A is 'superior' to the category B if every major category

---

[6]   Please keep in mind throughout the paper that the desired interpretation of Multiple Wh is distributive (or pair-list), as will be explained in more detail presently.
[7]   The relevance of "bare" vs. "complex" Wh-phrases will be illustrated shortly.

dominating A dominates B as well but not conversely.

Applying (4) to (3), X corresponds to the Wh-position (SpecCP) and Y to the Wh-object (*what*). The intermediate Z would be the Wh-subject (*who*) to which the rule of Wh-movement to SpecCP would also apply. Thus *who* blocks movement of *what*, rendering (3b) ungrammatical while allowing Superiority-obeying (3a).

There are some basic assumptions regarding (multiple) Wh-movement cross-linguistically. Essentially, it is assumed that Wh-movement targets a quantificational position outside IP, namely CP; this movement may be either partially overt (e.g., English and German), fully overt (such as Bulgarian or Romanian) or entirely covert (as in Japanese and Chinese). The latter two varieties are straightforward: all Wh-elements move in the overt syntax (i.e. before Spell Out) in one case and at LF in the other. In English and German, for example, (almost) any Wh-element may remain *in situ*; this may apply to single or multiple Wh-constructions.[8]

The underlying assumption, namely that at LF all Wh-elements must be in SpecCP, goes back to Higginbotham and May (1981) and Huang (1982), among others. The relevance of Wh-movement and Universal Grammar, then, is that all languages look alike at LF. One important aspect of the contribution intended by this paper is to extend the universality to Multiple Wh in a more coherent and detailed manner.[9]

Finally, concerning the interpretation of Multiple Wh, it should be noted that the object of investigation is a Multiple Wh-construction that leads to a pair-list answer (distributive reading) as opposed to Multiple Wh whose function is an echo question or the like (individual reading). Thus, question (1a) should elicit possible answers such as the following:

(5)a.   Jay kissed Kay, Bill kissed Gill, Paul kissed Paula…
    b.   Kay kissed Jay, Gill kissed Bill, Paula kissed Paul…

Likewise, possible replies to (6) would be those in (7), but crucially not in (8) where the Wh-elements are interpreted individually, which has the flavour of an echo question.[10] (Here and in the following, '#' is used to indicate infelicity.)

(6)a.   Who kissed Mary when?
    b.   Who kissed Mary where?
(7)a.   Martin kissed Mary in the garden, John in the living room, Peter on the couch…
    b.   Martin kissed Mary in the morning, John at noon, Peter at night…
(8)a. # John
    b. # On the couch/At night etc.

2.1.2 Past Approaches to Multiple Wh

The standard approach in the GB-framework, starting with Aoun et al. 1981, to account for the Superiority Condition on theoretical grounds is based on the ECP: all traces must be properly governed. (For a more detailed overview of GB-approaches to Multiple Wh and an alternative analysis, see Hornstein 1995:123ff.; for a brief summary, see Grohmann 1997c.)

This accounts well for the contrast in (1):

---

[8]   There is also the mixed case of French where the Wh may be fronted or stay *in situ*. For a possible approach to this apparent optionality, see Boskovic, in press.

[9]   Recent studies of Multiple Wh have also been concerned with this issue. Especially with respect to movement processes independent of Wh-movement, some research has been conducted not only in German (Wiltschko 1997) but also in other languages (Boskovic 1993, 1998, in press, Takahashi 1993 and numerous others). Where relevant, these proposals will be discussed here. The novelty of the current approach, however, is to make the involvement of movement operations such as scrambling to come out as a corollary, not as the driving force behind Multiple Wh-formation.

[10] It has also been observed (Comorovski 1996) that each Wh should denote more than one referent, i.e. the following is not acceptable either (but see Wiltschko 1997 for contrary view; it is not clear whether this applies to all Multiple Wh or only to D-linked contexts, but I will address this in section 3):

(i) a. # Martin [kissed Mary] in the garden
     b. # Martin [kissed Mary] in the morning

(9)a. $[_{CP} [who_i] whom_j [_{IP} t_i [_{VP} kiss t_j]]]$
    b. $[_{CP} [whom_j] who_i [_{IP} t_i [_{VP} kiss t_j]]]$

Under the assumption that Wh-elements move at LF, the trace of *whom* is lexically governed by the verb *kiss* in (25a), while the trace of *who* is antecedent governed by *who* (made possible through COMP-indexing). On the other hand, the trace of *who* in the LF-representation (25b) is not properly governed: it adjoins to *whom* which percolates its index to COMP (not *who*); thus the ECP is violated and the construction (1b) is ruled out.

    This approach cannot, however, account for "Pure Superiority", as pointed out by Hendrick and Rochemont (1982):[11]

(10) a.   Who did you persuade t to buy what?
      b. * What did you persuade who to buy t?

In the grammatical construction, both Wh-traces are properly governed at LF: *persuade* lexically governs the trace of *who* and *buy* the trace of *what*. But as can be trivially observed in (10b), the same government relations obtain. The generalization thus seems to be that only subjects (and adjuncts) fall under the ECP, namely when they need to be antecedent governed.

    Based on this and other empirical shortcomings, Pesetsky (1982, 1987) proposes to evaluate well-formedness of Wh-movement on the basis of the Nested Dependency Condition (NDC).[12] In cases of overlapping, the NDC forces one Wh-trace dependency to contain the other. This works well especially for cases in which two Wh-elements are generated in the same clause and either one could move to one of to SpecCP-positions:[13,14]

(11) a.   What did you wonder who (PRO) to see t in t?

      b. * Who did you wonder what (PRO) to see t in t?

Here, the path of the grammatical Wh-trace dependency is contained in the other, while in the ungrammatical case they overlap ("cross"). (The NDC will not be mentioned again; I refer the interested reader to Pesetsky's original discussion and to a recent return of some sort by Richards (1997).)

    Furthermore, and more relevant to the present treatment of Multiple Wh to be discussed in length in section 3, Pesetsky introduces the notion of D(iscourse)-linking. Some elements receive their interpretation not by moving to a specific position (to check a feature, in minimalist terminology), but by some way of linking to the discourse. Although the technical details of this approach may not be very clear at this point (i.e., the exact syntactic processes involved in Wh-constructions that contain D-linked Wh-elements), the idea is straightforward: D-linked Wh-elements do not have to move to SpecCP (in the overt syntax). This accounts, for example, for the grammaticality of *which*-phrases in Multiple Wh in

---

[11]   Hornstein and Weinberg (1987) set out to account for these facts by invoking a number of further stipulations, making the approach look rather "artificial" (Norbert Hornstein, personal communication).
[12]   Or, as in Pesetsky 1982, the Path Containment Condition.
[13]   For reasons that will become clear in section 3, I want to use bare Wh-phrases in (11) in order to not confuse matters with D-linking (see below). Maybe the supposedly grammatical sentence does not sound as good as a corresponding sentence with two *which*-phrases, which shall be not our concern here, though:

(i) a.   Which movie$_i$ did you wonder which actor$_j$ (PRO) to see t$_j$ in t$_i$?
     b. * Which actor$_j$ did you wonder which movie$_i$ (PRO) to see t$_j$ in t$_i$?

[14]   Note that (11) is an example of Wh-extraction out of a Wh-island. On the one hand, this allows contrasting (11) with the "heavy Wh"-version as in fn. 13; on the other hand, it serves as a precursor to the discussion on Wh-extractability in section 4.2. Needless to say, Pesetsky's approach was intended to capture all those regular cases of Superiority effects with Multiple Wh independent of islands found in English.

English, and further apparent exceptions to the Superiority Condition (see the next sub-section). We will return to this in section 3.

2.1.3 A Minimalist Conception of Multiple Wh

Let me briefly lay out how Superiority could be viewed from a minimalist perspective. The main motivation to reconsider past approaches is that the notion of "government" (which played a crucial role for the ECP) is eliminated. Thus the range of phenomena that the ECP could account for need to be re-evaluated to fit current assumptions. Recall that the main premise in the Minimalist Program for movement is that it takes place in order to check some morphological feature. If the feature is "weak," movement must be delayed until LF (according to the economy principle Procrastinate) while a "strong" feature forces the feature to move overtly and pied-pipe lexical material for PF-reasons.

As for locality and economy, the Minimal Link Condition (MLC) ensures that movement is confined to a well-defined range of possibilities. The MLC is defined as follows (from Chomsky 1995:311), interpreting "closer" as in (13), pace Chomsky (1995:299) and Kitahara (1997:15):[15]

(12) *Minimal Link Condition*
      K attracts α only if there is no β, β closer to K than α, such that K attracts β.
(13) *Closeness*
      β is closer to H(K) than α iff β c-commands α, and β is not in the minimal domain of CH, where CH is the chain headed by γ, and γ is adjoined to H(K).

The MLC rules out directly the standard cases of Superiority violations in English: K (in this case, following (12), C^0 being marked [+Wh]) cannot attract α containing a Wh-feature (*what*) because β (*who*) also contains the relevant Wh-feature, is closer and thus intervenes.[16] This is exemplified in (14) where the intervening, closer element ("β" from above) is marked in boldface:

(14) a. $[_{CP}$ who$_i$ C^0 $[_{AgrSP}$ t$_i$ bought what]]
      b. * $[_{CP}$ what$_i$ C^0 $[_{AgrSP}$ **who** bought t$_i$]]

Kitahara also adopts Huang's (1982) "null pronominal hypothesis" for *when/where*-type of Wh-expressions. Interrogative adverbs such as *when* and *where* (among others) thus differ from *why* and *how* in that they could be analysed as being part of a prepositional phrase with an empty preposition. Huang gives as evidence the grammaticality contrast between *from where* vs. *P why/how*. He also considers the existence of pronominal non-interrogative counterparts. I will return to this briefly. Again, not even an MLC-type of approach rules out immediately the behaviour of *why* and *how*.

In general terms, the MLC straightforwardly accounts for Superiority effects and suits well reformulating the Superiority Condition in current terminology. I will adopt the MLC as a general condition on movement and Superiority in particular, and extend the approach to capture the German data presented in the next sub-section.

2.1.4 Multiple Wh in German, Extended Data and More on Superiority

---

[15] Note that Chomsky (in press) dispenses with the notion of "(checking) domain" and seriously revises other minimalist concepts. As the current version of this paper is a draft, however, I will dispense with discussing many of its interesting implications.

[16] Note that while generally accepted, the MLC is not uncontroversial, neither in its conception nor its definition. I follow one possible approach here. For alternative instantiations of the MLC with respect to Multiple Wh, see Fanselow 1997, Grewendorf and Sabel 1997, and Müller 1997 among others; for a discussion that casts doubt over the MLC as a principle (as opposed to "Fewest Steps"), see Zwart 1996, 1997b.

It might also be the case, as Norbert Hornstein in recent class lectures suggests, that the conception of movement from the viewpoint of the target ("Attract") is not on the right track; instead he suggests a return to the standard operation "Move"—whatever its technical details (such as a composite of Merge and Copy, a primitive etc.). In this case, of course, the MLC makes no sense at all; rather locality conditions on movement (if they exist in the first place) must depend on a notion of "Shortest Move." I will not discuss this any further.

The same contrast as shown in (3) for English does not arise in German (cf. (1-2) above). Thus this subject-object asymmetry does not seem to hold for all languages as the Superiority Condition would suggest. Take (15), for example, which is the German equivalent of (3), repeated as (16).

(15) a.   Wer hat was gekauft?
      b.   Was hat wer gekauft?
(16) a.   Who bought what?
      b. *   What did who buy?

The situation is the same as in English, i.e. in one case the Wh-subject moves, in the other the object; as in English, the Wh-subject can be seen as the "higher" element (with respect to the Superiority Condition) as it is base-generated "above" the object. Unlike in English, however, both operations yield grammatical sentences. This seems to suggest that the Superiority Condition does not hold for German, at least not in the same way that it does for English.

But the situation is more complicated than shown so far. In particular, many of the ungrammatical English constructions are fine in German;[17] also, exceptional conditions do not render some constructions ungrammatical in English.

One of these exceptional conditions is the following.[18] If both bare Wh-phrases (*who, what* etc.) are replaced by *which*-phrases, Superiority effects do not arise even in English.

(17) a.   Which student read which book?
      b.   Which book did which student read?

The grammaticality in English naturally extends to their German equivalents:

(18) a.   Welcher Student hat welches Buch gelesen?
      b.   Welches Buch hat welcher Student gelesen?

Pesetsky (1987) accounts for the grammaticality of (17b) by D-linking (see section 3.1).

The other exception in English as noted by Baker (1970) and Bolinger (1978) involves insertion of an additional Wh-phrase, as shown in (19). (20) shows that the German equivalent is grammatical, too. (These constructions will not be discussed; see fn. 2 above.)

(19) a.   What did who give to whom?
      b.   To whom did who give what?
(20) a.   Was hat wer wem gegeben?
      b.   Wem hat wer was gegeben?

But the standard case as shown in (3)—and the asymmetry between English and German (13)—can be extended. There seems to be an asymmetry in English between certain Wh-adjuncts that is not found in German (but see fn. 19), as shown by Aoun et al. (1981), for example. Consider the English examples and compare them with the German equivalents:

(21) a. *   Whom did Martin kiss why?
      b.   Why did Martin kiss whom?
(22) a. *   Whom did Martin kiss how?

---

[17]   In this paper, I will employ data from a northern dialect of German. To my understanding, grammaticality judgements about Multiple Wh-constructions in German vary between northern and southern dialects. I will rely on judgements from my own dialect which are confirmed by speakers from the region of East Westphalia as well as places further north.

[18]   As a histo-sociological footnote, Pesetsky (1987) studies these constructions in depth and cites Chomsky (1980) as the first one who mentions these cases (accrediting the observation to Richard Kayne; also see Kayne 1983). In spirit, this observation goes back to Bolinger (1978), however, who first noted that discourse-contextual restrictions may render sentences such as (16b) in fact grammatical. Karttunen (1977:37) also mentions this fact briefly, although he simply attributes it to the "longer" Wh-phrase involved. See also Baker 1970 for early discussion.

  b. How did Martin kiss whom?
(23) a. Wen hat Martin warum geküßt?
  b. Warum hat Martin wen geküßt?
(24) a. Wen hat Martin wie geküßt?
  b. Wie hat Martin wen geküßt?

As the English a-examples show, the Wh-adjuncts *why* and *how* may not appear *in situ*. The b-examples lead us to the conclusion that if a Wh-adjunct and a Wh-argument appear in one clause, only the latter may stay *in situ*; the adjunct must move. In German, this restriction does not hold.

  If this generalization is correct, it is expected for English that the co-occurrence of *why* and *how* is ruled out, but not for German. This prediction turns out to be correct:[19]

(25) a. * Why did he kiss Mary how?
  b. * How did he kiss Mary why?
(26) a. Warum hat er Maria wie geküßt?
  b. Wie hat er Maria warum geküßt?

But this case does not extend to other Wh-adjuncts, as the following shows.

(27) a. Who kissed Mary where/when?
  b. Where/when did who kiss Mary?
(28) a. Wer hat Maria wo/wann geküßt?
  b. Wo/Wann hat wer Maria geküßt?

  Lastly, it should be pointed out that German correlates of Pure Superiority (Hendrick and Rochemont 1982) do not exist either (cf. (10), here repeated as (29)).

(29) a. Who did you persuade to buy what?
  b. * What did you persuade who to buy?
(30) a. Wen hast du überredet was zu kaufen?
  b. Was hast du wen überredet zu kaufen?

In these cases, there is no Wh-subject present; both Wh-elements are objects. From the observational viewpoint of the Superiority Condition, (29) can be accounted for straightforwardly: *who* is superior to *what*. And again, the Superiority Condition seems to be violated in German.[20]

  I have presented data from English and German Multiple Wh, i.e. those interrogatives constructions that contain (exactly) two Wh-elements. The set of data so far suggests certain asymmetries between the two languages as well as asymmetries arising among several

---

[19] As made clear in the text below, the desired reading of Multiple Wh-constructions is not individual (such as giving one Wh referents and keep the other indefinite, such as "Mary, Jane and Lynn" as a reply to (22b), for example) but distributive. Thus, a felicitous answer to (26a) or (26b) would be such as shown in (i); such a reading is not available for English (indicated by '#'), thus ruling out (25).

(i) Er hat Maria erst aus keinem besonderen Grund sanft auf die Wange geküßt, sie dann aus Zorn
  he has Mary first out no particular reason gently on the cheek kissed, her then out anger
  ins Ohr gebissen und ihr schließlich aus Zuneigung einen Zungenkuss gegeben
  in-the ear bitten and her finally out affection a tongue-kiss given
  #'He first kissed Mary gently on the cheek for no particular reason, then bit her in the ear out of anger and finally gave her a French kiss out of affection'

  Note that while not all German speakers share the judgement that *warum* 'why' and *wie* 'how' can occur low ("*in situ*"), it is the generally perceived judgement. Haider (1996), for example, proposes a solution as to why these elements can only occur in SpecCP based on their properties to quantify over propositions or predicates only. (Also see Hornstein 1995 for an elaborate discussion of the English data.) However, there might be an explanation why other German speakers find these constructions acceptable. I will present one possibility in section 4 based on the morphological make-up of German Wh-words.

[20] I present the data regarding Pure Superiority for the sake of completeness. I will not have to say much about these constructions though the analysis proposed in section 3 can also be used to account for these constructions.

constructions in English. The remainder of the paper will deal with the interpretation of Multiple Wh and the syntactic processes involved. It will culminate in a radically alternative proposal in the narrow sense; to some degree its application in a broader, cross-linguistic perspective will be evaluated.

2.2   Overt A-Movement Prior to Wh-Checking

The intuition behind the following is very much in the spirit of some recent proposals: NP-movement (A-movement) and Wh-movement (A'-movement) may interact. In particular, I will suggest that scrambling may obviate Superiority effects. I will first show what A-movement can do to Multiple Wh in English, before looking at scrambling and Multiple Wh in Japanese and Spanish—and possibly Hebrew—, before looking at a possible implementation for German.

2.2.1 Hornstein (1995): Raising and Multiple Wh in English

The standard instance of A-movement—alongside passivization—is raising. Moreover, unlike scrambling, (obligatory) raising can be observed in English, such as in (31):

(31) a. * Seems (to his friends) Paul to enjoy commercials
      b.   Paul seems (to his friends) to enjoy commercials

The numerations of (31a) and (31b) are identical; this rules out the other option that English allows, namely to insert *it* in subject-position in (31a) and make the embedded clause to finite (as in *It seems (to his friends) that Paul enjoys commercials*). Thus, when the numeration is exhausted (at the stage depicted in (31a), for example), the embedded subject *Paul* must raise out of the non-finite clause into the subject-position of the finite matrix clause (31b).

Interestingly, a Wh-subject in the non-finite embedded clause underlies the same constraint:

(32) a. * Seems (to his friends) who to enjoy commercials?
      b.   Who seems (to his friends) to enjoy commercials?

The embedded Wh-subject cannot stay *in situ*, nor can it simply move to an embedded SpecCP.[21] It must move to matrix SpecCP.

It may shed some light on Multiple Wh to see how raising contexts change the grammaticality status; let us insert expletive *it* in one case and raise the Wh in the other:

(33) a. * Who does it seem to whom enjoys commercials?
      b.   Who seems to whom to enjoy commercials?

In a Multiple Wh-constructions where the matrix subject-position is filled by *it*, the Wh-subject may not move from its embedded position to matrix SpecCP; in a construction where the matrix subject-position needs to be filled through Raising, it may. As a matter of fact, it must raise regardless of whether the raised element stays in this position. But in (33b), the MLC is violated: the closest element that C^0 can attract is not *who* but *to whom*. (The same applies to (33a).)

Now consider (34):

(34) a. ? To whom does it seem that who enjoys commercials?
      b. * To whom seems who to enjoy commercials?

In accordance with the MLC, *to whom* moves to SpecCP in (34a), forming a (more or less) grammatical sentence. But if the same operation takes place in the former grammatical (33b), the construction is ruled out (34b)—despite accordance with the MLC (presumably because *who* fails to be Case-marked).

---

[21]   For present purposes, it is immaterial whether this position exists as such, or what the status of vacuous movement may be.

The observation to take with us from these examples is that NP-movement prior to Wh-movement may have an effect on the application of the MLC and hence on the relevance of the Superiority Condition. Rather than treading further on Hornstein's steps, I will pursue an approach that does not rely solely on semantic properties of Wh-elements in general but on discourse-conditioned properties in Multiple Wh.

2.2.2 Takahashi (1993): Scrambling and Multiple Wh in Japanese

The intuition that scrambling and Wh-formation may interact was also pursued by Takahashi (1993). The generalization that emerged from his study is roughly that Japanese scrambling comes in two types, both of which differ with respect to Wh-elements. Short scrambling in Japanese can be viewed as involving A-movement, displaying many of its characteristics, unlike long scrambling: scrambling across clause boundaries should rather be analysed as A'-movement much in the spirit of Mahajan's (1990) cross-linguistic characterization of scrambling phenomena (also see Corver and van Riemsdijk 1994 among many other works on scrambling).[22]

The following shows both constructions (Takahashi 1993:663-664):

(35) a.   Dare-ga nani-o   tabeta no?
            who-NOM what-ACC ate   Q
            'Who ate what?'
      b.   Nani-o dare-ga tabeta no?
(36) a.   John-ga   dare-ni [Mary-ga   nani-o   tabeta to] itta no?
            John-NOM who-DAT Mary-NOM what-ACC ate   COMP said Q
            'Who did John tell that Mary ate what?'
      b. ?? $Nani_i$-o John-ga dare-ni [Mary-ga t_i tabeta to] itta no?
      c. $Pizza_i$-o John-ga dare-ni [Mary-ga t_i tabeta to] itta no?

While one construction exemplifies the possibility of short scrambling of two Wh-elements (35b)—with the base-order as in (35a)—the other one shows that this is not possible crossing clause boundaries (36b). Of particular interest is the contrast between (36b) where the Wh-element long scrambles, yielding ungrammaticality/strong marginality, and (36c) which is an instantiation of long scrambling of a non-Wh-element.

What the Japanese examples show is that Multiple Wh-constructions differ depending on whether short or long movement takes place. As this is also a crucial distinction for purposes of scrambling, an interaction does not immediately seem implausible. It is also clear that movement of the Wh-elements cannot be driven by standard feature-checking (i.e. that a strong Wh-feature needs to be checked) as Japanese Wh-formation is *in situ* and the Wh-feature thus weak.

2.2.3 Boskovic (1993): AgrOP and Multiple Wh in Spanish (and Hebrew)

In a study on Multiple Wh and interactions with other properties of the grammar, Boskovic (1993) considers Hebrew and Spanish, among other languages, and finds interesting similarities and contrasts.[23] The upshot of his discussion is that the presence of an AgrOP may have further consequences.

---

[22] The nature of scrambling is not topic of this paper (but see section 2.3). The data and judgements are taken from the literature, and to my understanding variations in acceptability exist. This might give us a clue that the input of A-relations in Multiple Wh is not as straightforward as the data seem to suggest initially.

While scrambling most likely affects Multiple Wh, it cannot be the only influence. It seems to be established that the Wh-elements do not move overtly for the purpose of Wh-checking; also, scrambling allows free order of all arguments. With this in mind, (i) (Akemi Matsuya, personal communication) is unexpected:

(i) a.   Dare-ga   Mary-ni nani-o   ageta no?
            Who-NOM Mary-DAT what-ACC gave Q
            'Who gave Mary what?'
     b.   Dare-ga nani-o Mary-ni ageta no?
     c. * Nani-o Mary-ni darega ageta no?

[23] Note that in the published version (Boskovic 1997), the data from Hebrew have been omitted, presumably because their status is not very clear. While the details may be more complicated than exposed here, this type approach is at least feasible to some degree and hence included in our present discussion. Nothing hinges on

Multiple Wh-movement looks like that in Type I languages (such as English or German) but apparently violates the Superiority Condition in both languages. Consider (37) from Hebrew and (38) from Spanish where there also seems to be no subject-object asymmetry (Boskovic 1993:2):

(37) a.   Mi   kana   ma?
             who bought what
             'Who bought what?'
      b.   Ma kana mi?

(38) a.   Quién dijo qué?
             who   said what
             'Who said what?'
      b.   Qué dijo quién?

In his discussion, Boskovic notes that usual WCO-configurations exist and lead to ungrammaticality. Also, the Superiority Condition is supposedly active in both languages as he shows with an example from Spanish. Thus it is not the case that Spanish or Hebrew differ in crucial respects from other languages that might otherwise account for the absence of Superiority above.

There are good reasons to assume that subjects in both languages may stay *in situ*, i.e. in the base-generated position SpecVP. In order to account for the grammaticality of (37-39), Boskovic proposes that a Wh-element passes through its corresponding AgrP on the way to the Wh-related position, SpecCP. Thus, an otherwise post-verbally appearing Wh-subject moves to CP via AgrSP in order to check the relevant features, while a Wh-object moves through AgrOP.[24]

Boskovic (1993:6) offers the following derivations for (37b) and (38b), respectively (but see fn. 23 on the status of Hebrew and Boskovic 1997 for more details):

(39) a. $[_{CP}$ ma$_i$ kana$_j$ $[_{AgrOP}$ t$_i$ $[_{VP}$ $[_{V'}$ t$_j$ t$_i]$ mi$]]]$
      b. $[_{CP}$ qué$_i$ dijo$_j$ $[_{AgrOP}$ t$_i$ $[_{VP}$ $[_{V'}$ t$_j$ t$_i]$ quién$]]]$

It is the lower Wh-element that moves. As it has to check its D-features (such as Case) on its way (see fn. 24), the Wh-object moves to the checking position in AgrOP. This is in fact the closest position to CP; thus, if this element then moves on to SpecCP for checking its Wh-feature, it satisfies the MLC trivially. (Similarly, if the Wh-subject were to move, it would pass through AgrSP, from where it would also be closest to SpecCP.[25])

There is thus some evidence that non-Wh-related movement operations might influence the outcome of well-formed Multiple Wh.

2.3   A Possible Application to German

I will now consider a possible implementation of the running idea—namely that other properties of the grammar may interact with (multiple) Wh-movement—to German. First I will lay out general assumptions regarding German clause structure, then consider such an implementation.

2.3.1 The Agr-Complex in German and Scrambling Below the Subject
Following recent minimalist work on West Germanic clause structure, I assume that German word order is underlyingly SVO. (This cannot be discussed in any depth here; see Zwart

---

this particular analysis; we will see in a later section, what role Hebrew Multiple Wh play in an attempt to classify the present observations across more languages than English and German.

[24]   From a minimalist perspective, this makes sense cross-linguistically (or: on theory-internal grounds based on Chomsky 1995, it is a very reasonable assumption following from the general framework): weak features need to be checked at LF; but if an element has already passed the relevant projection (as in the case of topicalization, Wh-movement etc.), it cannot do so, assuming that lowering for feature-checking (or any other purposes) is ruled out also at LF. Even if the strong/weak-dichotomy of features is abandoned (à la Chomsky, in press and other works), this type of operation is feasible, if not even desired.

[25]   I leave out the possibility of interpreting these constructions as an indication that the EPP might not be required, as suggested to me by Norbert Hornstein.

1993, 1997a, Hinterhölzl 1997, Lattewitz 1997). Let me simply give the following examples to show that arguments in German always leave VP, following the standard assumption that sentential negation marks a position very close to the VP-boundary:[26]

(40) a. Der Martin$_i$ hat der Maria$_j$ das Buch$_k$ nicht [$_{VP}$ t$_i$ gegeben t$_j$ t$_k$]
          the Martin has the Mary the book   not   given
          'Martin didn't give the book to Mary'
      b. Der Martin hat der Maria ein (bestimmtes) Buch nicht [$_{VP}$ gegeben]
(41) a. * Der Martin hat der Maria nicht [$_{VP}$ das Buch gegeben]
      b. * Der Martin hat der Maria nicht [$_{VP}$ ein (bestimmtes) Buch gegeben]
(42) a. Der Martin hat der Maria nicht [$_{VP}$ (irgend)ein Buch/was gegeben]
      b. * Der Martin hat der Maria (irgend)ein Buch/was nicht [$_{VP}$ gegeben]

While (40a) shows that all definite arguments leave VP (cf. (41a)), (40b) shows the same for specific indefinites (cf. (41b)); the specificity is made explicit by *ein bestimmtes* 'a specific'. (42) highlights the contrast between an indefinite non-specific argument preceding sentential negation vs. following it.[27] (On the non-specificity of non-interrogative *was* 'what/ something', see Haiden 1995 and the discussion in section 3.3.2 below.)

Scrambling can then be seen as movement out of VP into the Agr-complex. In spite of many attempts to account for scrambling more accurately, no perfect solution is known to me. I thus adapt one of the available accounts and assume the ordering of the two AgrOPs (AgrDOP and AgrIOP for direct objects and indirect objects, respectively) to be free.[28,29] This yields the following structural configurations:

(43) a. [$_{AgrSP}$ Der Martin$_i$ hat [$_{AgrIOP}$ der Maria$_j$ [$_{AgrDOP}$ das Buch$_k$ nicht [$_{VP}$ t$_i$ gegeben t$_j$ t$_k$]]]]
      b. [$_{AgrSP}$ Der Martin$_i$ hat [$_{AgrDOP}$ das Buch$_k$ [$_{AgrIOP}$ der Maria$_j$ nicht [$_{VP}$ t$_i$ gegeben t$_j$ t$_k$]]]]

In minimalist spirit, overt movement takes place for checking strong features. Here I follow current assumptions that movement out of VP takes place to check strong (D-, φ- and/or) Case-features in AgrP. The exact properties of Agr are not relevant for current purposes.

2.3.2 The CP-Domain: Pre-Subject Scrambling as Topicalization

---

[26] I will not discuss further whether this only applies to definite arguments (Diesing 1992) or specific ones (Moltmann 1991), though the latter seems to be on the right track (viz. (40b)); furthermore, as we will see below, specificity is also involved in Multiple Wh. (Haiden (1995) presents good evidence that all arguments leave VP.)

[27] As a matter of fact, indefinites under negation morphologically amalgamate *nicht* and *ein*:

(i) Der Martin hat der Maria kein Buch/nichts   gegeben
     the Martin has the Mary no   book/nothing given
     'Martin didn't give Mary a book/anything'

I am not concerned here whether this is evidence in favour of the indefinite object having moved out of VP over negation or whether negation sits higher than the object; even in the latter case it is not clear that the indefinite has not moved out of VP (see Haiden 1995).

[28] Free order of the AgrO-projections is an adaptation of Haeberli's (1995) proposal that all Agr-projections may be freely ordered. Independently, however, I assume the subject to be in a fixed position in relation to objects, which I mark here as AgrSP. This projection must be set higher than either AgrOP. (See Grohmann, in progress for a detailed and unified approach for this phenomenon.)
The free ordering of AgrIOP and AgrDOP can be derived by assuming Chomsky's (1995) notions of Minimal Domain and Equidistance; an application of these relations results in the lower AgrOP to be equidistant for either object to move to (more in section 3.3.1 below). Let us then assume further that the specification of AgrOP as AgrDOP or AgrIOP depends on the element that moves and checks Case.

[29] Note that one possible alternative may be free ordering of the base-generated arguments in VP. It is not directly relevant to the present issues how middlefield-scrambling can be accounted for and I will hence simply adopt the idea that only AgrIOP and AgrDOP may appear in any linear order for reasons of exposition. A rich body of literature exists about the nature of scrambling and the role of base-generation vs. A- or A'-Movement (see, among many others, the collection of papers in Corver and van Riemsdijk 1994). To briefly mention yet another approach, Richards (1997) adopts the framework of multiple specifiers (Ura 1994, Chomsky 1995) about which I have nothing to say at this point. Some arguments against a multiple specifier analysis of German scrambling can be found in Grewendorf and Sabel 1997.

Scrambling as construed above (i.e. as feature-driven A-movement) is not a uniform process in German. As claimed in previous work, I assume there to be a substantial difference between scrambling below the subject and scrambling above the subject, henceforth pre-subject scrambling (Grohmann 1996, 1997d, in progress). This relies crucially on the assumption that the subject in German sits in a unique position in unmarked clauses (SpecAgrSP in the present configurations), an assumption that seems to be borne out (Grohmann 1998c).

The subject is arguably in its assigned position, SpecAgrSP.[30] It thus functions as an indicator for the IP-boundary: all material preceding the subject is outside IP, in the CP-domain. I will refer to this as the left periphery. Following Rizzi (1997), CP could be conceived as an articulated domain, much in the spirit of the finer structure of IP, developed in detail ever since Pollock's (1989) seminal paper. The most relevant functional projection in this domain is a recursive Top(ic)P, independently established for German in recent work (see Müller and Sternefeld 1993, Haftka 1995).

As head movement is not of much concern here (especially verb movement and the V2-phenomenon, see fn. 33 below), the following are the major projections including their intrinsic features, immediately relevant for present purposes:

(44)   CP [Wh] — TopP* [Top] — AgrSP [Nom] — AgrI/DOP [Dat/Acc] — VP [θ]

All arguments leave VP (where they are θ-marked) and move to the specifiers of the respective Agr-projections (for Case-checking);[31] arguments scrambled over the subject are in addition endowed with a Top-feature, moving to the specifier(s) of TopP(s). The Wh-feature sits on the head of CP.[32] I will leave open the process of positioning adjuncts, but nothing a priori prohibits their presence in TopP.

Implementing these ideas, we would receive something like the following representations (leaving out negation or adverbs) for instances of pre-subject scrambling:[33]

(45)   a.   [$_{TopP}$ Das Buch$_i$ hat [$_{AgrSP}$ der Martin [$_{AgrDOP}$ t$_i$ [$_{AgrIOP}$ der Maria gegeben]]]]
       b.   [$_{TopP}$ Der Maria$_j$ hat [$_{TopP}$ das Buch$_i$ [$_{AgrSP}$ der Martin [$_{AgrDOP}$ t$_i$ [$_{AgrIOP}$ t$_j$ gegeben]]]]]
       c.   [$_{TopP}$ Das Buch$_i$ hat [$_{TopP}$ der Maria$_j$ [$_{AgrSP}$ der Martin [$_{AgrIOP}$ t$_j$ [$_{AgrDOP}$ t$_i$ gegeben]]]]]

With freely ordered AgrOPs, both objects satisfy the MLC in movement out of VP.[34] In that configuration, they also satisfy the MLC for further movement to TopP.

The same, of course, applies to adjuncts as well. A temporal adverbial such as *gestern* 'yesterday' would also be topicalized when it appears in front of the subject.[35] It is also shown that the same process takes place in embedded contexts (see Grohmann 1996, 1997d).

---

[30]   Note that in Chomsky 1995:sec. 4.10 and much current work, the view of Agreement projecting is dispensed with altogether in favour of an implementation of multiple specifiers (Kuroda 1988). The main objection to AgrP is that it carries no interpretable feature and is hence not "virtually conceptually necessary" at the interfaces. However, assuming the existence of an AgrP does not necessarily involve its survival at LF, and alternatives to deal with interpretability issues are feasible (see Zwart 1997b).

[31]   It is Case, as indicated here; I leave out the possibility that this move simply checks D-features (which may even be conceptualized as the EPP) on DPs rather than Case. There is recent literature on either possibility and even more, all irrelevant for the present purposes.

[32]   Note that Rizzi also suggests a recursive TopP to the left of CP (FocP), an assumption empirically supported by data from Romance languages, among others, and also by German (as shown in earlier work, see Grohmann 1997a). I will return to the relevance of this position later.

[33]   In this paper, I am not concerned with the specifications of deriving V2. Although it undoubtedly plays a crucial role even in these constructions, I will remain agnostic about the nature of movement or whether movement plays a role at all. In the light of this, I will not indicate possible traces of the verb in derivations.

[34]   Norbert Hornstein (personal communication) reminds me that the assumption of freely ordered AgrOP is not immediately compatible with recent minimalist proposals concerning movement, such as Richards (1997). Note, however, that Richards' approach crucially builds on the availability of multiple specifiers in the grammar—an assumption that is not shared here (at least not with respect to German; see also Grewendorf and Sabel 1997 on this).

[35]   A severe difficulty arises when we want to account for adjuncts, simply for the reason that no satisfactory theory of their syntactic (and semantic) behaviour exists. One of the main problems is their position in the clause structure: (where) are adjuncts base-generated, (where) do they move, what categorial status do they

(46) a. $[_{TopP}$ Gestern$_j$ hat $[_{AgrSP}$ der Martin t$_j$ das Buch der Maria gegeben]]
             yesterday has   the Martin   the book the Mary given
             'Yesterday, Martin gave the book to Mary'

      b.   Er sagte, $[_{ForceP}$ daß $[_{TopP}$ die Maria$_j$ $[_{AgrSP}$ der Martin t$_j$ geliebt hat]]]
             he said   that   the Mary   the Martin loved has
             'He said that Mary, Martin loved'

In (46a), the sentential adverb in pre-subject position sits in TopP (where I leave out the exact derivation, cf. fn. 35); in (46b) the fronted object moves there.[36]

2.3.3 Grohmann (1997b): Scrambling and Multiple Wh in German

What the previous presentations have shown is that A-movement may obviate Superiority effects in (Multiple) Wh-constructions. One clear case is Raising in English, the others are scrambling (Japanese) or nominal feature-checking (Spanish); in order to consider a similar application—namely, movement operations prior to Wh-movement—to German, the status of scrambling needs to be addressed as well as the requirement of it being A-movement. For reasons of space, I cannot indulge in a deeper treatment here. I would like to hold fast to the resulting observation that there are movement processes that occur prior to Wh-movement, and whether it is solely constrained to A-movement is not relevant. German scrambling is such a process, and as we have seen above, pre-subject scrambling is topicalization. The following will be an implementation of scrambling and Multiple Wh in German.

In Grohmann 1997b, a proposal is made that takes into account such operations. In order to observe the MLC, C^0 can attract only the closest Wh-element. In simple constructions such as those we have seen so far, an Wh-object may be fronted over a Wh-subject if it scrambles prior to Wh-movement. Pre-subject scrambling conceived of as topicalization may thus obviate Superiority effects. The following exposition will be elaborated thoroughly in section 3; in particular, while the earlier approach aspired to find a link between scrambling and Multiple Wh, empirical evidence will now be presented to correlate topicalization and Multiple Wh, and by doing so, deriving the apparent influence of scrambling as a corollary.

The following data from embedded Multiple Wh illustrate (Grohmann 1997b:101):[37]

(47) a.   Ich frage mich $[_{CP}$ wer$_i$ C^0 $[_{AgrSP}$ t$_i$ $[_{AgrOP}$ was$_j$ gekauft hat]]]
             I   ask   myself who   what bought has
             'I wonder who bought what'

      b.   Ich frage mich $[_{CP}$ was$_j$ C^0 $[_{TopP}$ t$_j$ $[_{AgrSP}$ wer$_i$ $[_{AgrOP}$ t$_j$ gekauft hat]]]

The derivation of (47a) is as straightforward as in the English case: the Wh-subject, stopping on its way to CP in AgrSP, satisfies the MLC as it is closest to C^0. In (47b) we observe that the Wh-object is the element overtly Wh-fronted. We can now capitalize on the analysis of scrambling assumed here: scrambling over the subject, *was* targets SpecTopP to check off a Top-feature. In this position, above the subject, only *was* can satisfy the MLC in the following Wh-movement.

One prediction is that any Wh-argument may occur in SpecCP, with all other (Wh-) arguments lower. If all Agr-heads in German are endowed with a strong feature which drives the arguments to move out of VP overtly, these movement operations must also take place

---

have, are they right- or left-adjoined or both? Some proposals towards the syntax of adverbs or adjuncts in general by Laenzlinger (1996), Alexiadou (1997), Ernst (in press), Cinque (in press) may give us a clue.

[36]   The position of the complementizer is here indicated as ForceP, following Rizzi's finer structure of CP. This will not be relevant here, though the distinction of a position for Wh-elements vs. complementizers might be.

   The issue of these projections in interaction will be explored in more detail very soon, as well as a unified account of generalized pied-piping into the left periphery rather than the apparently unmotivated "massive topicalization" that I assume here (Grohmann, in progress).

[37]   The approach to be developed in the following sections deviates from Grohmann 1997b in that all Wh-elements obligatorily undergo topicalization; in the earlier version, this was just a means to circumvent Superiority effects in the guise of scrambling. Topicalization here is motivated on different grounds.

when endowed with an additional Wh-feature. The consequence is that prior to Wh-movement, the particular element has moved to the specifier of either TopP or AgrP.[38]

(48)   Ich frage mich…
   a.   … [$_{CP}$ wem$_i$ [$_{TopP}$ t$_i$ [$_{AgrSP}$ wer [$_{AgrOP}$ t$_i$ [$_{AgrOP}$ das Buch gegeben hat]]]]]
   b.   … [$_{CP}$ wem$_i$ [$_{TopP}$ t$_i$ [$_{AgrSP}$ der Hans [$_{AgrOP}$ t$_i$ [$_{AgrOP}$ was gegeben hat]]]]]
   c.   … [$_{CP}$ wem$_i$ [$_{TopP}$ t$_i$ [$_{TopP}$ was$_j$ [$_{AgrSP}$ der Hans [$_{AgrOP}$ t$_j$ [$_{AgrOP}$ t$_i$ gegeben hat]]]]]]
   d.   … [$_{CP}$ wem$_i$ [$_{TopP}$ t$_i$ [$_{TopP}$ das Buch$_j$ [$_{AgrSP}$ wer [$_{AgrOP}$ t$_j$ [$_{AgrOP}$ t$_i$ gegeben hat]]]]]]
(49)   Ich frage mich…
   a.   … [$_{CP}$ was$_i$ [$_{TopP}$ t$_i$ [$_{AgrSP}$ wer [$_{AgrOP}$ t$_i$ [$_{AgrOP}$ der Maria gegeben hat]]]]]
   b.   … [$_{CP}$ was$_i$ [$_{TopP}$ t$_i$ [$_{AgrSP}$ der Hans [$_{AgrOP}$ t$_i$ [$_{AgrOP}$ wem gegeben hat]]]]]
   c.   … [$_{CP}$ was$_i$ [$_{TopP}$ t$_i$ [$_{TopP}$ wem$_j$ [$_{AgrSP}$ der Hans [$_{AgrOP}$ t$_j$ [$_{AgrOP}$ t$_i$ gegeben hat]]]]]]
   d.   … [$_{CP}$ was$_i$ [$_{TopP}$ t$_i$ [$_{TopP}$ der Maria$_j$ [$_{AgrSP}$ wer [$_{AgrOP}$ t$_j$ [$_{AgrOP}$ t$_i$ gegeben hat]]]]]]

These simplified representations (adapted from Grohmann 1997b:102) show the general idea, namely a possible approach to Multiple Wh in terms of topicalization of Wh-elements.

In sum, pre-subject scrambling may obviate Superiority effects in German. One big question that arises will be addressed in the next section; in fact, it lies at the heart of the present proposal which also eliminates Wh-movement from these Multiple Wh-constructions in German: what does it mean for a Wh to move through (or even to) TopP?

3   DRQ as "Operations Instead of Wh-Movement"

Building on the intuition that D-linked Wh-phrases do not undergo operator-movement to SpecCP, I will demonstrate how D-linking differs from Discourse-Restricted Quantification and propose topicalization as the movement operation in lieu of "standard" Wh-movement. It will be shown in how much this analysis carries over to English, as well as other languages. The upshot is that languages differ with respect to the realization of D-linking.[39]

3.1   D-Linking Is Not Enough

Pesetsky (1987), building on Bolinger's (1978) original observation, lays out the basics of how context may be used to account for apparent Superiority violations in English. He proposes that certain Wh-phrases are D-linked and thus underlie different syntactic constraints than non-D-linked ones. The most prominent of inherently D-linked Wh-phrases is the *which N*-type; Pesetsky thus suggests that D-linking can account for the absence of subject-object asymmetries; the relevance of partitive Wh-phrases will become clear shortly.

(50) a.   Which student read which book?
      b.   Which book did which student read?
(51) a.   Mary asked [[which] of the students] read [[which] of the books]
      b.   Mary asked [[which] of the books] [[which] of the students] read

The following will contain a discussion of D-linking in English and an extended approach to German Multiple Wh. It will become clear that Multiple Wh underlie stricter contextual restrictions, here understood as Discourse-Restricted Quantification (henceforth, DRQ). I will then propose a syntactic approach to these semantic and pragmatic constraints.

3.1.1 Context in English and D-Linking
As a starting point, D-linking can be defined as a restriction on (multiple) questions where a felicitous answer can only denote sets of referents for the Wh-element(s) that have previously

---

[38]   While the basic gist of this approach will be adopted and elaborated here, I will assume movement to TopP of both Wh-elements. This will be presented in section 3.2.3 and the subsequent discussion will address some of the problems but also some of the consequences. Crucially, movement to TopP differs dramatically from the earlier approach.
[39]   The proposal put forward, based on the following discussion, differs in some details from Grohmann 1998b.

been established in the discourse.[40] Such a view of D-linking immediately conjures up a correlation with partitives where the set of referents for the NP is already established. Thus, Wh-phrases of the types *which N(s)* and *which of the Ns* are inherently D-linked.[41] This accounts for the grammaticality of (50-51) on the one hand and the grammaticality difference in (1a) vs. (1b) on the other; *who, what* etc. are not—and in general, cannot be[42]—D-linked.

The two major discussions in the literature regarding D-linking are the works by Bolinger (1978) and Pesetsky (1987). Curiously, both crucially differ in what is the trigger for apparent Superiority violations with D-linked phrases (Comorovski 1996:85): Bolinger takes the D-linked property of the fronted Wh to be the crucial factor, while Pesetsky claims it is the D-linking of the non-fronted Wh (Wh-*in-situ*).

The following two situations from Bolinger 1978:108 and Pesetsky 1987:109, respectively, show how context may D-link a bare Wh (emphasis from Comorovski 1996:84):

(52) a. *   What did who break?
      b.   I know that among all the disasters in that kitchen, Jane scorched the beans and Lydia put salt in the ice tea; but *whát* did *whó* bréak? I know that somebody broke something, so stop evading my question.
(53) a.   I wonder where what goes
      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 *whére whát* goes!

In both contexts, the fronting of the lower Wh is possible; but do these examples show that it is the D-linking of the fronted Wh (52b) or of the lower Wh (53b) that allows this? Presumably not; remarkably (as also pointed out by Comorovski), neither author employs examples with one D-linked and one bare Wh to support their claims. Besides, (53a) is good even without context (see section 2.1.3 above).[43]

Now consider the following sets of questions.

(54) a. *   Who did who kiss this week?
      b. *   What did you persuade who to buy?
(55) a. *   Which girl did who kiss this week?
      b. *   Which book did you persuade who to buy?

While (54) repeats the by now well-known Superiority facts, (55) shows that if a D-linked Wh is fronted over a bare Wh, grammaticality does not automatically arise.

On the other hand, if a non-D-linked Wh is fronted over a D-linked one (56), the question is acceptable.[44]

---

[40]   Pesetsky (1987:107f.) formulates it as "[w]hen a speaker asks a question like *Which book did you read?*, the range of felicitous answers is limited by a set of books both speaker and hearer have in mind," which Comorovski (1996:2) assumes when calling "*discourse-linked* those Wh-phrases whose range of felicitous answers is limited by a set of objects already referred to in the discourse or salient in the context of utterance."
[41]   It is a widely held view that D-linking and partitivity are alike in their semantic (and ultimately syntactic) properties. See e.g., Kiss 1993, Comorovski 1996, Wiltschko 1997 and references cited there for discussion.
[42]   D-linking of bare Wh-phrases is possible under heavy stress, for instance, or in even more restricted context. See Baker 1970 and Bolinger 1978, among others, and the discussion by Pesetsky (1987); see also (52-53).
[43]   The other example Pesetsky (1987:109) gives is (ib) which does indeed contrast with (ia):

(i) a. *   What did who do?
     b.   I know what just about everybody was asked to do, but *whát* did *whó* (actually) do?

While (ia) is not part of any specific context, a rough reference of both Wh-elements in (ib) has been introduced in the discourse. The crucial role of this will be explored in much more detail in the following.
[44]   It is well-known to anyone who studied Multiple Wh in some depth, that judgements are very intricate. My own (limited set of) informants agree with the judgements given here for (55-56). Comorovski (1996:85), however, presents the following judgements to show that neither Bolinger nor Pesetsky were correct with their particular accounts:

(i) a. ?   Which book did how many people buy?

(56)   a.   Who did which boy kiss this week?
       b.   What did you persuade which of your friends to buy?

     The data above show—if anything—that context plays a crucial role; in how much we can make any appropriate generalizations, however, differs from the individual viewpoint.
    Comorovski argues that the leftmost Wh must be D-linked. (But see Dobrovie-Sorin 1997:16 for criticism on Comorovski's strong claim.) This carries over to bare Wh:

(57)   a.   It's nice to have all those times scheduled, but when are you doing what?
           (#But what are you doing when?)
       b.   It's nice to have all those activities ahead of you, but what are you doing when?
           (#But when are you doing what?)

Comorovski (1996:143) illustrates with these examples taken from Bolinger 1978 that wide-scoped Wh-phrases are always D-linked. Here we encounter a dilemma, however: wide scope is established by movement to CP (where it takes scope over the entire clause); on the other hand, Pesetsky argues that D-linking does not induce movement. I will return to this after an exposition of the German facts and an alternative approach to Multiple Wh.
     Compare lastly (57), where bare Wh-phrases may also be D-linked, with the situation depicted in (58). Bare D-linked Wh-phrases differ from inherently D-linked Wh-phrases, in that only the latter may appear in any order:[45]

(58)   It's nice to have all those activities scheduled for all those times,
       a.   … but at which time are you going to do which activity?
       b.   … but which activity are you going to do at which time?

3.1.2 Context in German and DRQ

In German, Superiority effects cannot as easily be found as in English. I propose on the basis of the following data and contexts that D-linking occurs in a much stricter version, which I call Discourse-Restricted Quantification (DRQ): the quantificational force of both Wh-elements in a German Multiple Wh-construction is severely restricted by the discourse.[46]
     We have seen in the previous sub-section that context plays a role in English. D-linked Wh may violate the Superiority Condition without rendering a question ungrammatical. But we have not clearly established which Wh must be D-linked. Context plays a much more constraining role in German and the issue of which Wh should be D-linked is much clearer: both Wh-elements must be D-linked; this I will refer to as DRQ. Consider the following:

---

    b.   ?   What did which student read?

     The conflicting judgements on the various sets of data (i.e. fronted bare Wh and D-linked Wh-*in-situ*, fronted D-linked Wh and bare Wh-*in-situ*, fronted bare Wh and bare Wh-*in-situ*, fronted D-linked Wh and D-linked Wh-*in-situ*) can only lead to the conclusion that something else besides D-linking is also at stake in English.
[45]   But note the contextual difference between (57) where only one of the Wh-referents is mentioned and (58) where both possible referents are introduced, as illustrated in (i):

(i)   It's nice to have all those activities scheduled for all those times,
       a.   … but when are you going to do what?
       b.   #   … but what are you going to do when?

If two bare Wh-phrases are used in the context of (58), only the order *when…what* is felicitous (presumably due to the proximity of the referent of the fronted Wh, namely *all those times*).
[46]   It is generally assumed that Wh-elements are quantificational. Under an interpretive approach to quantification (and hence Wh)—as suggested by Postma (1995), for example—a Wh does not carry its interpretation inherently but receives it; it is not unreasonable to assume that this reception is the result of checking a particular feature. While Wh-phrases are often taken to be indefinites, their interpretation may vary: they may be existential or quasi-universal. A D-linked Wh is quasi-universal; this, I argue, is due to the particular Top-feature involved. This restricts the quantificational force, i.e. it confines the quantificational interpretation such Wh-elements may receive.

(59)   Situation I: An author sold three of his top-selling biographies to three leading movie
       studios. Delighted, he calls his mother on the cell phone:
       "I sold *As I Lay Sighing: A Day in the Life of William Jefferson Clinton*,
Housekeeping: The Yeltsin Years and *Tropic of Virgo: Did You Know Suharto
Was a Virgin and Other   Interesting Facts about Indonesian Politics*."
    a.   Who bought what?
    b. * What did who buy?
    c. # Wer hat was gekauft?
    d. # Was hat wer gekauft?

The English questions show what we have seen already: a question involving two Wh-elements is well-formed if it obeys the Superiority Condition (59a), but ruled out when not (59b). It becomes clear that D-linking a bare Wh is not very easy; *what* in (59b) fails to be D-linked by the (limited) context set up.[47] The German equivalents in (59c-d), however, show something else: given the context, the question that asks for the distributive meaning of *wer* 'who' and *was* 'what' is infelicitous, regardless of which Wh-element is fronted.[48]
    Compare Situation I with another context, given in (60):

(60)   Situation II: Same author, same books but now he informs his boyfriend a little better.
       "I sold *As I Lay Sighing: A Day in the Life of William Jefferson Clinton*,
Housekeeping: The Yeltsin Years and *Tropic of Virgo: Did You Know Suharto
Was a Virgin and Other   Interesting Facts about Indonesian Politics*. And
guess what: Disney, DreamWorks   and Fox bought them."
    a.   Who bought what?
    b. * What did who buy?
    c.   Wer hat was gekauft?
    d.   Was hat wer gekauft?

In this context, the acceptability and grammaticality of the English double questions does not change from (59);[49] the German equivalents, on the other hand, are felicitous and well-formed, both with the Wh-subject preceding the Wh-object and vice versa.
    The difference between Situations I and II (or between the contexts set up by the questions) is that in one case, only the objects of selling are mentioned, while in the other both are named, the objects of selling and the buyers of them. I would like to propose that only in the latter case do both Wh-elements conform to DRQ: the sets of referents to both Wh-elements are given.[50]

---

[47]   Note that (59b) would not be grammatical even if the fronted Wh is an inherently D-linked Wh-phrase (ia):

(i) a. *   Which book did who buy?
    b.   Which of these books did who buy?

Although judgements vary (and I adopt those judgements that do not accept this type of construction), (ib) is fine where a partitive Wh is used. This observation will be addressed again below.
    Note further that in the limited context given, the Wh-subject cannot be D-linked:

(ii) a. #   What did which person/which company/what kind of company/… buy?
    b. #   Which book did which person/which company/what kind of company/… buy?

The subject can presumably not be D-linked due to the fact that the context has not mentioned anything about a possible set of buyers. (I mark the infelicity of these constructions with '#', as usual.)
[48]   Naturally, a pair-list answer would look as follows, which is a possible answer in English, but not in German:

(i) Fox bought *As I Lay Sighing*, Disney *Housekeeping* and DreamWorks *Tropic of Virgo*.

Obviously, (i) is a possible answer in German, but not in this context (where no reference to the buyers has been made).
[49]   It does with respect to D-linking of the subject; (ii) from fn. 47 is a possible option in this context.
[50]   In the original formulation (Grohmann 1997c, 1998b), DRQ was taken to refer to "exhaustive" lists of referents. This is too strong a claim, as pointed out by Gregory Ward:

(61) *Discourse-Restricted Quantification (DRQ)*
Questions involving two Wh-expressions are well-formed iff the referent sets of both Wh-expressions are determined by the context; determination of referent sets is satisfied by providing a list of at least two possible referents in the discourse.

In the sense of this informal definition of DRQ, both Wh-elements in German Multiple Wh must obligatorily be "D-linked" where D-linking is taken to denote sets of referents that are discourse-old in the sense that at least two referents have been introduced into the discourse. (The more interesting questions, namely why exactly at least two referents in constructions with exactly two Wh-elements, and how DRQ applies—if at all—to English have to be kept in mind for further research; I appreciate Juan Uriagereka's continuous puzzlement but cannot yet provide a satisfactory account.)

We can now skim through the previous examples from English and test their acceptability, felicity and, ultimately, grammaticality in German.

The following are approximate translations of (52b) and (53b):

(62) a. # Ich weiß, daß von all den Katastrophen in dieser Küche Jane die Bohnen verbrannt hat und Lydia Salz in den Eistee getan hat; aber *wás* hat *wér* zerbróchen? Ich weiß, daß jemand etwas zerbrochen hat, also hört endlich auf, meiner Frage auszuweichen.
      b. Ich weiß, daß wir Transistor A, Transistor B und Transistor C installieren müssen und ich weiß, daß diese drei Löcher für Transistoren da sind, aber der Teufel soll mich holen, wenn ich aus der Anleitung herausfinden könnte, wo was hin geht!

The judgements are as predicted, both in English and in German. While (52b) is rendered grammatical due to the context (in contrast to the ungrammatical (52a)), (62a) is not grammatical (felicitous) because it does not conform to DRQ: no reference has been made to the breakers or the broken items.[51] Likewise, (53b) is okay (as is (53a))—thus not showing much for English in the first place—and the German equivalent (62b) is, too: reference to both, the *where* and *what* has been introduced in the discourse.[52]

The other contexts given above for English concern (58) (and (i) in fn. 47), translated here:

(63) Es ist ja schön, all diese Aktivitäten zu all diesen Zeiten geplant zu haben,
      a. … aber zu welcher Zeit wirst du welche Aktivität machen?
      b. … aber welche Aktivität wirst du zu welcher Zeit machen?
      a. … aber wann wirst du was machen?
      b. … aber was wirst du wann machen?

(i) Situation (from (60) above): "… And some movie companies bought them."
     a. Wer hat was gekauft?
     b. Was hat wer gekauft?

In this modified version of (60), reference to *wer* 'who' is made only in passing, namely by referring to the buyers in a non-specific manner as *some movie companies*. Nevertheless, the Multiple Wh is well-formed with either Wh fronted, implying that DRQ is observed. In order to distributively interpret the question, each Wh needs to have at least two referents (otherwise, an echo interpretation might arise).

Indefinite quantifiers such as *some* imply such reference, and so do non-exhaustive sets such as "Disney and DreamWorks, for example…" in place of "Disney, DreamWorks and Fox…" in (60).

[51] This question is only felicitous when DRQ is observed, such as in (i) where the breakers and the broken items are mentioned prior to the question *aber was hat wer zerbrochen?* 'but what did who break?':

(i) … Ich weiß auch, daß Peter, Paul und Maria die Teller, Tassen und Schalen zerbrochen haben; aber…
     …*I   know also   that Peter, Paul and Mary the plates, cups   and bowls   broken   have; but…*
     '… I also know that Peter, Paul and Mary broke the plates, cups and bowls; but…'

[52] As we have seen in section 2.1.3, there is no asymmetry in English (nor in German) with *where/when* and Wh-arguments. But this is irrelevant here, as reference to both Wh-elements has been introduced.

As both, a reference to time and to activities, has been made (*all diese* 'all these', referring to something known to speaker and hearer), (63a-d) are all well-formed, felicitous questions: with the subject *which N*-phrase preceding the object and vice versa as well as either order for the bare Wh-phrases. (All other examples from above are fine when DRQ is strictly observed.)

3.1.3 More Data on Contexts in German
One might conclude that felicitous German Multiple Wh necessarily involve definite DPs or referring expressions, as can be seen in (60). This is not the case, however:

(64)   Situation III: A sex shop owner tells a reporter of items he sold and his customers.
       "I just sold a dildo, a magazine and a porno. A lesbian, an old guy and a pretty hooker bought all that stuff, but it really could be anyone coming in and buying stuff here."
       a.   Who bought what?
       b. *   What did who buy?
       c.   Wer hat was gekauft?
       d.   Was hat wer gekauft?

In this situation indefinite DPs denote the sold objects and the same grammaticality judgements apply as in (60). But on top of the information of sold objects (cf. (59)), the buyers of these items are also mentioned in (64), thus priming for a felicitous double question in German.

This supports the view that at least specificity is involved in D-linked contexts (see, among many others, Kiss 1993 or Comorovski 1996) and further that interrogative Wh-elements are indefinites (since Karttunen 1977). I will return to this discussion.

The main conclusion one can draw from the previous examples is that a felicitous Multiple Wh-construction in German depends on the context which may depend solely on the situation or more accurately, as exemplified here, on the information provided by the statement priming for the question. The quantification of the Wh-elements is thus restricted by the discourse, in other words: the Wh-elements underlie Discourse-Restricted Quantification.[53]

This is not the case in simple Wh-questions which extends to German as well. Compare the following with (59) from above:

(65)   Situation I: (from above)
       "I sold a book today."
       a.   Who bought it?
       b.   Wer hat es gekauft?

---

[53]   To come back to the remark from fn. 10 that the referent sets must be more than one for each Wh, it may be relevant to mention Uriagereka's (1988:209ff.) observation regarding multiple questions, as also brought to my attention by Howard Lasnik (personal communication). He distinguishes between "incriminatory" and "inquisitory" questions: while the latter is the type of Multiple Wh that we are interested in here, the former gives a framework for those types of Multiple Wh where each Wh has one referent and the purpose of the question is to assign an ordering among them.

Thus imagine a situation where two boys come crying to their mothers, obviously having beaten each other. If the mother asks (i), she knows exactly that one of the boys hit the other, but she does not yet know who started the fight:

(i)   Who hit whom?

This question pairs the referents of the Wh-elements but does not intend to inquire more (i.e. it does not yield a list), whereas in the cases shown here, an exhaustive pair-list answer is intended. It might be the case that Uriagereka's incriminatory question type extends further than the case illustrated here (or even in Uriagereka 1988). Note that in the German equivalent, only (i) is fine; the Wh-object fronted is ruled out. This clearly shows that these types of questions are different from the type of Multiple Wh under investigation here. It is not immediately clear in how much this distinction bears on the object of the present study.

These are speculations, however, and should be part of a broader investigation than can be pursued here. Thus, let us constrain the present goal to inquisitory-type Multiple Wh where the intended answer pairs a list of Wh-referents containing more than one referent for each Wh exhaustively.

In both cases, the simple Wh leads to the formation of a grammatical and felicitous question, although the set of referents is anything but known. This is the classical case of a Wh: one person does not know the referent set of some entity and inquires it from someone who knows; in this sense, the Wh is an indefinite expression.

Let me finally show that the discourse restriction set up by the information provided not only applies to a Wh-subject and a Wh-object, but more generally to all kinds of double questions, involving any two Wh-elements: arguments (DPs and PPs alike) as well as adjuncts (temporal, locative or any other) and combinations of all of them.[54]

In (66), the question seeks to get a distributive answer for the direct and indirect objects of a ditransitive verb.

(66)   Situation III: (from above)
       "This guy gave away his mag collection, a special ring and handcuffs."
       a.   To whom did he give what?
       b. *   What did he give to whom?
       c. #   Wem hat er was gegeben?
       d. #   Was hat er wem gegeben?

(67)   Situation III: (from above)
       "This guy gave away his mag collection, a special ring and handcuffs. Remember the biker lady? And this old fart and the young high school teacher? He gave it to them."
       a.   To whom did he give what?
       b. *   What did he give to whom?
       c.   Wem hat er was gegeben?
       d.   Was hat er wem gegeben?

The same restrictions we have observed above apply: both referent sets must be given in the discourse in German, but not in English (contrast with (67)).

The same restriction holds for the combination of an argument and adjunct:

(68)   Situation: Ray's jealous girl-friend tells her friend about some of his kissing locations.
       "Ray kissed in the car, behind the house and on the couch."
       a.   Where did Ray kiss who?
       b.   Who did Ray kiss where?
       c. #   Wo hat Ray wen geküßt?
       d. #   Wen hat Ray wo geküßt?

(69)   Situation: She informs another friend of Ray's kissing locations and the girls involved.
       "Ray kissed Mary, Jane and Emma. He did it in the car, on the loo and on the couch."
       a.   Where did Ray kiss who?
       b.   Who did Ray kiss where?
       c.   Wo hat Ray wen geküßt?
       d.   Wen hat Ray wo geküßt?

Again, we can observe the discourse restriction applying to a Wh-argument and a Wh-adjunct: in order to be felicitous, reference to both kissees and kissing locations must be made.

As can be seen more clearly in the German cases, the adjunct may also be a PP.[55]

(70)   Situation: Martin's aunt informs a friend of his thinking of certain people who died.
       "Martin thinks of his father, his mother and his brother."
       a.   When does Martin think about who?
       b.   Who does Martin think about when?
       c. #   Wann denkt Martin an wen?
       d. #   An wen denkt Martin wann?

---

[54]   Note that all examples provided here involve matrix questions. As Norbert Hornstein (personal communication) points out to me, embedded questions possibly behave differently. As this study can only be preliminary in many respects, deeper considerations of embedded questions must be left for further research.
[55]   It is irrelevant here whether it is an adjunct, an argument or a quasi-argument.

(71)   Situation: His aunt informs a friend of his thinking of those people at certain times.
       "Martin thinks of his father, his mother and his brother. He does so especially in
       March, May and July, when they died."
       a.   When does Martin think about who?
       b.   Who does Martin think about when?
       c.   Wann denkt Martin an wen?
       d.   An wen denkt Martin wann?

       Lastly, DRQ is also at work when the classical problematic adjuncts *how* or *why* (see
section 2.1) are involved. (72-73) show the interaction of *why/warum* and the (indirect)
object.
(72)   Situation: Someone informs someone else of Norbert's lending of a specific book.
       "Norbert lent the book to Peter, Paul and Mary."
       a.   Why did Norbert lend whom the book?
       b. *   Whom did Norbert lend the book why?
       c. #   Warum hat Norbert das Buch wem geliehen?
       d. #   Wem hat Norbert das Buch warum geliehen?
(73)   Situation: Now the information about the lending includes Norbert's reasons.
       "Norbert lent the book to Peter, Paul and Mary. One needed it for a reference, the other
       never read it before and the third guy needs to write a review and doesn't have it."
       a.   Why did Norbert lend whom the book?
       b. *   Whom did Norbert lend the book why?
       c.   Warum hat Norbert das Buch wem geliehen?
       d.   Wem hat Norbert das Buch warum geliehen?

       The various scenarios depicted here show that the restrictions laid upon the felicity
conditions of Multiple Wh by the discourse vary among English and German, in particular.[56]
Before moving ahead, I should mention that the primary interest here are the general
conditions upon questions; I am not concerned with pairing word-order sequences of answers
to word-order sequences of questions: thus, it is immaterial for present purposes whether
subject-object order in a question primes for subject-object order in an appropriate answer.[57]

3.1.4 Crucial Differences
The previous three subsections have presented us with a wide range of data from German and
English Multiple Wh. On the face of it, contextual constraints differ widely in both
languages. While in German Multiple Wh-constructions, the referent sets of both Wh-
elements must have been introduced in the discourse (DRQ), which then allows the speaker
to front either Wh, the situation in English is more complicated. In general, the Superiority
Condition must be obeyed. It may be overridden by contextual information, too, which has
been called D-linking in the literature. The exact role of D-linking is, however, still
somewhat blurry.[58]
       With D-linked and non-D-linked Wh-elements involved in double questions, four
possible orders arise, as already noted. These possibilities are furthermore enriched by
(language-internal) asymmetries between subject and object, for example, or certain Wh-
adjuncts. Concentrating on the subject-object asymmetries in English, the above data allow
the following categorizations for fronted non-D-linked Wh-elements:

---

[56]   Especially the last case (72-73) might give us a clue that *why/how* and *warum/wie* are syntactically
completely different. I will suggest below something along the lines that the German counterparts may move,
unlike *why* and *how*. This is in line with suggestions from the literature and captures variation among German
speakers as well.
[57]   See, for example, Kiss 1993 who uses this test to distinguish properties of Multiple Wh in English vs.
Hungarian; also, Kuno (1982) proposes the "Sorting Key Hypothesis" which makes reference to linear order of
the Wh-elements and the corresponding answers. (On the latter, see the discussion in Kuno and Takami
1993:112ff. and Comorovski 1996:109f.)
[58]   DRQ, as understood here, may be seen as a special kind of D-linking. In how much it concerns the same
phenomenon is another question, though. Ultimately we would like to know why discourse restricts well-
formedness of Multiple Wh cross-linguistically. While some suggestions are made here, deeper issues may
involve mental representations of the respective Wh-elements in the lexicon of speakers. The current treatment
can only be seen as one attempt to uncover the special role of D-linking in this respect.

(74) a.   *Non-D-linked Wh-SU… Non-D-linked Wh-OB*
          Who bought what?
      b.   *Non-D-linked Wh-SU… D-linked Wh-OB*
          Who bought which book?
      c. * *Non-D-linked Wh-OB… Non-D-linked Wh-SU*
          What did who buy?
      d.   *Non-D-linked Wh-OB… D-linked Wh-SU*
          What did which girl buy?

Superiority is obeyed in (74a) and (74b), where the Wh-subject is fronted; it is apparently violated in (74d), where the Wh-object is fronted; in this case, however, the Wh-subject is D-linked. That D-linking of the non-fronted Wh is crucial with respect to Superiority can be seen from (74c) which is ungrammatical.

The same list can be construed for fronted D-linked Wh-phrases:

(75) a.   *D-linked Wh-SU… D-linked Wh-OB*
          Which girl bought which book?
      b.   *D-linked Wh-SU… Non-D-linked Wh-OB*
          Which girl bought what?
      c.   *D-linked Wh-OB… D-linked Wh-SU*
          Which book did which girl buy?
      d. * *D-linked Wh-OB… Non-D-linked Wh-SU*
          Which book did who buy?

The inverse is the case here: a D-linked Wh-object cannot be fronted over a non-D-linked Wh-subject (cf. (75c) vs. (75d)). In other words, English questions involving a Wh-subject and a Wh-object seem to not permit a non-D-linked subject Wh to remain *in situ*. Interestingly, even this generalization can be violated:[59]

(76) a. * What did who buy?
      b. * Which book did who buy?
      c.   Which of these books did who buy?

In (76a), the bare Wh has no properties that allow it to be fronted over the Wh-subject, while the Wh in (76b) is at least D-linked; this, however, is not enough in English to allow fronting over the bare (= non-D-linked) Wh. (76c), on the other hand, shows that fronting of a partitive Wh-object over a non-D-linked Wh-subject is grammatical.

The upshot of the intricate English data is then that the Superiority Condition rules out fronting of a lower bare Wh over a higher one; this movement can only occur to check a Wh-feature, and the higher Wh is closer to CP than the lower one (MLC). Superiority effects can be obviated by D-linking, where no Wh-movement applies; constraints on ordering Wh-elements underlie unselective binding (D-linking). Furthermore, partitive Wh-phrases may move even if another Wh is apparently closer to CP; this, I will show, can also be found in German (DRQ). I will return to all three accounts after a thorough discussion of a syntactic analysis of DRQ and subsequently "Wh-topics."

The observations concerning German Multiple Wh can be summarized as follows. If the possible referent sets for both Wh-elements has been introduced in the discourse, anything goes; if no or only one referent set has been established, nothing goes (DRQ). Unlike English, German allows this special kind of D-linking for all Wh-elements.

---

[59] I would like to thank Jennifer Graham for pointing out to me the special role that partitives play. I will return to this contrast below. While it has been observed that D-linking and specificity, and even partitivity, are connected in some sense, there is no consensus about this interaction; furthermore, to my knowledge, the contrast in (76) has not been noted in the literature. Kiss (1993) notes that partitives are easier to extract out of Wh-islands, as is well-known by now; Wh-islands will be addressed in section 4.2. I will return to the role that partitives play in general below.

The task is now to (i) give a syntactic account for DRQ, (ii) find evidence for this approach from other languages, and (iii) establish a connection between the variation in Multiple Wh among languages. I suggest to (i) understand DRQ as topicalization, (ii) look at Chinese for obvious instances of Wh-topics, and (iii) consider the different syntactic realization of semantic and/or pragmatic factors across languages.[60]

3.2   DRQ: Multiple Wh Is Topicalization

After clarifying the issues involved with Multiple Wh-constructions and D-linking/DRQ, I will propose to analyse DRQ as overt topicalization, i.e. that in German Multiple Wh-questions all Wh-phrases move overtly to the specifier of (a recursive) TopP, situated between AgrSP and CP (as part of the CP-domain following Rizzi (1997) and others). As a consequence of how DRQ has to be interpreted, some (related) instances of English D-linking will be shown to also involve topicalization.

3.2.1 Outset: The Problem
It is by now pretty well established that in a simple Wh-question the Wh moves to SpecCP. Languages differ in this respect (see section 2.1.1) whether they move overtly (before Spell-Out) or covertly (at LF). This is also the standard view for Multiple Wh-questions. However, D-linked Wh-phrases are not assumed to move to CP. One property of D-linking is that it does not require movement for interpretation, but rather that the D-linked Wh is interpreted *in situ*, via unselective binding (following Lewis (1975) and Heim (1982) regarding unselective binding for indefinites as well as Pesetsky (1987) and subsequent work for implementation to the syntax of Wh-elements).

With the previous discussion in mind, some problems immediately arise. For one, it is not exactly clear what "unselective binding" means in a minimalist framework**.** Or more to the point, is it really enough to say that some Wh-phrases receive interpretation by movement to SpecCP, while others do not move anywhere at any stage? What is it that allows discourse to bind something unselectively? It would be nice if we could relate the special discourse-properties to something else.

In the minimalist framework, where relevant properties are understood as morphological features, D-linking must correlate to some feature. While it has been argued that D-linking relates to specificity (Kiss 1993), others argue that it correlates to partitivity (Comorovski 1996).[61] If a D-linked Wh would be a partitive Wh, we would not be able to account for the contrast in (75), repeated slightly extended below as (77-78):

(77) a. * Which book did who buy?
       b. * Which books did who buy?
(78) a.   Which of the books did who buy?
       b.   Which of these books did who buy?

I thus assume that D-linking involves "specificity" (à la Enç 1991). In the feature-checking framework adopted here, specificity is a feature encoded on the relevant Wh-phrase. Without delving too deep into the syntax and semantics of specificity, this property does not force movement to a specially designated position—unlike movement to check Wh- or nominal features, for example; it thus is sufficient for a D-linked Wh to remain *in situ* and check the

---

[60]   The task of providing a principled distinction between German and English (or even among German dialects) will be left virtually unexplored in this paper. What might be of interest is an observation by Bruce Morén (personal communication) about dialectal variation in English. Apparently, some dialects (his among them) allow for fronting of a lower bare Wh-element (such as (1b)). Curiously, though, this is only a legitimate construction when it occurs in the contexts shown here for German Multiple Wh; furthermore, if not in these contexts, the standard Multiple Wh (i.e. (1a)) cannot asked either. It thus looks like DRQ is at work in certain English dialects in the same manner as in German. This suggests something about DRQ to be on the right track.
[61]   David Pesetsky (personal communication) claims that the relevant property of D-linked Wh-phrases is not specificity but definiteness. However, I have not seen the relevant empirical data or theoretical support for this view. The standard definitions of both properties are vague enough to allow for this assumption; given that I have not seen proposals in favour of definiteness in the literature (unlike specificity or partitivity), I will ignore this for the remainder of the discussion.

Wh-feature at LF (if checked at all, viz. Pesetsky 1987). "Unselective binding" might then be understood as a property of the discourse (possibly encoded in ForceP, following Rizzi (1997)) to assign the interrogative interpretation correctly—such as the "Sorting Key" (Kuno 1982), perhaps[62]—given an otherwise satisfying derivation. This part of Multiple Wh is not the most interesting one in the present study, and I refer the reader to Kiss' approach of correlating specificity and D-linking.

Following Pesetsky, D-linking does not involve Wh-checking. Whether it involves checking of some other features is not relevant here. What is crucial is that English D-linking does not involve the same overt operations that English D-linked partitive Wh-phrases undergo or German DRQed Wh-phrases. That some sort of movement must take place becomes clear from examples such as (17b), i.e. from those cases where a D-linked object is fronted over a D-linked subject: the object must sit in CP as can be deduced from the presence of *do*-support:

(79) a.   Which book did which student read?
      b. * Which book which student read?

The syntactic representations of Multiple Wh in English involving D-linked Wh-phrases thus look (at least) something like the following:

(80) a. $[_{CP}$ Which student$_i$ $[_{IP}$ t$_i$ $[_{VP}$ t$_i$ read which book]]]?
      b. $[_{CP}$ Which book$_j$ did $[_{IP}$ which student$_i$ $[_{VP}$ t$_i$ read t$_j$]]]?

What is more interesting is how German Wh-elements differ from English ones in Multiple Wh. The constraints on the well-formedness of German Multiple Wh that DRQ inflicts seem to be at first sight intimately related to D-linking: the referent sets have to be somehow established. It goes a step further, however, in that in every Multiple Wh both Wh-elements must be "D-linked"; this special case of D-linking is, I argue, topicalization which allows movement for feature-checking purposes where discourse plays a crucial role.

I take these interrelated issues as the outset to reconsider the syntax of Multiple Wh not only in German and English but across languages. The hypothesis is that not all Multiple Wh are the same; discourse factors play a role. From what we have seen so far, we can say with certainty that some of these factors have not been looked at, namely the role of context in German Multiple Wh with bare Wh-elements. With the limited set of languages considered in this study, it becomes clear that this hypothesis extends to other languages as well, basically revolving around the distinction we have established already.

3.2.2 DRQ and Wh-Topics

A Wh-element in German Multiple Wh can only have a referent previously established in the discourse. In this sense, it must be specific. Specificity alone, however, cannot distinguish successfully the difference of D-linked Wh in English and DRQed Wh in German.

Specificity of certain Wh-elements implies a different semantics involved in Multiple Wh, but it has little if anything to say about the syntax. Correlations have been drawn between such discourse-restricted (= D-linked) Wh, specificity and partitivity. These correlations are mainly driven by semantic properties. However, partitivity may also lead towards a different syntax (although rather tacitly). Consider the structure of a partitive DP:[63]

---

[62]   Kuno (1982:141) defines the Sorting Key as follows:

(i) In a multiple Wh-word question, the fronted Wh-word represents the key for sorting relevant pieces of information in the answer.

    Thus, discourse plays a role in establishing the conditions whether the higher or lower D-linked Wh must be fronted (cf. (57-58) above and discussion).

[63]   The structures depicted in (81-82) are crude descriptions, of course. Most accounts take into consideration semantic and pragmatic properties and hence constraints on the syntactic distribution of partitives, but not on the particular structure that such elements should be assigned. A useful overview of the partitivity research can be found in de Hoop 1998, most of which is not utterly relevant to the present issues; explorations into finer structure of DPs along the lines suggested here, can be found in the works by Uriagereka (1993), Castillo (1998) and Muromatsu (1998).

(81) a. $[_{DP}$ many Q $[_{NP}$ students]]
    b. $[_{DP}$ many Q $[_{Part}$ of those $[_{NP}$ students]]]
(82) a. $[_{DP}$ which Wh $[_{NP}$ student(s)]]
    b. $[_{DP}$ which Wh $[_{Part}$ of those $[_{NP}$ students]]]

What is relevant here is that the specifier of the entire DP may be a quantifier for an indefinite (81) which also includes a Q-feature, or a Wh-phrase with its Wh-feature for an interrogative (82). Furthermore, partitivity is somehow encoded on certain DPs—here labeled simplified as "Part" but more generally (under minimalist assumptions) to be conceived of as a feature internal to the entire partitive DP. The following is one attempt to capture this as a morphological feature.

    I will expand on this assumption with explicit consequences for the syntax. Let us give this bracketing more content in a representation where much more goes on than I can show here:

(83) a. $[_{DP}$ which [+Wh] of $[_{DP}$ those [+Top] students]]
    b.
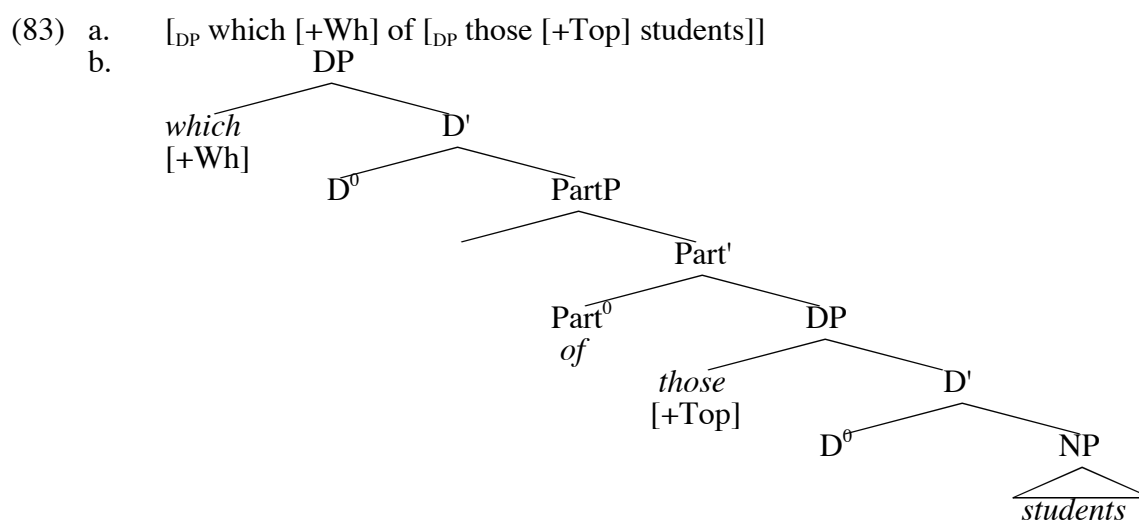
The bracketed structure in (83a) shows the unchecked features, namely those DP-features that need to engage in a feature-checking process in an appropriate configurations with another head bearing the same feature. This is clarified a little in (83b): partitivity itself is satisfied in its own phrase, part of a deeper DP-structure; specificity is probably saturated inside the lower (specific) NP *students*. These features are thus checked DP-internally (like nominal NP-features are checked DP-internally). Furthermore, I take [Top] to be a feature of the complement of Part^0; in how much this assumption can be verified across a larger set of partitive construction is beyond the scope of this study.[64] As a result, the two features construed with the two DPs remain to be checked via movement. As both DPs are part of the same DP-constituent, the assumption is that [Top] percolates to the top of the projection and is the feature that needs to be checked first.

    Applying the structure in (83b) to all partitive interrogative phrases, DRQed Wh-phrases contain two features that need to be checked via movement: a Wh- and a Top-feature. As the Top-feature needs to be checked prior to the Wh-feature, the relevant functional projection to look at is TopP.

    The part that has been assumed to denote specificity and partitivity will thus be taken one step further: it represents given information—a property well-known from topicalized elements. Analysing a DRQed Wh as a topic has syntactic consequences alongside the well-established semantic ones. Are these consequences desired at all? Do we need them?

    The answer to these questions will, of course, be "yes." Before proceeding, let me briefly mention three cases in support of this view, one of which we have seen already (with

---

[64]   The correct, more detailed structure of these are not relevant here; no claim is made that this is enough of a representation. The important properties of (83b) to note are the two DP-features ([Wh], [Top]) and the (not unreasonable) assumption that all other features are satisfied.

the other two to be outlined in more detail below). I take these observations to be good reasons to pursue an approach to Multiple Wh based on topicalization. Details will be laid out in the following.

Firstly, we have seen that contextual restrictions apply to the felicity—and ultimately grammaticality—of Multiple Wh-constructions. Thus different semantic and pragmatic properties directly lead to two different syntactic outputs (as described in detail in section 3.1 above).
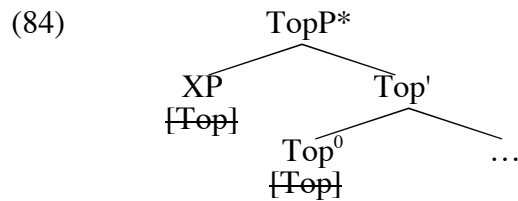
Secondly, there is good evidence in Chinese that certain constructions involving a Wh are actually instances of Wh-topicalizations. On top of the evidence in favour of such an analysis, a closer look at the discourse restrictions upon these constructions shows striking resemblances with those operative in German. I will lay out Wu's (1996) study in some detail in section 4.1.1 below.

Thirdly, Wh-extraction out of Wh-islands seem to be cross-linguistically only possible with "heavy" or D-linked Wh-phrases (see Comorovski 1996:ch. 5 for discussion). Looking at this phenomenon, we will be able to make generalizations about the possibility of extraction on the grounds of different structural positions—and hence, different syntactic processes—involved. I will return to this issue in section 4.2 below.

3.2.3 The Rough Syntax

Let me propose a rough outline how to implement topicalization to German Multiple Wh.[65] The restriction on possible referents has the flavour of a topic; namely, it refers to given information. As implied by the definition of DRQ, this property seems to take scope over the quantificational force of the Wh-element: it restricts it (see fn. 46 above). In a (minimalist) framework that expresses all relevant information in terms of formal features, topichood is an instantiation of positioning the topic in such a way that it can be licensed, namely by entering it in a specifier-head relationship with a topic-feature.

The formal feature [Top] heads its own projection. The locus of this is inside the CP-domain (following Rizzi (1997)) and the structural configuration of TopP looks as follows:

(84)

$$\begin{array}{c} \text{TopP*} \\ \diagup \quad \diagdown \\ \text{XP} \qquad \text{Top}' \\ \text{[Top]} \quad \diagup \quad \diagdown \\ \text{Top}^0 \qquad \dots \\ \text{[Top]} \end{array}$$

The topic-feature, which is part of the feature-bundle of XP (the topic), must be checked off against the topic-feature of the head; this is indicated in (83) by strikethrough typeface.[66] As shown by Rizzi, languages may employ more than one topic per sentence (thus having multiple topic projections, indicated as recursive TopP*); they may also differ with respect to the position of TopP* relative to CP: in some languages TopP precedes CP, in others TopP follows CP—and yet other languages may have TopP* preceding and following CP; in principle, there could also be languages where topichood is not checked in its own projection in the left periphery. German has the option of moving a topic to a position preceding as well as following CP (such as Left Dislocation, as proposed in Grohmann 1997a), with preference for the latter; in both, English and German, a topic may also appear base-generated before CP: we can witness such constructions in Hanging Topics (see Grohmann 1997a for more and the collection of papers in Anagnostopoulou et al. 1997), as will be discussed below.

In the case of a Wh-phrase subject to DRQ, the restriction is expressed on the Wh-phrase and can be conceptualized at least as follows (cf. (83) above):

(85)   [Wh [Top]]

Thus, whenever the set of referents (that is, referents that are part of a felicitous answer to a given question) is restricted, this information is encoded on the Wh-phrase and needs to be checked off. Moreover, the encoding can be thought of as a topic-feature; as the Top-feature needs to be checked off before the Wh-feature, it is the former feature only that needs to be checked (see (83b) and discussion above). This drives DRQed Wh-phrases to move to the specifier of TopP, prior to further movement to satisfy the Wh-feature on C^0.

The following shows a sample derivation (cf. (48c)):

(86)
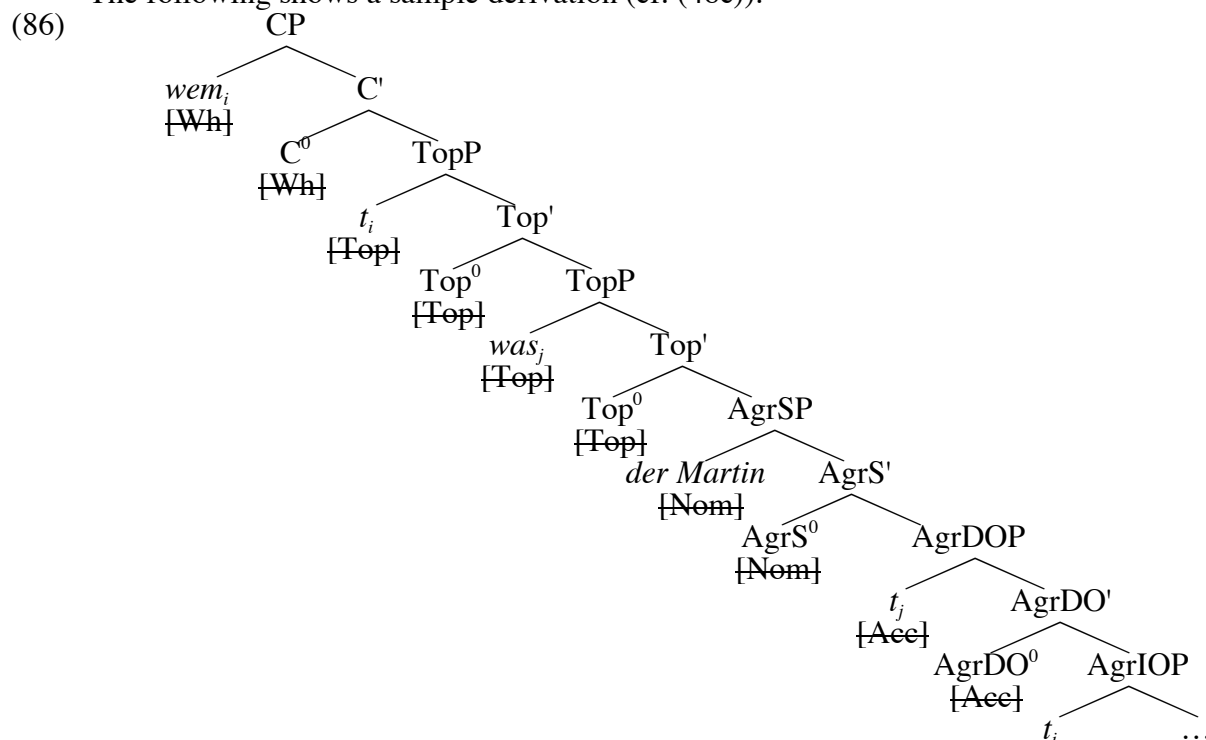
As already laid out in section 2.3.3, "scrambling" results in argument-movement to the specifiers of AgrP.[67] When AgrOP enters the derivation, either object may move there (Equidistance), giving it the appropriate label (AgrIOP or AgrDOP); subsequent movement of the subject results in formation of IP (AgrSP). Next, Top^0 enters the derivation with a strong feature, attracting the closest element that is also marked [Top]. (I will return to this step and the role of Equidistance below.) With three possible candidates (*der Martin, wem, was*) this can only concern either Wh: both are marked with a Wh- and a Top-feature (DRQ). Top^0 attracts the closest, *was*, which subsequently moves to check [Top]. Then another Top^0 enters the derivation, attracting *wem*. When C^0, marked with a strong Wh-feature, finally enters the derivation, it attracts the highest element that also contains [Wh]; in this case, *wem* is the closest and it moves to SpecCP.

We have seen already how movement prior to Wh-movement affects violability of the Superiority Condition. We have also seen that scrambling plays a role in languages that exhibit scrambling. Wiltschko (1997) argues that scrambling triggers D-linking in German and thus obviates Superiority effects. D-linking on its own does not seem to be enough; first and foremost, it could not account for the differences between English and German. More is at stake which I identify as a necessary syntactic process on top of the semantic and pragmatic properties of such Wh-phrases; topicalization as I argue. This follows from DRQ which ultimately is the crucial difference between "common D-linking" (English *which*-phrases) and "special D-linking" (German DRQ). Hence, scrambling alone cannot not

---

[67]   To repeat, it is irrelevant whether the arguments receive their respective Case or check any other feature.

enough. But the connection to scrambling arises as a corollary: scrambling over the subject is independently taken to be topicalization (following Grohmann 1996, 1997d).[68]

Suppose that pre-subject scrambling is topicalization. Suppose further that DRQed Wh-phrases are topics. Add the observation that scrambling obviates Superiority effects. The result is that DRQed Wh-elements scramble over the subject to TopP where they satisfy the Top-feature, induced by DRQ, and Superiority effects disappear.

Given our assumptions regarding German clause structure (sections 2.3.1 and 2.3.2), this accounts straightforwardly for the lack of Superiority violations in a number of Multiple Wh-constructions. Let us look at the derivations of a sample of German Multiple Wh-constructions.

First, the often noted lack of Superiority effects when two Wh-object are involved. As we have seen already, either object may precede the other:

(87) a. $[_{CP}$ Wem$_i$ hat $[_{TopP}$ t$_i$ $[_{TopP}$ was$_j$ $[_{AgrSP}$ der Martin $[_{AgrDOP}$ t$_j$ $[_{AgrIOP}$ t$_i$ gegeben]]]]]?
      b. $[_{CP}$ Was$_i$ hat $[_{TopP}$ t$_i$ $[_{TopP}$ wem$_j$ $[_{AgrSP}$ der Martin $[_{AgrIOP}$ t$_j$ $[_{AgrDOP}$ t$_i$ gegeben]]]]]?

This is a clear case that both Wh-elements have moved far beyond their VP-generated positions; both are fronted over the subject in the overt syntax. Both Wh-phrases scramble out of VP into their respective specifier-positions. As argued in section 2.3.1, these can be thought of as being freely ordered (Equidistance). Once the subject has moved to AgrSP, the functional head Top^0 is introduced into the derivation which attracts the closest element bearing a Top-feature (MLC), which in the case of (87a) is *was*; the Wh moves to SpecTopP (but see below). Next, another Top^0 enters the derivation, itself attracting the closest element with an unchecked Top-feature; this time, *wem* moves from the lower AgrOP to the new specifier position. Regardless of how V2 is derived, the final structure at Spell-Out is that shown in (87a) where the highest topic checks the Wh-feature. An alternative question is derived in (87b), where the Wh-phrases scramble in a different order. In both cases, all moves conform to the MLC.

Checking of the Top-feature prior to the Wh-feature in the overt syntax satisfies DRQ—the movement-based D-linking variant of Multiple Wh in German.

Wh-elements in German Multiple Wh are Wh-topics and undergo a derivation in accordance with DRQ. This means that the underlying, deeper structure associated with these Wh-elements is invariably that of partitive Wh-phrases, i.e. the one shown in (83b). DRQ is then a more powerful version of D-linking: while the latter has the force to interpret Wh-elements, the former assigns deeper syntactic structure based on semantic and pragmatic properties.[69] But English may also employ this strategy: namely, in exactly those cases, where the Wh is inherently partitive (83a).

(88) a.   Which of these students bought which of these books?
      b.   Which of these books did who/which student/which of these students buy?

(89)
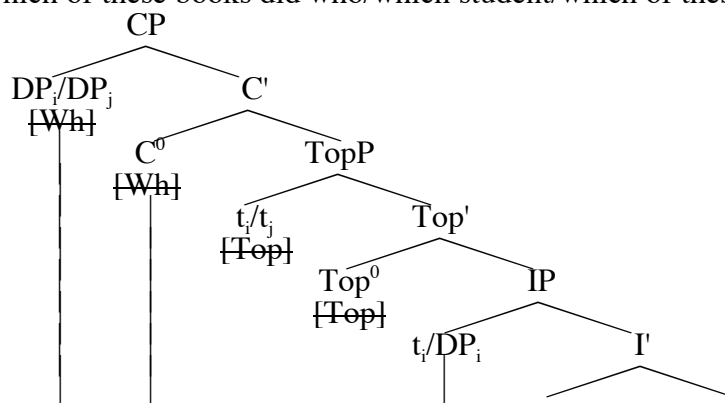
$$I^0 \qquad VP$$

$$t_i \qquad\qquad V'$$

$$V^0 \qquad DP/t_j$$

a.	*Wh-SU*			*bought*	*Wh-OB*
b.	*Wh-OB*	*did*	*Wh-SU*	*buy*	
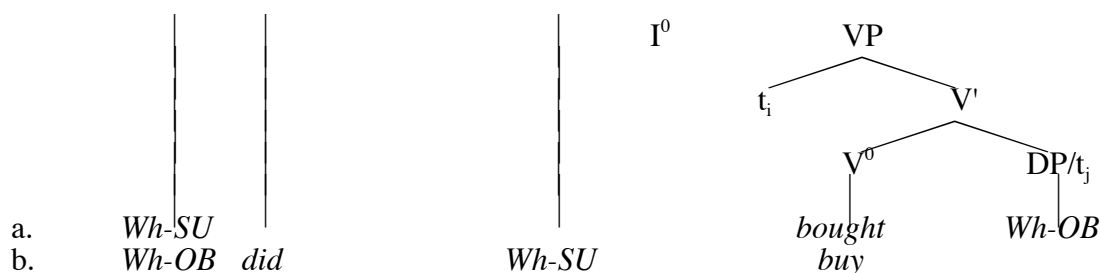
On analogy with the previous derivation, (89) represents the derivations for (88).[70] The partitive Wh needs to check both, the Wh-feature and the Top-feature. The generalization, then, is that all German Wh-elements in Multiple Wh-constructions (i.e. under DRQ) have the pragmatic, semantic and syntactic properties of partitives/topics, while only overt partitive Wh-phrases have these properties in English. Further restrictions apply which I will return to.

Let me briefly mention additional evidence—tentative, though, they may be—from English that the general Wh-Topic approach suggested here can be extended to cover certain facts in English, too.[71]

(90) a.   What kind of bike, it was?
      b.   Hey, hey LBJ, how many did cigs you smoke today?

If some specially D-linked Wh-phrases in English may also contain [Top], one might analyse (90a) as an "incomplete" derivation where for some reason the last step, i.e. that of checking the Wh-feature, was left out. This results in a Wh-question with the surface structure of a topic-construction. Likewise, (90b) can be understood as checking of the Top-feature by the whole Wh (*how many* cigs) and further Wh-checking by only the interrogative part (*how many*); here, the topicalized part (*cigs*) stays in TopP.

3.3   Extensions of Wh-Topics

The above exposition reveals the direction that the idea of DRQ conceived of as topicalization of Wh-phrases heads in, but it needs refinements. Some of these will be provided here.

3.3.1 Subjects and Multiple Wh: Some Consequences

One question that immediately arises is whether we can decide which Wh-phrase is topicalized or whether there can be more than one Wh-topic. Given the conception of DRQ for double questions as presented in (61), we would expect both Wh-elements to move to SpecTopP. The sample derivations shown above build on this assumption also. This is precisely what the following discussion will be concerned with.

To yield a derivation that conforms to the MLC it was assumed that the Wh-phrase from the lowest AgrP moves to the highest SpecTopP. This is what we have seen in (87) already.

(91) a.   [$_{CP}$ Wen$_j$ hat [$_{TopP}$ t$_j$ [$_{TopP}$ wem$_i$ [$_{AgrSP}$ der Martin [$_{AgrIOP}$ t$_i$ [$_{AgrDOP}$ t$_j$ vorgestellt]]]]]?
              whom has   to-who   the Martin   introduced
              'Whom did Martin introduce to who yesterday?'
      b.   [$_{CP}$ Wem$_j$ hat [$_{TopP}$ t$_j$ [$_{TopP}$ wen$_i$ [$_{AgrSP}$ der Martin [$_{AgrDOP}$ t$_i$ [$_{AgrIOP}$ t$_j$ vorgestellt]]]]]?

---

[70]   Some details are completely irrelevant here: I thus label the subject-position in English IP and employ a very simple VP-structure, which suffices to show that the subject moves from its VP-position to a higher position, before moving to TopP and CP (analogously to German) in one case, and the object moves overtly to TopP—as in German—and then to CP in the other. It is also irrelevant which features are checked when, especially whether the object checks Case- or D-features on its way to TopP.

[71]   These examples were found by the author. (79a) was overheard from a policewoman in College Park, MD when inquiring about a stolen bike; by the time of uttering the question it was already established that the topic of the discourse was a specific bike. (79b) was found in *The Smoking Life* by Ilene Barth (Columbus, MS: The Genesis Press, 1997) on page 80 as cited here. If nothing else, the reader is encouraged to smile at this point;-)

The two Wh-phrases in (91a) are ordered direct object—indirect object. The surface derivation is derived by ordering AgrIOP above AgrDOP. The opposite is the case in (91b). On the other hand, free ordering of AgrOP has so far been justified with the Equidistance Principle. It is time now to go through some technicalities regarding Equidistance and the MLC.

First, we are confronted with another question. If German Multiple Wh is always an instance of topicalization (driven by DRQ), should the subject not also move to TopP? The reasoning behind this is, of course, that if two non-subject Wh-phrases occur with the subject in between and both Wh-phrases are in TopP, we need to account for the subject-position.

Consider (92), where a non-Wh-subject occurs in between two Wh-elements:

(92)   a.   Was hat Martin wann gekauft?
            what has Martin when bought
            'What did Martin buy when?'
       b.   Was hat Martin wem gegeben?
            what has Martin whom given?
            'What did Martin give to whom?'

For cases such as (92), I do indeed propose that the subject moves higher to the specifier of a further TopP. We assume the subject-position to be fixed, an assumption that is certainly shared by other linguists (starting in the minimalist framework at least with Zwart 1993 where the subject needs to move to AgrSP for feature-checking). But in cases such as (92), it is plausible that the subject moves further, once the relevant features are checked by $AgrS^0$.[72]

I thus take the subject to be inherently marked with the property of (thematic) topichood which it can "check off" in AgrSP. The reason for the inverted commas is that I take this checking process to be inherent, not formal. The underlying assumption is the thematic dominance of the subject which manifests itself by virtue of being clause-initial. When a clause is not subject-initial, this inherent thematic checking is overridden: in this case, the subject has to move to the specifier of a TopP in order check the relevant thematic feature; and only in this case, as subject-initial declarative clauses do not invoke the CP-structure.

A Wh-subject may be interpreted as "Wh-in-Agr," without further movement for exactly this reason. But when there is one Wh-phrase preceding and one following the subject, AgrSP does not satisfactorily serve as a site of interpretation anymore, hence further movement is required; neither is AgrSP sufficient to mark the subject's dominance if an element enters the CP-domain without preceding the subject at Spell-Out (under the current assumptions of why and how elements enter the CP-domain).[73]

Equidistance, as assumed here (section 2.3), allows either object to move out of their generated VP-position into the specifier-position of AgrOP first in ditransitive constructions. The reasoning behind this is, of course, that AgrOP enters the derivation "underspecified" in that it receives its ultimate label (AgrIOP or AgrDOP) through spec-head agreement with either an indirect object or a direct object, respectively. Once the verb adjoins to $AgrO^0$, the minimal domain of the complex head $AgrO^0$ is VP. This makes the two possible candidates—indirect object and direct object—equidistant, meaning that either one may move first.

---

[72]   This is also motivated independently: as is well-known, the subject always bears a certain notion of topichood, in the vague sense of "thematic topic" as outlined here (for example, see Erteschik-Shir 1998 for an elaborate discussion and consequences for her focus-structural approach to synthesize syntactic and semantic conditions, to name just one recent study; also see Comorovski 1996 for a similar concern regarding subjects and Wh-questions) which in the common cases receives its interpretation in the subject position.

[73]   This is part of what Comorovski had in mind when formulating D-linking in terms of leftmost Wh-phrases. That subjects are highly thematic is well-known: in many languages it is the subject that carries old, given information. Languages that exhibit a strong left periphery (in having frequent movement operations targeting functional projections above the subject)—such as German or Romanian—, this property might be more dominant than in others (such as English). In English, where topics need not always be realized by movement, the intrinsic topichood of the subject can be maintained more easily; as a consequence, German subjects must also move in some cases.

With respect to an intervening non-Wh-subject, as in (92), Equidistance applies again, and again it is the special status of AgrP that plays the crucial role. While in the former case AgrOP receives its label through spec-head agreement with the relevant argument, in this case the label is "open" again; moreover, all labels of AgrP are open, as notions such as "subject," "indirect object," or "direct object" do not play a role once the features are checked off (presumably happening for LF-purposes as suggested in Zwart 1997b). If then the complex Agr-head adjoins to Top^0, all material dominated by AgrP—the minimal domain to Top^0—is equidistant and either one may move first. In other words, after agreement features in the highest AgrP, AgrSP, are checked, AgrSP, AgrIOP and AgrDOP collapse to a formally equivalent AgrP.

Following the preceding exposition, (92) is derived as follows:

(93) a. $[_{CP}$ Was$_i$ hat $[_{TopP}$ t$_i$ $[_{TopP}$ Martin$_k$ $[_{TopP}$ wann$_j$ $[_{AgrSP}$ t$_k$ $[$t$_j$ $[_{AgrOP}$ t$_i$ gekauft$]]]]]$?
      b. $[_{CP}$ Was$_j$ hat $[_{TopP}$ t$_j$ $[_{TopP}$ Martin$_k$ $[_{TopP}$ wem$_i$ $[_{AgrSP}$ t$_k$ $[_{AgrIOP}$ t$_i$ $[_{AgrDOP}$ t$_j$ gegeben$]]]]]]$?

In (93a), the relevant operation takes place once AgrSP is formed (by movement of *Martin* to its specifier), assuming that *wann* is adjoined somewhere inside AgrP: *wann* moves first to TopP, *Martin* second and *was* third (all being equidistant); in (93b), the first move at the same timestep is by the indirect object, followed by the subject and lastly the direct object. In both cases it is the highest Wh-topic that moves to SpecCP, conforming to the MLC.

With these technicalities in mind, there are a number of other possibilities for the equidistant elements to move, as shown in (94):[74]

(94) a.   Wann hat Martin was gekauft?
      b.   Was hat wann Martin gekauft?
      c.   Wann hat was Martin gekauft?

On the other hand, (94') shows some derivations that are ill-formed under the same interpretation for the obvious reason that the last step of the derivation—movement of *Martin* to SpecCP—is illegitimate. (Alternatively, (94') may be ruled out due to absence of CP altogether.)

(94') a. * Martin hat wann was gekauft?
      b. * Martin hat was wann gekauft?

In sum, subjects are regarded as highly thematic elements by themselves. While under normal conditions, this thematic property may be realized in the canonical subject-position, in cases of interleaving with topics, the subject must be topicalized in the overt syntax in order to check its Top-feature. This means for Multiple Wh that both Wh-phrases as well as the subject may sit in TopP, depending on whether the subject's topichood may be saturated.

3.3.2 Scrambling and Multiple Wh: In Response to Wiltschko (1997)

With this outline of the approach in mind, let us look at data concerning the interaction of scrambling and D-linking. Wiltschko (1997:107) gives an apparently supporting example that also reveals the presence of Superiority effects in German:

(95) a.   Wer hat denn schon oft   was gesehen?
            who has PRT   already often what seen
            'Who has already often seen what?'
      b. * Was hat denn schon oft wer gesehen?

Following the assumption that the particle-adverbs used in (95) are VP-related—i.e. their presence indicates that the following material is VP-internal, including arguments—(see Diesing 1992), Wiltschko argues that the Wh-subject in (95b) has not scrambled, cannot be D-linked and hence induces a Superiority violation on the assumption that only D-linked Wh-elements may violate Superiority and D-linking is established by scrambling.

However, it is not so clear that the adverb-particle combination in (95) occupies a fixed structural position. Haiden (1995), for example, shows convincingly that what can be said is

---

[74] (94) shows the three other possibilities for (93a); equivalent combinations can be construed for (93b).

that sentential adverbs are positioned high and manner adverbs low; adverbs of quantification sit higher than manner adverbs and negation and also above the landing site of scrambling.

With this in mind, (95) does not really support what Wiltschko intends to show:[75] it does not show that *was* cannot have scrambled in (95a), and it thus cannot account for why (95b) should be ungrammatical. Rather *wer* has scrambled out of VP and (95b) has to be explained otherwise. The following discussion will make this even clearer.

Adverbs in the German middlefield vary a great deal with respect to their structural positions. Haiden (1995:50ff.) convincingly uses a class of (manner) adverbs which really seem to be "low" in the structure, very close to VP. This class contains *gut* 'well', *schlecht* 'badly, *richtig* 'correctly', *schön* 'pretty' etc.

(96)   a.   daß es Hans ja wohl richtig   gemacht hat
                 that it Hans indeed   correctly made   has
                 'that Hans has indeed done it correctly'
   b. *   daß es Hans richtig ja wohl gemacht hat

Furthermore, Haiden (1995:55) shows that a "maximally non-specific" argument such as non-interrogative *was* may only appear to the left of manner adverbs.[76]

(97)   a.   daß Hans endlich einmal was gut/richtig   machen will
                 that Hans finally   once   what well/correctly make   wants
                 'that Hans finally wants to do something well/right'
   b. ?? daß Hans endlich einmal richtig was machen will
   c. *   daß Hans endlich einmal gut was machen will

This shows that the claim that any position to the left of manner adverbs renders arguments specific, cannot be maintained; non-interrogative *was* may never be specific showing that (i) scrambling out of VP is not only restricted to specificity and (ii) existential interpretation is not dependent on a projectional domain but on the domain of the existential operator.[77]

Another consequence of Haiden's study is that manner adverbs do indeed sit low in the clause. Building on this finding, one would predict—under current assumptions regarding Multiple Wh-constructions—that double questions cannot be formed with either Wh-element following a manner adverb.

(98)   a.   Wer hat was richtig gemacht?
      b.   Was hat wer richtig gemacht?
(99)   a. *   Wer hat richtig was gemacht?
      b. *   Was hat richtig wer gemacht?

Recall that the desired interpretation of Multiple Wh is a pair-list reading. While the questions in (99) may not bear this reading, they may serve as Echo questions. Postma (1995:106) makes a similar observation. He employs the Dutch equivalents (*netjes* 'neatly') to show the impact on Wh-elements with respect to a pair-list interpretation based on relative distribution. His data and conclusions carry over directly to German:

(100) a. ?? Welcher Schüler hat richtig welche Aufgabe(n) lösen können?
           Which   pupil   has right   which exercise(s) solve can
           'Which pupil could solve which exercise(s) correctly?'
     b.   Welcher Schüler hat welche Aufgabe(n) richtig lösen können?

---

[75]   See fn. 65 above: this section intends to show that alternative approach has problems accounting for the data accurately, too. *Ceteris paribus* the need to conclusively and unambiguously show that both Wh-elements are in the high TopP assumed here is not as crucial at the moment: it probably is as impossible to show as well-known instances of vacuous movement (e.g., movement of a Wh-subject to CP in English)—if on the right track—and thus follows from theory-internal reasons (Checking Theory à la Chomsky 1995) and the range of data this approach covers.

[76]   Why non-interrogative *was* can only be non-specific, see Haiden 1995 or Postma 1995, among many others.

[77]   Diesing (1992) proposes VP as the invariant domain for existential interpretation, while Postma (1995) suggests to link it to an existential operator.

(101) a. * Welcher Schüler hat richtig was lösen können?
     b.   Welcher Schüler hat was richtig lösen können?

The Multiple Wh-constructions in the b-examples are both grammatical and felicitous: a pair-list reading can be construed. This is not the case in the a-examples; here, the only possible interpretation is that of an Echo question.

     The preceding discussion has shown that Wiltschko's data show Superiority effects even in German; it has also shown, however, that Wiltschko's analysis is not sufficient: she rules out the ungrammatical constructions on the grounds of the lower Wh's failure of being D-linked; D-linking occurs through scrambling. As other authors have shown, however, even the supposedly non-scrambled Wh-elements in Wiltschko's corpus have scrambled.

     Wiltschko's approach can be reconciled with the present analysis in a straightforward manner. As briefly touched upon already, DRQ might be seen as a movement-variant of D-linking. Thus, elements may be D-linked by not moving at all (inherent D-linking as laid out by Pesetsky) or by moving to a topic position. English does not make use of the movement-variant,[78] but German does so (solely). Hence, Wiltschko's D-linked phrases are actually topicalized, following DRQ. To conclusively show that a high TopP in the left periphery is involved (and at the same time account for Wiltschko's apparent counterexamples) shall be a topic for further research; for the time being, I conclude that Wiltschko's analysis cannot accurately account for her data either (see also fnn. 65 and 75 above).

     As a last word for the time being on this issue, note that (95b) may be rendered grammatical when a manner adverb follows the lower Wh:

(102) a.   Was hat denn schon oft wer gut gesehen?
     b.   Was hat denn schon oft wer richtig gemacht?

Take another set of examples from her study (Wiltschko 1997:117-8), concerning Pesetsky's notion of "aggressively non-D-linked" Wh-elements:

(103) a.   Wer zum   Teufel hat wen gesehen?
          who to-the devil has who seen
          'Who the devil saw who?'
     b. * Wen zum Teufel hat wer gesehen?
(104) a.   Who the hell caught what
     b. * Who caught what the hell?

Pesetsky (1987:124f., fn. 20) presents the contrast in English between an aggressively non-D-linked Wh-subject (104a) and an aggressively non-D-linked object (104b). Aggressive non-D-linking also exists in German (103), as Wiltschko convincingly shows, and patterns with English in that the Wh-phrase in question may not be scrambled.

     We can capture the ungrammaticality of another of Wiltschko's examples—namely, (103b) containing an aggressively non-D-linked Wh—differently: the Wh-phrase *wen zum Teufel* need not—and hence, under minimalist assumptions, cannot—be topicalized, as it is clearly non-D-linked. A non-D-linked Wh must sit in SpecCP in order to be interpreted and such movement evokes a completely different interpretation of the question.[79]

     The following sentences can be accounted for accordingly (Wiltschko 1997:118):

(105) a. * Wer zum   Teufel hat wen zum   Teufel gesehen?
          who to-the devil has who to-the devil   seen
          '*Who the devil saw who the devil?'
     b. * Wen zum Teufel hat wer zum Teufel gesehen?

---

[78]   Presumably only with partitive Wh-elements, though this is not crucial.
[79]   Recall that the German Multiple Wh-constructions investigated here are those that follow DRQ. The case of aggressively non-D-linked Wh-elements might blur the object of study: it may in fact not be a "real" multiple interrogative, in the sense of Uriagereka 1988, for example. In this sense, it is plausible that no topics are involved and consequently (103a) conforms to the Superiority Condition, while (103b) does not: according to the MLC, the first Wh that C^0 attracts is *wer zum Teufel* and *wer*, respectively.

Aggressively non-D-linked Wh-elements must move to SpecCP overtly; German C^0 can only attract one Wh prior to Spell-Out, hence ruling out the constructions in (105).

3.3.3 Wh-Adjuncts in Multiple Wh-Constructions: Some Speculations

Take for one Huang's (1982) null preposition hypothesis that aims at accounting for the syntactic behaviour of Wh-adverbs such as *where* and *when* vs. *why* and *how*.[80] Only the former may be preceded by an overt preposition such as *from where, since when*. We can schematize this as follows with the null preposition represented as *p* (from Kitahara 1997:101):

(106) a.   [PP *p* where/when]
　　　 b. * [PP *p* why/how]

Huang uses this fact to propose that *where*/when pattern like *why*/how in the syntax, while they behave like *who*/what at LF. (See also Tsai 1994 for a discussion of this in Chinese.) As Norbert Hornstein (personal communication) points out, it is not immediately clear how this might account for their asymmetry with respect to Superiority effects (but see Kitahara 1997:101ff. for some suggestions).

　　　This would also predict an asymmetry for the German equivalents. But the asymmetry is not as uniform in German. The reason may lie in the second part of Huang's analysis which concerns non-interrogative pronominal counterparts. While *where* and *when* have a pronominal counterpart in Modern English (*there* and *then*), *thy and *thow do not exist for *why* and *how*. German *warum* and *wie*, on the other hand, do have such forms: *darum* for *warum* (and also *deshalb* for *weshalb*) and *so* for *wie*.[81]

　　　We have already seen that German and English differ with respect to licensing conditions for Multiple Wh. The pronominal evidence from Wh-adjuncts seems to suggest further that the current approach is on the right track: DRQ is not active in English in the sense proposed for German, a conclusion that is evident from the above discussion. There is thus a restriction in English that only one Wh-element moves overtly, and the target site is SpecCP. In German, both elements move, and they both move to TopP. The richness of morphology in German vs. the absence of it in English might be one of the reasons in that nominal features in German need to be checked off overtly. Moreover, *why* and *how* can only occur in the sentence-initial operator position (SpecCP), presumably due to their property of acting as a "generator" (see Chierchia 1991, Hornstein 1995; for functional approach, see Reinhart 1995). They do not seem to have these properties in German (but see Haider 1996). Consequently, the German Wh-adjuncts are all subject to the standard movement restrictions (MLC); English *why*/how vs. *where*/when do not act uniformly due to the differences in the nature of movement. This might be a clue that *why* and *how* are generated in a specific position, rather moving there.[82,83]

　　　Let us capture this as the following hypothesis:

(107) *Differences in English* why/how *vs. German* warum/wie
　　　 a.　　 *why, how* are base-generated in CP and cannot move anywhere
　　　 b.　　 *warum, wie* are not base-generated in CP but may move there

---

[80]　Jianxin Wu (personal communication) informs me that this distinction was also intended to capture the Chinese data, which has subsequently been modified. The exact behaviour of *why*/how remains a mystery.
[81]　Werner Abraham (personal communication) reassures me that *so* is the pronominal counterpart for *wie*. Why this analogy does not extend to English (*so* for *how*) is not clear to me; neither am I competent to speculate about the historical roots of these elements in order to support the current suggestion.
[82]　Reinhart (1995 and earlier work) argues for this and identifies the position uniformly as SpecCP.
[83]　Rizzi (1997) proposes the existence of TopP preceding FocP ("CP") as well as following it. The discussion so far has only been concerned with the ordering CP — TopP; but if there is anything to the need of *why*/how to sit in SpecCP (and their counterparts across languages), we might want to consider TopP — FocP/CP — TopP above IP. One possible consequences for the German cases where *warum*/wie seem to be "*in situ*" would be that the adjunct sits in the specifier of FocP/CP, while the other Wh moves to either the lower TopP or the higher one. This is sheer speculation at the moment but it might turn out to be interesting to pursue this further.

(On a more general note, it has been assumed so far that adjuncts are base-generated and move, in analogy to arguments. This need not be the case, as briefly mentioned already. While one view of *why* and *how*, for example, is that they are base-generated in SpecCP (see Reinhart 1995 and much earlier work), this might be extended to all Wh-adjuncts; if this can be done, we might actually want to explore the possibility of late insertion for all (adverbial) adjuncts. It could thus be an option that adverbs are introduced into the derivation after all obligatory (= overt) movement has taken place. Under this conception, there is no multiple movement to TopP in those cases where adverbs are positioned in between two Wh-phrases. Rather these adverbs are generated in those adjoined positions instead of being generated in adjoined positions inside IP.[84] Generalized, this would mean that in Multiple Wh-constructions containing a Wh-adverb, this element is inserted into the position that it fills at Spell-Out.)

On a different note, the following sets of data (from Kuno and Takami 1993:90ff.) show some differences in the position of *why/how* vs. *where/when*. With support from German, I will argue that these data do not necessarily show different positions of these elements, but differences in possible landing sites for some types of topics in the two languages, drawing from Grohmann 1997a and extended data.

(108) a.   Why, man, did you come to the United States?
        b.   How, man, can you drink hot coffee so quickly?
(109) a. * When, man, did you come to the United States?
        b. * Where, man, did you meet Mary?

Kuno and Takami argue that these data show that *why/how* sit in a different position from *where/when*; in their framework, they equate the position of the former as dominated by S'' (like left-dislocated elements) and that of the latter as dominated by S' (like topicalized elements). I will turn to the impact of topicalization and left dislocation presently.

Interestingly, the German equivalents to (107-108) are fine in German:[85]

(110) a.   Warum, Mann, bist du in die Vereinigten Staaten gekommen?
        b.   Wie, Mann, kannst du heißen Kaffee so schnell trinken?
(111) a.   Wann, Mann, bist du in the Vereinigten Staaten gekommen?
        b.   Wo, Mann, hast du Maria getroffen?

Kuno and Takami assume that left-dislocated elements and topics sit in different positions. Using the finer terminology from Grohmann 1997a (see Anagnostopoulou et al. 1997, especially the introduction (van Riemsdijk 1997) for more), Kuno and Takami's left-dislocated elements are, in fact, hanging topics. Put crudely, hanging topics differ from left-dislocated elements in not agreeing in Case with the resumptive pronoun. I argue that topics and left-dislocated elements move to TopP, while hanging topics are base-generated in TopP.[86]

The following paradigm and the corresponding simplified representations show this (where indices indiacte derivational and italics semantic co-reference):

(112) a.   *Diesen Mann$_i$ hat Maria t_i geküßt*
        this-ACC man has Mary   kissed
        'This man, Mary kissed'
        b.   *Diesen Mann$_i$, den$_i$ hat Maria t_i geküßt*
        this-ACC man   the-ACC has Mary   kissed
        'This man, Mary kissed him'

---

[84]   This option—if desirable in the first place—will have to address various constraints on derivations, such as the Extension Condition, the (Strict) Cycle Condition and others. I will leave this open for further investigation.
[85]   The interjection *Alter* ('Old' as in 'buddy, mate, dude') is colloquially more common than *Mann* in today's language. Although (110-111) may not be used as such, other interjections such as *mein Freund* 'my friend' sound even better and are used as shown here.
[86]   Note that this notation differs slightly from the one traditionally used for English. Lasnik and Saito (1992), for example, maintain that English exhibits left dislocation; this type, however, is base-generated there and hence correlates directly to hanging topics as used here. For more, see Grohmann 1997a.

c.   *Dieser   Mann, den$_i$*   hat Maria *t$_i$* geküßt
      this-NOM man   the-ACC has Mary   kissed
      'This man, Mary kissed him'

In (112a), the topic moves from its AgrOP-position to TopP where it checks [Top]; all other features have been checked at AgrOP. The left-dislocated element is also a topic (i.e. endowed with [Top]); however, it moves from its Agr-position to an intermediate position where it spells out as the demonstrative pronoun (presumably for V2-reasons) before it moves on to TopP. In the hanging topic-construction, on the other hand, the demonstrative pronoun is the base-generated object that moves for the usual reasons and the hanging topic is generated in TopP; crucially, this element does not agree in Case with the pronoun. In English, left dislocation does not exist, but the comparable constructions involve hanging topics.

Furthermore, this TopP is available in both languages preceding CP. Only in German, though, is this position a possible landing site for movement; in English it may only serve as a position for base-generated elements (i.e. hanging topics[87]).[88]

(113) a.   This man, why did Mary kiss him?
      b. * This man, why did Mary kiss?
(114) a.   Diesen Mann, warum hat Maria den geküßt?
      b.   Dieser Mann, warum hat Maria den geküßt?

These facts can account for the differences that Kuno and Takami note without positing different positions for Wh-elements in the following way. Interjections sit in such a position that they intervene movement over them in English but not in German. Thus, if an element occurs in a position preceding the interjection in English it must be base-generated. The analogy is drawn to the distinction of base-generated hanging topics and movement-based left dislocation. These show also that a left-peripheral position blocks movement.[89]

Returning to the present topic, we now have evidence that left-peripheral positions vary in German and English with the possibility of movement. The syntax of interjection has, to my knowledge, not been the topic of much recent research. We can see from the limited data that it must be in a fixed position in the CP-domain. Interjections presumably sit right below the Wh-position (or even adjoined to C^0) and somehow intervene with movement to that position in a similar vein that we observe in hanging topics. The conjecture is that the hypothesis (107) can be retained and these restrictions account for the differences observed by Kuno and Takami.[90]

3.3.4 Multiple Wh and Quantifier Interaction: A Reaction to Beck (1996)
Wu (1996) shows that scope asymmetries exist in Chinese when Wh-elements and quantifiers interact (see below for more on Wu's study). In particular, a topicalized Wh-element yields only an individual reading when fronted over a quantifier. Interestingly, Beck

---

[87]   Note that Case-agreement is different from German: the fact that the hanging topic may also bear accusative Case in English is rather a reflex of morphology (with accusative being the default Case). Every dislocated element is marked accusative, just like all hanging topics in German are marked nominative (default Case in this language). See Grohmann 1997a for thorough discussion.

[88]   With the Wh in CP, V2 has not to be saturated in any special form in (114b). I take the low pronoun to be a Spell-Out for Case-reasons. Note that topics are not allowed without the resumptive pronoun:

(i) * Diesen Mann, warum hat Maria geküßt?

I have to refer the interested reader to Grohmann 1997a for more discussion and evidence.

[89]   Another argument in favour of base-generation of *why* and *how* comes from the observation that in conjunction with interjections, they may only modify the clause they are in (pointed out by Norbert Hornstein):

(i) a. * *Why*, man, did you say Bill *left*
     b.   *Why*, man, did you *say* Bill left

[90]   It is clear that the previous discussion is just another tentative approach. The details to this approach would be beyond the scope of this paper and I have to leave them open for further considerations, as so many issues.

(1996) also notes a similar peculiarity for German. She presents data that show interpretive differences in the interaction between quantifiers and Wh-phrases, and proposes an analysis based on the Minimal Quantified Structure Constraint.[91]

Consider (115), taken from Beck 1996:1 and 6:

(115) a. # Wen   hat niemand wo   gesehen?
         whom has nobody where seen
         'Where did nobody see whom?'
      b. # Wann hat niemand wen   eingeladen?
         when has nobody whom invited
         'When did nobody invite whom?'

The two Wh-elements are argument and adjunct in both cases. In (115a), the Wh-adjunct *wo* occurs low, while *wann* appears high in (115b); both positions are fine under normal circumstances (cf. (28)). What makes these cases special is the presence of the quantifier *niemand* in between. It seems to render the sentences ungrammatical, or rather, deviant (marked by '#'): no coherent meaning can be assigned to the sentences in (115), as Beck remarks.

Compare with (116):

(116) a.   Wen hat wo niemand gesehen?
      b.   Wann hat wen niemand eingeladen?

Here, the quantifier appears lower than both Wh-elements. Both sentences are grammatical and can be interpreted distributively.

Under the proposal laid out here, the difference between (115) and (116) can be accounted for directly: as we have seen in section 3.3.1, a non-Wh-subject must move overtly to TopP in case it is preceded and followed by a Wh-element as it could otherwise not be interpreted as the highly thematic subject. This implies for (115), that the quantifier *niemand* would be topicalized. We know independently, however, that certain quantified expressions cannot be topicalized.

This is shown in the following.:[92]

(117) a. * Nobody, Karl likes
      b.   Most people, Karl likes
(118) a. * Niemanden mag Karl
      b.   Die meisten Menschen mag Karl

The lack of topic-properties of *niemand* in (115) rules out the derivation that it would receive under the present proposal. The subject-quantifier cannot move to TopP, and thus *wo* cannot move to TopP either, violating DRQ. On the other hand, if the quantifier stays in AgrSP, both Wh-elements can freely move to TopP, resulting in a grammatical output (116).

This rationale rules out (119a), while allowing (119b) (from Beck 1996:6):

(119) a. # Wer hat niemanden wo   angetroffen?
         who has nobody   where met
         'Who didn't meet anybody where?'
      b.   Wer hat wo niemanden angetroffen?

---

[91] The following gives a straightforward account for Beck's data throughout. While lending strong support to the present analysis, it also does away with the various constraints that Beck proposes. A more careful investigation of her data may be needed, but the most relevant parts will be dealt with here rather smoothly.

[92] I cannot indulge here in a deeper discussion of the exact reason for this restriction. Only some quantifiers can be topicalized, and for the present purpose it suffices to note that some quantifiers such as *nobody* cannot topicalized. One might classify the topicalizable vs. non-topicalizable quantifiers along the weak/strong-dichotomy of quantifiers following Milsark (1977) and much subsequent work; however, as pointed out to me by Paul Portner, the correct dichotomy seems to be increasing vs. decreasing quantifiers. (Gregory Ward (personal communication) suggests referentiality of quantifiers to be relevant; under this view, referential quantifiers may be topicalized, non-referential ones may not. For a number of reasons, the increasing/decreasing distinction seems to be more relevant, leading to a discussion beyond the scope of this paper.)

Quantifiers such as *niemand* in general cannot be topicalized, whether subject or object. Following from the syntax of DRQ, *wo* in (119a) needs to move to TopP which in turn implies that the quantifier topicalizes, too, given the not unreasonable view that explicit topic phrases are situated in the left periphery. The quantifier stays in its AgrOP-position at Spell-Out in (115b), allowing the second Wh-element to move to TopP, yielding grammaticality.

Needless to say, constructions involving two Wh-adjuncts underlie the same restrictions and conditions (from Beck 1996:16):

(120) a. # Wann hat Karl niemanden wo   angetroffen?
            when has Karl nobody   where found
            'When did Karl where find nobody (at home)?'
      b.   Wann hat Karl wo niemanden angetroffen?

While the non-Wh-subject may freely topicalize, the quantified object may not (120a), rendering the sentence ungrammatical; if the quantifier stays in its Agr-position, the by now familiar configuration arises where both Wh-elements as well as the subject topicalize.

On the other hand, comparable constructions to (115) with an increasing quantifier (cf. fn. 92) in between the two Wh-elements are predicted to be fine: such a quantified expression should be able to topicalize.[93]

This prediction is borne out, lending further support to the present approach:

(121) a.   Wen   hat jeder   wo   gesehen?
            whom has everyone where seen
            'Whom did everyone see where?'
      b.   Wen   haben die meisten (Menschen) wann eingeladen?
            whom have   the   most   (people)   when invited
            'Whom did most (people) invite when?'

The structure assigned to (121) should by now be very familiar: the first Wh sits in CP (after having moved through TopP), while the lower Wh sits in TopP. The subject needs to move to TopP, too, in order to be interpreted as the highly thematic element that it is;[94] in these cases, the subject is an increasing quantifier which can legitimately move to TopP.

Beck (1996:4) also has another interesting set of examples. These are Wh-phrases that are split; crucially, these phrases are clear partitives in the sense assumed here throughout.

(122) a. # Wen   hat keine Studentin von den Musikern getroffen?
            whom has no   student   of   the musicians met
            'Which of the musicians did no student meet?'
      b.   Wen hat Luise von den Musikern getroffen?

The Wh-phrase in question is *wen von den Musikern*—obviously a partitive Wh. While (122) is not a Multiple Wh, it contains this partitive Wh. This, as I have argued for in considerable detail, initiates the movement to TopP; [Top] is a feature inherent to such elements. Interestingly, German allows splitting up such Wh-phrases. The derivation for (122) is thus parallel to those we have seen so far: the Wh moves completely to TopP, from where the interrogative part moves on its own overtly to CP to check the Wh-feature. As predicted, the thematic interpretation of the subject is blurred by the presence of a filled TopP and the subject thus has to move to TopP as well. As before, a decreasing quantified subject cannot do so (122a), unlike a non-quantified one (122b).

A very similar case is the following (Beck 1996:7):

---

[93] Note that this would only be predicted if (i) both Wh-elements in Multiple Wh move to TopP and (ii) the subject raises to TopP also. This sub-section shows how smoothly the present proposal works for these data which otherwise would be subject to rather elaborate constraints. This in and of itself supports this approach.
[94] Beck has more to say on the status of *jeder* which is not relevant to the present discussion; in fact, it is not exactly clear how relevant it is at all.

(123) a.   Was für Bücher hat niemand gelesen?
            what for books has nobody read
            'What kind of books did nobody read?'
      b. * Was hat niemand für Bücher gelesen?

Again, a partitive Wh moves to CP; for theory-internal reason, it has to stop at TopP on the way. This is the sort of vacuous movement that is hard to show but assumed throughout. We can only give the circumstantial evidence from (123b) where the partitive element of the Wh remains at this intermediate position, on analogy with the previous examples. In (123a), the subject does not have to raise to TopP, as it is not overtly filled; *niemand* may receive interpretation in AgrSP just fine. In (123b) TopP above the subject is filled, forcing illegitimate movement of *niemand* to TopP.

   As predicted, (122) is grammatical with an increasing quantifier as the subject, and so is (123).[95]

(124) a.   Wen hat jede Studentin von den Musikern getroffen?
      b.   Was hat jeder für Bücher gelesen?

The increasing quantifiers in (124) may move to TopP without any further ado.

   The approach as just outlined can thus account for all of Beck's data without further assumptions. The role of quantifiers and topicalization may need some more attention, and so does Beck's discussion on apparent intervention effects of *jeder* 'every'. The upshot of this discussion, however, is that a topicalization-approach to Multiple Wh (driven by DRQ) seems feasible not only on its own grounds, but also with respect to other relevant phenomena observed in the literature. Furthermore, this section lends strong support to the proposal that under certain circumstances subjects must topicalize overtly.

4   Cross-Linguistic Variation, Learnability and More Support

In the last section, I will present evidence from other languages; the most prominent one—and the one which this study leans on heavily—is Wu 1996 for Chinese. I will then discuss the role of Wh-islands to some extent. After looking at cross-linguistic evidence for Wh-topics, it will be more generally considered how the D-linking/DRQ-dichotomy from English and German can be realized across languages. I will close with some remarks on learnability and child language.

4.1   The Availability of Wh-Topics in UG

This section will contain an elaborate discussion of the Chinese findings by Wu (1996) that allow us to construe the notion of a Wh-topic; I will also give additional evidence from Japanese.

4.1.1 Chinese Wh-Topics
In Chinese questions there is no overt movement of the Wh-element(s). Presumably, the Wh-elements move to SpecCP at LF in both single and double questions.[96] This is considered the standard approach to Chinese interrogatives ever since Huang 1982.

---

[95]   For the sake of completeness, the version of (123b) without a quantified subject is also grammatical:
(i)   Was hat Luise für Bücher gelesen?

   In fact, a thorough consideration of all examples shown here really do exhibit the same pattern: a non-quantificational subject or an increasing quantifier may occur in between two Wh-elements freely; decreasing quantifiers may not. The same goes for all instances of split partitives and for the cases of *was für*-split. Other cases of splitting up phrases (as studied in a number of papers in Lutz and Pafel 1996) need to be considered in more detail in future research; as a direct reply to Beck's discussion, this section should suffice.

[96]   Aoun and Li (1993) and Watanabe (1992) suggest movement of parts of the Wh-phrase (or an empty Wh-operator) in the overt syntax, parameterizing Wh-movement in a different way. I will adopt the traditional view, although nothing much hinges on this assumption.

I will now present some interesting findings by Wu (1996) who concludes that in some cases Wh-elements may move. These cases are strictly discourse-restricted in that they are instances of topicalization. After citing his empirical evidence, I will recapitulate Wu's speculations on a possible interaction between Wh-movement and topicalization—two grammatical processes that seem at first glance to contradict each other but may well be conceived of as being related. The general idea is very much in line with the proposal.

First, consider standard examples of Wh-formation in Chinese where no movement takes place at all, as presented by Wu (1996:178):

(125) a.   Shei mai-le   shenme?
           who buy-ASP what
           'Who bought what?'
    b.   Ni   quan   shei bu yao mai shenme dongxi?
           you persuade who not   buy what   thing
           'You persuade who not to buy what?'

In (125a) we have an instance of a subject- and Wh-object, while (125b) parallels the construction where Pure Superiority could possibly arise (by movement of the lower Wh-element). In neither construction can we observe movement taking place.

But compare (125) to (126) below (from Wu 1996:177):

(126) a.   Shenme$_i$ shei mai le   t$_i$?
           what   who buy ASP
           '*What did who buy?'
    b.   [Shenme dongxi]$_i$ ni   quan   shei bu yao mai t$_i$?
           what   thing   you persuade who not   buy
           '*What did you persuade who not to buy?'

We can see that the constructions in (125) are the equivalents of (126) with the exception that the lower Wh-element is fronted in both.[97] This gives rise to Superiority violations in English but not in Chinese.

In an interesting discussion, Wu reaches the conclusion that the fronted Wh-elements in (126)—and comparable cases—must be topicalized. For economical reasons within the minimalist framework, the movement of the Wh cannot be for Wh-feature-checking purposes: the standard Wh-checking at LF is more economical than overt checking.

One of the most convincing pieces of evidence he presents is the following restriction on Multiple Wh (adapted from Wu 1996:181):

(127) Situation: Zhangsan went to the supermarket to do some shopping.
    a.   Zhangsan mai-le   shenme?
           Zhangsan buy ASP what
           'What did Zhangsan buy?'
    b. #   Shenme Zhangsan mai-le?

---

[97]   Wu notes that they are not equivalent in their meaning: while (126a) may be interpreted in a distributive or an individual reading, (126b) may only be assigned the latter. But this is a difference not directly relevant here: this section serves to show that the conception of certain Wh-elements as topics is not as alien as it may sound at first sight; in how much languages vary with respect to the semantics of Wh-topics needs to be established in further research.

    Wu (1996:182) also points out that (ib) may only have a distributive meaning (unlike (ia)), a statement that he qualifies somehow in his fn. 2 on the same page according to which some speakers may get a distributive reading.

(i) a.   What did everyone buy?
   b.   Which book did everyone buy?

    In order to make clear predictions about the possibilities of distributivity and topicalization of Wh-elements (and quantifiers), more work needs to be done.

In the given context, only (127a) is an appropriate question. The situation does not imply that Zhangsan actually bought anything. The question in (127b), on the other hand, can only be posed if there is a referent to the Wh-element, i.e. if Zhangsan has really bought something such as a number of items of a shopping list. (Compare this with the felicity of questions for German.)

(128) Situation: Z. went to the supermarket and bought items from a mutually known list.
       Shenme Zhangsan mai-le?

At the face of it, German and Chinese are different in that DRQed Wh-phrases always topicalize overtly in German, while in Chinese it is only one Wh that moves (also, see fn. 97 on interpretation). This might again be related to the general absence of morphological features in Chinese driving overt movement.[98]

While the discourse restrictions show similarities between Chinese Wh-topics and German Wh-topics, weak crossover effects give us a further unambiguous clue that the fronted Wh is indeed in TopP and not in a Wh-marked CP.
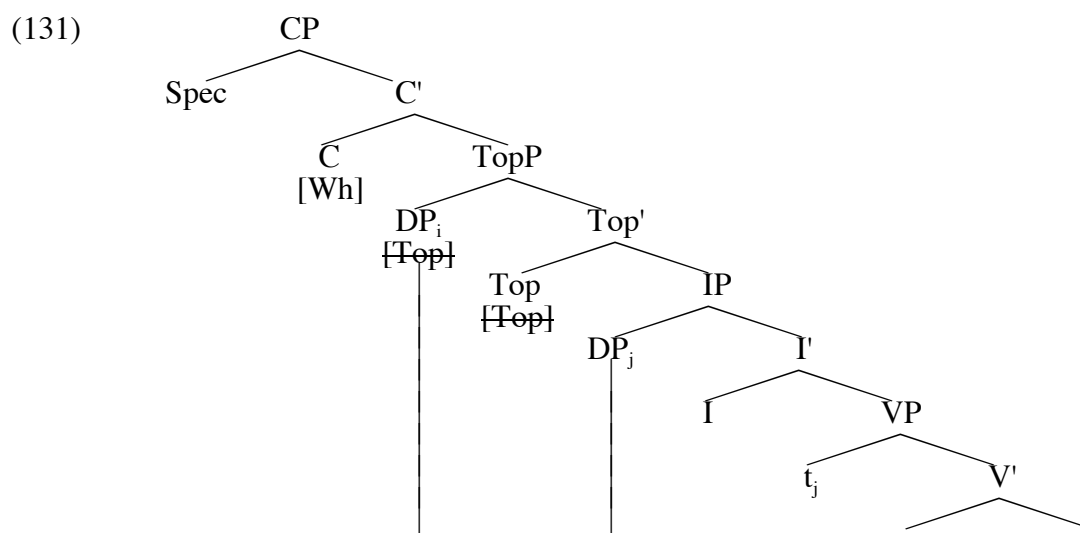
Consider the following contrast in English (Wu 1996:178):

(129) a. *   Who$_i$ does his$_i$ mother like t$_i$?
      b.   John$_i$, his$_i$ mother likes t$_i$

A weak crossover effect occurs in English when the Wh is co-referent with a pronoun; but in topicalized structures the pronoun may be a bound variable co-referent with the topic.

(130) a.   Shei$_i$ ta$_i$ de muqing hen xihuan t$_i$?
              who he DE mother very like
              '*Who does his mother like?'
      b. *   Ta$_i$ de muqing hen xihuan shei$_i$?

Wu (1996:179) shows that the same contrast can be observed in Chinese: if the Wh stays *in situ*, weak crossover effects can be found (130b), which disappear if the Wh moves to TopP.

Wu assumes—as I have throughout the paper—that topics move to their own projection, namely TopP, outside the IP-domain. (126b) and (128) thus receives the following structural representation:[99]

(131)

					V	t$_i$
a.	*Shenme*		*she*		*mai-le*	
b.	*Shenme*		*Zhangsan*		*mai-le*	

Overt movement to TopP satisfies [Top], subsequent movement to CP will occur at LF. Note the crucial difference: only one TopP is available, not forcing the subject to topicalize.

However, Wu (personal communication) also informs me that the relation between Multiple Wh and context as described in section 3.1.2 for German holds the same for Chinese (as shown in section 4.3.1 below) and thus differs crucially from comparable English constructions.[100] The final piece of evidence from Chinese will be presented in section 4.2 where the issue of Wh-extraction out of weak islands will be discussed. The confirmation from Chinese, and the Japanese context-restriction shown below, indicate that the present approach hits upon something novel that is worth pursuing.

4.1.2 Japanese Wh and the Topic Marker *wa*

As shown by Miyagawa (1987), the Japanese topic marker *wa* may under certain circumstances also appear in interrogative contexts. This has not been widely noticed, for exactly the reason that these circumstances are very restricted. But it shows that Wh-questions and topicalization are not mutually exclusive by nature.[101]

Miyagawa (1987:186) notes that while *wa* is ungrammatical in simple interrogatives, it may occur in those questions that severely restrict the set of referents. Here, *wa* is interpreted thematically and thus refers to a definite individual:

(132) a. * Dare-wa kita no?
        who-TOP came Q
        'Who came?'
    b.   Dare-wa kite,   dare-wa konakatta   no?
        who-TOP come-GER who-TOP didn't-come Q
        'Who came, and who didn't?'

Miyagawa shows more instances of the connection of the topic marker and Wh-elements. Note that this is also reminiscent of what Wu (1996) has shown for Chinese and what section 3 was concerned with for German. The upshot is, again, that a connection between Wh and topicalization need not be ruled out a priori; furthermore, it is found cross-linguistically.

It should be noted that the nature of *wa* on Wh-elements differs from the remainder of this paper. While the contrast between German and English showed different constraints by the discourse (further supported by the Chinese data), something else is going on in Japanese; *wa* is a contrastive topic-marker and thus induces a different interpretation than a D-linked topic.

4.1.3 Wh-Topics in General

The limitations set for this project have not allowed a more thorough survey of Wh-topics across languages; this task remains for future research. Interestingly, evidence from child language acquisition might support the present approach (deVilliers et al. 1996, Thornton 1995) which will be laid out in section 4.4; incidentally, non-verbal languages might do so, too.[102] The outcome of this will also have to be addressed in another project. In sum, the cross-linguistic observations in general, and the thorough application to German in particular, go hand in hand towards a theory of Wh-topics that might turn out to be found more frequently than previously assumed.

---

[100] It exceeds the scope of this paper to thoroughly look at more examples from Chinese in order to say that all cases of Multiple Wh in Chinese underlie the same discourse restrictions in German.

[101] Judgements regarding the acceptability of *wa* vary; it should also be noted that *wa* in these cases can only mean contrastive topic. This might be additional evidence that a focus-part is involved in these constructions, maybe even responsible for Wh-interpretation across languages (as briefly mentioned above).

[102] In a recent study, Neidle et al. (1997) are concerned with Wh-formation in American Sign Language (ASL). Their findings seem to suggest that ASL employs a sentence-initial Wh-topic in addition to a sentence-final "real" Wh-element. Robert Lee (personal communication) brought this research from ASL to my attention.

4.2   Wh-Islands and Wh-Topics

Comorovski (1996) connects Cinque's (1990) characterization concerning D-linking and the possibility for Wh-extraction out of a Wh-island (see her chapter 5 for detailed discussion) into her general account of D-linking.

4.2.1 Some Basic Data

Comorovski's (1996:162) data from Swedish are taken from Maling 1978 and Engdahl 1980 which are shown below, respectively:

(133) a. * Vad visste ingen vem som skrev?
            what knows no-one who   wrote
            'What does no one know who wrote?'
      b.   Sven undrar   vilken bok alla studenter minns   vilken författare som skrev
             Sven wonders which book all students   remember which author   wrote
             'Sven wonder which book all students remember which author wrote'

A bare Wh may not be extracted out of a Wh-island (133a), but a D-linked one may (133b).

Rizzi (1978) noted a similar property of Italian, where Multiple Wh-constructions are generally ruled out. Again taken from Comorovski 1996:163, the relevant data are shown here:

(134) a. ?? A chi   non ti ricordi   quanti   soldi hai   dato?
            to whom not you remember how-much money have-you given
            'To whom don't you remember how much money you gave?'
      b.   A quale dei tuoi figli   non ti ricordi   quanti   soldi hai   dato?
             to which of your sons not you remember how-much money have-you given
             'To which one of your sons don't you remember how much money you gave?'

Original speculations regarding the "heaviness" of Wh-island-extractable Wh-phrases seem to point to the phenomenon of D-linking, as fairly well established by Cinque and Comorovski.[103] With respect to the present approach—where multiple Wh-phrases are analysed as topics—this might have far-reaching consequences.

With this in mind, does Chinese—where Wu (1996) has conclusively shown that Wh-topics exist—allow Wh-extraction out of Wh-islands?

(134) a.   Shenme dongxi$_i$ Zhangsan xiang zhidao Lisi mai mei mai t$_i$?
             what   thing Zhangsan want know Lisi buy-not-buy
             'What did Zhangsan want to know whether Lisi bought?'
      b. * Shenme$_i$ Zhangsan xiangxin Lisi tou le t$_i$ de yaochuan?
             what Zhangsan believe Lisi steal DE rumour
             'What is x such that Zhangsan believes the rumour that Lisi stole x?'

As Wu (1996:74) shows a Wh-topic may be extracted out of a Wh-island (134a); (134b) illustrates with a CNP-island that strong islands are still boundaries for extraction.[104]

One can imagine how the by now familiar connection between D-linked Wh-phrases and Wh-topics would account for this within a general explanation why only these may be extracted and not bare, non-D-linked Wh-phrases; namely, the D-linked Wh uses embedded TopP as an "escape hatch" to move out of the Wh-islands. (I will lay out the details below.)

4.2.2 Slavic and the Position of TopP

---

[103] Cinque refers to the difference of D-linked Wh vs. non-D-linked Wh as a difference in "referentiality." In the terms of partitives, this referential property follows immediately (see also 4.4.3 below where a further explicit reference to referentiality made by Thornton (1995) is reviewed).
[104] As Comorovski (1996) sets out to account for, extraction of a D-linked Wh is possible not only out of Wh-islands but (often) out of all weak islands, as Wu also shows for Chinese (negative islands, factive islands). On the other hand, strong islands can never be moved out by Wh-elements.

Bulgarian also has a TopP in the left periphery, independently from Multiple Wh. The standard assumption is that projection dominates CP (or FocP in Rizzi's (1997) terms). This has been argued by Rudin (1985) and is mentioned by Richards (1997:111f.) who cites (135):

(135) a.   [$_{TopP}$ Ivan vcera [$_{CP}$ kakvo kupi]]?
             Ivan yesterday what   bought
             'What did Ivan buy yesterday?'
       b.   [$_{TopP}$ Vcera Ivan [$_{CP}$ kakvo kupi]]?

With respect to Wh-elements moving to this projection, Richards (1997:112) notes that only a D-linked Wh may move there:

(136) a. * [$_{TopP}$ Koj vcera [$_{CP}$ kakvo kupi]]?
             who yesterday what   bought
             'Who bought what yesterday?'
       b.   [$_{TopP}$ Koja zena   vcera [$_{CP}$ koja   kniga kupi]]?
             which woman yesterday which book bought
             'Which woman bought which book yesterday?'

While Wh-extraction out of islands is generally not possible (Richards shows this for strong islands), it is possible out of weak islands, such as Wh-islands—under the condition that the extracted Wh is D-linked. This observation was explored in detail by Comorovski (1989 and earlier) who presents the following from Romanian (which patterns much like Bulgarian):[105,106]

(137) a.   Despre care   stii   cine i   a   povestit?
             about   which you-know who to-him has told
             'Which one do you know who told him about?'
       b. * Despre ce stii cine i a povestit?

In (137)—from Comorovski 1996:164—, the D-linked Wh *despre care* 'about which (one)' is extracted out of the embedded question, yielding a grammatical question (137a); if *care* is replaced by *ce* 'who' (= non-D-linked), as in (137b), the construction is ungrammatical.

Suppose that D-linked Wh-elements are topics and that if these Wh-phrases move overtly (as clearly the case in all examples shown here), they move to SpecTopP—as argued all along—then the contrast in grammaticality between extraction of a D-linked and a non-D-linked Wh follows straightforwardly: the embedded CP does not serve as an escape hatch for the moved Wh (especially not if it is filled), but nothing prevents movement to the embedded TopP and further movement to the matrix TopP.

We would thus yield the following structural generalizations (which will be modified below):

(138) a. * [$_{CP}$ Wh$_i$ [… [$_{IP}$ … [$_{CP}$ Wh … [$_{IP}$ … t$_i$…]]]]]
       b.   [$_{TopP}$ Wh$_i$ [… [$_{IP}$ … ([$_{TopP}$ t$_i$ …) [$_{CP}$ Wh … ([$_{TopP}$ t$_i$ …) [$_{IP}$ … t$_i$…]]]]]]

A Wh may be extracted out of a Wh-island if TopP may serve as an intermediate landing site (the brackets in (138b) indicate that TopP may either precede or follow CP). A filled CP blocks movement (or whatever the particular assumptions about the ungrammaticality of Wh-islands are). This is supported by the well-known facts that topics can freely move out of Wh-islands, and weak islands in general, as we will see below in more detail.

Without going into any details, this approach has as a number of consequences for the analyses of Richards (1997) and Pesetsky (1998). Richards' cleverly designed theory of

---

[105] The seminal work on Slavic Multiple Wh is Rudin 1988. Another thorough study on Slavic (Romanian) syntax can be found in Dobrovie-Sorin 1993 and the collection of papers in Cinque and Giusti 1995.

[106] An interesting recent study on Multiple Wh in Slavic is Boskovic 1998. For reasons of space, I cannot indulge in further discussion; let me simply remark that Boskovic follows a different track based on focus-movement. While this approach makes explicit use of movement driven by a strong feature on the moved element (rather than on the attracting head only via "pure" Attract as assumed here), one can envision focus-movement in terms of Attract too (Boskovic, personal communication).

features and movement allows for multiple CP-specifiers where every Wh-phrase is tucked in after the first moved Wh—following Closest Attract and Shortest Move, respectively. While it covers a wide range of data, it cannot capture the fact that Romanian (and Bulgarian) allow Wh-extraction out of Wh-islands; to be precise, his system does not differentiate between D-linked and non-D-linked Wh-phrases. While the collaboration of a number of movement conditions allows a derivation in which a Wh is extracted out of an island, it does not predict—nor account for—that only D-linked Wh-phrases are extractable. If the theory is enriched by the conception of certain Wh-elements as (overt) topics, we lose the necessity for multiple CP-specifiers and consequently the data presented in those works can be re-analysed.

4.2.3 English and German: Some Apparent Problems[107]
As observed by Kiss (1993), for example, extraction of a Wh out of a Wh-island is only possible if the Wh is D-linked. Kiss (1993:91) presents the following contrast:

(139) a. *   How much money$_i$ do you wonder who to give t$_i$ to?
     b. ?   How much of the money$_i$ do you wonder who to give t$_i$ to?
The D-linked Wh in (139b) may be fronted to the initial position of the higher clause. But note that this Wh is not any D-linked Wh: it is a partitive Wh. This is exactly the type of Wh that is assumed here to move to TopP prior to Wh-movement.

As expected, those speakers who clearly differentiate between "normal" D-linked Wh-phrases and partitive D-linked Wh-phrases, the following contrast arises:[108]

(140) a. *   Which book do you wonder who bought?
     b.   Which of these books do you wonder who bought?

This shows that only those Wh-elements that even in English topicalize can be extracted.

Let us now look at German, where Wh-extraction out of Wh-islands is not possible—not even for D-linked Wh-phrases (see Müller and Sternefeld 1993 and especially d'Avis 1996 for discussion and references).

(141) a. *   Was weißt du, wer gelesen hat?
        what know you who read   has
        'What do you know who read?'
     b. *   Welches Buch weißt du welcher Student gelesen hat?

At first sight, this might pose a problem: if the language exhibits Wh-topics elsewhere (namely in Multiple Wh, conforming to DRQ), this is rather unexpected, especially with the cross-linguistic data in mind we have seen above.

This is even more puzzling as non-interrogative topics may be extracted out of Wh-islands.

(142) a.   Bücher weiß ich, wer gelesen hat
        books know I   who read   has
        'Books, I know who read'
     b.   Dieses Buch weiß ich, wer gelesen hat

Topicalized plural nouns (142a) or DPs with a demonstrative (142b) can be extracted out of a Wh-island (see Müller and Sternefeld 1993, Culicover 1996 for more). Obviously, they must get somehow to the matrix-TopP. So what could prohibit DRQed Wh-elements to move?

What Italian, Romanian and Bulgarian have in common is that all languages allow TopP to precede CP.[109] In particular, they all make use of this position in non-interrogative contexts. English, on the other hand, has this position available but the only evidence we

---

[107] For an easier-to-swallow version of this entire section, please see Grohmann, to appear.
[108] Space does not allow for a more accurate discussion of when Wh-extraction is allowed in English.
[109] I have not looked at the possibility of the ordering [$_{TopP}$ Top^0 [$_{CP}$ C^0…]] in Swedish or in Chinese rather than only allowing [$_{CP}$ C^0 [$_{TopP}$ Top^0 …]]; both might be feasible, however, given the following discussion.

have seen comes from base-generated hanging topics. In German, this position can also be filled, as we have witnessed from the examples of left dislocation.

I suggest to look at these constraints more closely. One way of interpreting the differences between English and German pre-CP topic-positions is the following: pre-CP TopP cannot be filled through overt movement in English; on the other hand, in German movement can only target it if is subsequently filled. If this characterization is correct, it should be possible in principle to evoke the projection in English if the element moves further; likewise, it cannot be activated in German if the element does not stay. Can we motivate this somehow?

Recall from section 3.3.3 that the resumptive pronoun in left dislocation-constructions is assumed to be the Spell-Out of the trace of the left-dislocated element. In Grohmann 1997a, 1998a I suggest that the Spell-Out occurs in order to satisfy constraints on Case-expression (low pronoun) or on V2 (high pronoun). Especially the latter condition might explain why left dislocation can only be found in V2-contexts, namely in matrix clauses and embedded under so-called "bridge" verbs that optionally may select for an embedded V2-clause.

Now consider (143), which seems to be a very strange exception at first glance:

(143) a.   Welche Bücher$_i$ weißt du, die$_i$ welche Studenten gekauft haben?
            which books know you the which   students bought has
            'Which books do you know that which students bought them?'
      b. * Welche Bücher weißt du, welche Studenten gekauft haben?
      c. * Welche Bücher$_i$ weißt du, die$_i$ haben welche Studenten gekauft?
      d. * Diese Bücher$_i$ weiß ich, die$_i$ diese Studenten gekauft haben

The grammaticality of (143a) comes as a huge surprise, given what we know about left dislocation. Left dislocation concerns a topic and its co-referent demonstrative pronoun; there is no (obvious) topic here. However, there is a Wh dislocated which is co-referent with a demonstrative pronoun and this particular Wh has been analysed throughout as a Wh-topic. This contrasts with (143b), where there is no demonstrative pronoun related to the Wh-topic. Also, as just noted, V2-context is crucial for left dislocation; it is not present here. This contrasts with (143c) where the same construction is set up within a V2-context. Lastly, (144d) shows that a non-interrogative topic may not be left-dislocated, violating the standardly assumed constraints.

Within the present proposal, (143a) receives a straightforward analysis.[110]

(144)

---
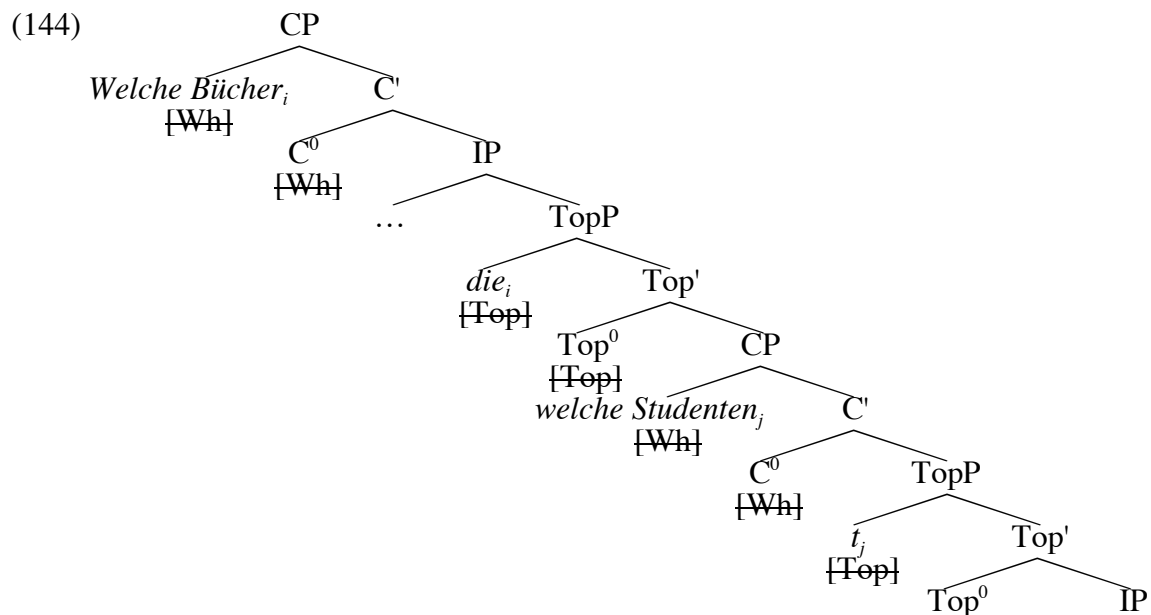
[110] I dispense with illustrating those parts of the derivation that are not relevant here. I will thus only show the derivational steps that are crucial to derive the left-dislocated Wh and the Wh-topic.

$$[\text{Top}] \qquad \overbrace{}$$
$$t_j \ldots t_i \ldots$$
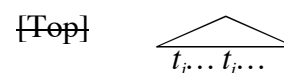
The overall constructions is a Multiple Wh; both Wh-elements thus bear [Top]. Both also bear [Wh] which in the normal case one Wh checks overtly, the other covertly. The derivation runs normally until the lower IP is formed. Top^0 enters the derivation and attracts the closest element, *welche Studenten*.[111] The Wh-subject thus checks [Top]. Next, C^0 attracts the closest element bearing a Wh-feature; again, the Wh-subject moves and checks [Wh]. It is only here that a further Top^0 comes in the picture. As *welche Studenten* has already checked [Top], *welche Bücher* is the closest element with an unchecked Top-feature; thus, the Wh-object is attracted, moves and checks its feature [Top]. After the matrix IP is formed, another C^0 arrives. There is only one element with a Wh-feature, namely the Wh-object. While in simple Multiple Wh-constructions there is only one C^0 and hence only one [Wh] can be checked overtly, in the case of two clauses a second C^0 is available, thus allowing the last move. Interestingly, however, the TopP dominating CP may not be empty. As in standard left-dislocated constructions certain requirements need to be met that force spelling out of a trace as a demonstrative pronoun, it is a further requirement in this case that forces to spell out the trace of the last move.[112]

We have now evidence that even German allows Wh-extraction out of Wh-islands.[113] To further test the proposal made here, we would expect that the bare Wh-equivalents of (143a) pattern just the same: the overall clause type is that of a Multiple Wh, thus any two Wh-elements should do (according to DRQ which is strictly at work in German).

This prediction turns out to be correct:

(145) a.   Was weißt du das wer gekauft hat?
           what know you the who bought has
           'What do you know that who bought it?'
      b.   Wen   weißt du   den wer gesehen hat?
           whom know you the who seen   has
           'Whom do you know that who saw him?'
(146) a.   Welches von diesen Büchern weißt du das welcher von diesen Studenten gekauft hat?

           which   of   these   books   know you the which   of   these   students   bought has
           'Which of these books do you know that which of these students bought it?'
      b.   Welchen von diesen Männern weißt du den welche von diesen Frauen gesehen hat?

           which   of   these   men   know you the which   of   these   women seen   has
           'Which of these men do you know that which of these women saw him?'
(147) a.   Was weißt du das welche Studenten gekauft haben?
      b.   Welche Männer weißt du die wer gesehen hat?

The extraction can apply to a bare Wh (145) or a partitive Wh (146); furthermore, any combination of bare, heavy or partitive Wh is fine (147).

With (145) at hand we can also test another prediction. The left-dislocated Wh cannot be base-generated, i.e. it cannot be a hanging topic.[114]

---

[111] Once the details concerning Wh-subjects are completely clarified, we can give a rule-driven account why in this case not the Wh-object is attracted; the apparent optionality we have already seen also holds for this example:

(i) Welche Studenten weißt du, die welche Bücher gekauft haben?

[112] The observation is that this position must be filled. I will not indulge in speculations how to formalize this requirement. Analogies might be brought up that in standard left dislocation Case or V2 need to be satisfied.

[113] While at first glance the examples presented seem to be limited in scope, it should be noted in my defense that we now have at least some instances of Wh-extraction out of Wh-islands. (Adjuncts might here actually be very different in general, see Boskovic 1998). These findings, however, point to an interesting research topic.

[114] Recall that the crucial difference is that left dislocation forces Case-agreement between topic and pronoun; this is not the case for hanging topics. However, plurals are a bad test case due to the Case system in German. Ideally, one would like to have a left-dislocated element that is marked accusative, singular, masculine; this should contrast with the hanging topic which is marked for default Case: nominative in German.

(148) a. * Wer weißt du den wer gesehen hat?
      b. * Welcher von diesen Männern weißt du den welche von diesen Frauen gesehen hat?

Again, this prediction is borne out. The grammatical case of Wh-extraction out of a Wh-island in German can only concern a Wh-topic with a resumptive pronoun, i.e. a Wh-left-dislocated element. This finding is very much in line with the present proposal. It also lends further support to the analysis of left-dislocated elements in general, advocated in Grohmann 1997a. Lastly, it shows a close connection between left-dislocated elements and topics.

     Concluding this sub-section, I would like to show that Wh-left dislocation not only occurs in cases of extraction out of a Wh-island.

     On the one hand, a Wh-topic may be left-dislocated in the presence of another Wh just like a normal topic may (cf. (142)):

(149) a.   Welchen von diesen Männern, wer hat den gesehen?
     b.   Welchen von diesen Männern, den hat wer gesehen?

The derivation is by now well-known: two Top-features are checked overtly by movement (by both Wh-elements), only one Wh-feature is checked (by the lower Wh), and the left-dislocated Wh-topic spells out a resumptive (demonstrative) pronoun.

     However, in the presence of only one Wh-topic, left dislocation is not possible as the Wh-feature cannot be checked:[115]

(150) a. * Welchen von diesen Männern, die Frau hat den gesehen?
     b. * Welchen von diesen Männern, den hat die Frau gesehen?

On the other hand, in these cases the Wh-left-dislocated element underlies the usual restrictions, i.e. it can only apply in V2-contexts:

(151) a. * Welchen von diesen Männern, den wer gesehen hat?
     b. * Welchen von diesen Männern, wer den gesehen hat?

     We have seen that Wh-topics are not so unusual; this sub-section has further shown that Wh-topics can also be left-dislocated; crucially, however, they cannot function as hanging topics.

4.2.4 Results

After an extensive discussion of the data and possibilities of Wh-extraction out of Wh-islands across languages, we can confirm the observation that only D-linked Wh-elements may be extracted to some extent. Under the view proposed here, these elements need to check a Top-feature on their way to the initial position of the matrix clause. It is thus more accurate to say that only Wh-topics can be extracted out of Wh-islands.[116]

     As the data have shown, it cannot be the Top-feature that allows extraction; rather the Top-feature allows movement beyond the lower Wh, yet still inside the embedded clause. It is the Wh-feature on matrix C^0 that forces further movement.

     We can now modify the generalizations from (138) as follows:

(152) a. $[_{CP}$ Wh$_i$ [… $[_{IP}$ … $[_{TopP}$ t$_i$ … $[_{CP}$ Wh … $[_{IP}$ … t$_i$…]]]]]
     b. * $[_{CP}$ Wh$_i$ [… $[_{IP}$ … $[_{CP}$ Wh … $[_{IP}$ … t$_i$…]]]]]
     c. * $[_{TopP}$ Wh$_i$ [… $[_{IP}$ … $[_{TopP}$ t$_i$ … $[_{CP}$ Wh … $[_{IP}$ … t$_i$…]]]]]

---

[115] The ungrammaticality of (150) also shows that [Top] must be checked prior to [Wh].

[116] Admittedly, the fact that (certain) D-linked Wh-phrases are endowed with [Top] in all languages has not been established here; the prediction is clear, though, based on the present proposal. It is also most certainly and smoothly in accordance with all data presented in this study.

     With respect to German and the cases of apparent "Wh-left dislocation," more data will be needed to conclusively show a generally allowed pattern. What the data presented here show, however, is clearly an exception to the standardly assumed generalization that Wh-extraction out of Wh-islands is always and categorically ungrammatical.

d. * $[_{CP}$ Wh$_i$ $[\ldots$ $[_{IP}$ $\ldots$ $[_{CP}$ Wh $\ldots$ $[_{TopP}$ t$_i$ $\ldots$ $[_{IP}$ $\ldots$ t$_i$$\ldots]]]]]$
e. * $[_{TopP}$ Wh$_i$ $[\ldots$ $[_{IP}$ $\ldots$ $[_{CP}$ Wh $\ldots$ $[_{TopP}$ t$_i$ $\ldots$ $[_{IP}$ $\ldots$ t$_i$$\ldots]]]]]$

The only way to escape a Wh-island is by topicalizing above the lower Wh and further Wh-move to the matrix CP (152a). All other possibilities are ruled out: long movement to matrix CP over a filled embedded CP (152b), movement to matrix TopP (153b) and extraction out of embedded TopP situated below CP to either matrix CP (152d) or matrix TopP (152e).

4.3 Cross-Linguistic Differences Among Multiple Wh

I will now turn to cross-linguistic considerations of D-linking in the bigger framework outlined here.[117] I propose a tripartite classification of languages and the felicity constraints on Multiple Wh. We already have three types of languages with respect to Wh-movement to CP (overt, covert and mixed). The present classification is similar in spirit: do languages overtly topicalize multiple Wh-phrases, covertly or mixed? This classification is very tentative; moreover it just looks at some obvious syntactic possibilities (fronting of Wh-subject vs. Wh-object) based on pragmatic (= discourse) restrictions. It does not make any deeper predictions as to what might be going on in other languages. It is, however, useful to get a theory-independent grip on these restrictions.

4.3.1 Chinese and German

German and Chinese Multiple Wh differ in some respects. For one, all Wh-elements obligatorily move overtly in German but not in Chinese. Secondly, a distributive reading can be applied to any felicitous Multiple Wh in German, while it can only be assigned to those constructions in Chinese where both Wh-elements stay *in situ*.

But the contextual influence on Wh-topics is the same in both languages. This sub-section is concerned with finding similarities between these two languages that cannot be found in the other languages under investigation (see below). In this respect, it is important to bear in mind that both languages (i) exhibit Wh-topics and, more importantly, (ii) underlie the same contextual restrictions on the felicity of multiple interrogatives. We have seen the contrast between English and German in some detail in section 3.1—let us now turn to the similarities between German and Chinese given the same circumstances.[118]

Consider thus the following situations (verified by Jianxin Wu, personal communication):[119]

(153) Situation I: A salesman comes home and speaks to his wife.
    "I sold a few things today."
    a. # She mai-le shenme?
    b. # Shenme mai-le she?
    c. # Wer hat was gekauft?
    d. # Was hat wer gekauft?
(154) Situation II: A salesman comes home and speaks to his wife.
    "I sold a car, a fridge and a TV today."
    a. # She mai-le shenme?
    b. # Shenme mai-le she?
    c. # Wer hat was gekauft?
    d. # Was hat wer gekauft?

In neither language is it possible to ask a Multiple Wh if reference has been made to only one Wh-element. That this is the right conjecture can be witnessed from the following:

(155) Situation III: A salesman comes home and speaks to his wife.
    "I sold a car, a fridge and a TV to an old man, a young woman and a teenager."

---

[117] This project allows only for a very limited and crude look into discourse restrictions across languages.
[118] For the sake of exposition, I only present data for the felicity of Multiple Wh involving a Wh-subject and a Wh-object, here and in the following.
[119] Here and in the remainder of the section, the questions in all languages are to be translated as *Who bought what?* in the a- and c-examples, and *What did who buy* in the b- and d-examples, similar in essence to (59).

      a.   She mai-le shenme?
      b.   Shenme mai-le she?
      c.   Wer hat was gekauft?
      d.   Was hat wer gekauft?

If both Wh-elements are specified to some degree (basically following DRQ), German and Chinese allow a felicitous Multiple Wh.

4.3.2 Spanish and Hebrew

The second type of discourse restrictions on Multiple Wh can be observed in Spanish and Hebrew. Although we have only seen limited data from these languages, the important aspect to note is that both languages allow the Wh-object to be fronted over the Wh-subject. The syntax of Multiple Wh in Hebrew and Spanish shall not be the topic of this paper; rather, I am interested in showing that the contextual restrictions applying to Multiple Wh are the same in both languages, while they differ from German and Chinese on the one hand, and from English on the other.

These two languages differ from German and Chinese in that not the referent sets of all Wh-phrases must be established in order to ask a Multiple Wh, but only of one.[120]

Employing the same situations as above, here are the results. (The following data are provided by Danny Fox, Idan Landau for Hebrew and Juan Carlos Castillo for Spanish.):[121]

(156) Situation I: A salesman comes home and speaks to his wife.
    "I sold a few things today."
      a.   Mi kana ma?
      b. # Ma kana mi?
      c.   Quién compró qué?
      d. # Qué compró quién?
(157) Situation II: A salesman comes home and speaks to his wife.
    "I sold a car, a fridge and a TV today."
      a.   Mi kana ma?
      b. # Ma kana mi?
      c.   Quién compró qué?
      d. # Qué compró quién?

We can see from the above that if the referent of the Wh-object is known (157)—or even if only the object is mentioned, but no known (156)—a Multiple Wh is felicitous; it is only grammatical (and felicitous), however, if the Wh-subject precedes the Wh-object. In this way, Hebrew and Spanish pattern like English.

They pattern like German and Chinese if the referents of both Wh-elements are established by the discourse (what we also observed for English D-linked Wh).

(158) Situation III: A salesman comes home and speaks to his wife.
    "I sold a car, a fridge and a TV to an old man, a young woman and a teenager."
      a.   Mi kana ma?
      b.   Ma kana mi?
      c.   Quién compró qué?
      d.   Qué compró quién?

Also, Spanish and Hebrew differ from the other languages considered: if only reference to the Wh-subject has been made, a Multiple Wh is well-formed with a fronted Wh-object.[122]

(159) Situation IV: A salesman comes home and speaks to his wife.

---

[120] In this respect, they pattern much like English.
[121] The range of examples and the underlying interpretations are illustrated in a very restricted way here; more research will be necessary, not only in Hebrew and Spanish but more cross-linguistically.
[122] Recall that this section yields but preliminary results. A more thorough investigation of these issues will have to wait for a future project.

"An old man, a young woman and a teenager came into my store today."
a. # Mi kana ma?
b.   Ma kana mi?
c. # Quién compró qué?
d.   Qué compró quién?

4.3.3 English and the Comparison

The two classes as presented above both differ from English in several respects.[123] While Multiple Wh-questions are only well-formed in Chinese and German (Class I) if the referent sets of both, Wh-subject and Wh-object, have been introduced in the discourse, reference to one of the two is sufficient to license a Multiple Wh in Spanish and Hebrew (Class II). Furthermore, once a Multiple Wh is felicitous in Class I languages, either order of the Wh-elements is fine; in Class II languages, known reference to the Wh-object forces fronting of the Wh-subject and known reference of the Wh-object forces fronting of the Wh-subject—if both are know, either order is grammatical. English (Class III) differs from both classes in that the Wh-object may not be fronted (unless clearly D-linked, more clearly than in the contexts above).

The following table summarizes how the three classes differ with respect to established Wh-referent sets and subject/object-order of Wh-elements.

(160) *Table 1: Obviation of Standard Superiority Effects Based On Discourse Information*

Given:	Wh-SU… Wh-OB	Wh-OB… Wh-SU
Referent of Wh-SU	Class III	Class II
Referent of Wh-OB	Class II, Class III	*
Referent of both Wh	Class I, Class II, Class III	Class I, Class II

That DRQ and D-linking are related in certain ways has become clear in the course of the discussion. D-linking is an option for English, where no movement of the Wh-element is possible (Wh-*in-situ*). Apparent D-linking properties of German Wh-phrases can be accounted for under a DRQ-approach. In this sense, DRQ might be seen as the movement variant of D-linking.

4.4   Child Language and Learnability[124]

An obvious question that arises from any discussion of a phenomenon that seems to be present in a number of languages yet is instantiated differently is how to account for the variation. Parts of the discussion so far have been concerned with this question. The follow-up, however, is what—if anything—does this variation have to do with the language learner?

4.4.1 Cues for Wh-Adjuncts

One issue we have dealt with to some extent is that of the difference of *why*/*how* vs. *warum*/*wie* (English vs. German) and vs. *when*/*where* (in English). One possibility is that *why* and *how* are different from the other Wh-adjuncts in English by virtue of being obligatorily base-generated in CP. All German Wh-adjuncts, on the other hand, move to CP, just like *where* and *when*. The task for the learner is then to come across positive evidence that allows it to assign the correct structures for the respective elements.

Under a cue-based approach to learnability (Lightfoot 1998), unambiguously simple ("Degree-0," Lightfoot 1991) positive data enable the learner to figure out the correct structure (and hence correct derivation) for the construction in question. With respect to the Wh-adjunct asymmetry in English and the question of movement vs. generation, the examples with interjections plausibly serve as cues:

---

[123] Observe that all examples employed bare Wh-phrases which are not necessarily D-linked in English, and plausibly not in other languages, too. A more thorough discussion will have to take D-linking into consideration and examine in detail what accounts for D-linking in which language (and what other restrictions exist).
[124] This section contains speculations and hence theory-driven considerations of the issue. I have not conducted experiments or collected data from child language. This, again, suggests plenty of research material.

(161) a.   Why, child, did you eat my dinner?
     b.   How, child, can you eat so quickly?
(162) a. * When, child, did you eat my dinner?
     b. * Where, child, did you hide my dinner?

While (161) shows that some Wh-elements are fine in the pre-interjectional position, the ungrammaticality of others (162) shows that this contrast is due to the prohibition of moving to this position. Of course, (162) is not a possible input for the learner. But hanging topics are, especially those constructions where the hanging topic precedes a Wh; hanging topics furthermore are base-generated—no movement ever targets a pre-CP-position.[125] The child has thus evidence that this left-peripheral position can only be filled via base-generation.

     Let us thus assume that the learner can differentiate between base-generation and movement of Wh-adjuncts. In German, all four constructions are fine. Furthermore, the contrast between hanging topics and left dislocated elements (that does not appear in English) allows the learner to differentiate movement and base-generation correctly. Crucial for the child is that these constructions can be possible evidence, as they are grammatical.[126]

4.4.2 Cues for Wh-Topics

But a more interesting issue—which also has farther-reaching consequences—is the difference among languages with respect to contextual restrictions on multiple interrogatives. Why languages should differ in such restrictions should be a topic for a different study; there seems no obvious explanation based on the (limited) observations this study could provide. A related question, however, can be given a preliminary explanation: what causes the syntactic variation given the contextual restrictions in Multiple Wh?

     Let us start with the most restricted case: DRQ which is at work in German, for example. The contextual restriction is that the reference of both Wh-elements must have been introduced prior to asking a Multiple Wh. Syntactically, this is expressed by movement to a topic-position in this language where this "referentiality" can be formally checked. Aside from other reasons, evidence for this movement can be found in the structure of the Wh itself: while the topic-feature may be encoded on a functional head attracting this movement, it must have a reflex coded onto the Wh. This reflex, I argue, roots inside the more delicate structure construed for the Wh: these Wh-elements behave like partitive interrogatives and hence consist underlyingly of the same structure that partitives do. The topic-feature can then be seen as a requirement of the partitive.

     In German all Wh-elements in Multiple Wh contain the partitive—whether it is overtly expressed or not. The simplest case is that of an overt partitive Wh: here the learner gets a cue from the peculiar behaviour, namely that partitive interrogatives may be split and intervening material is constrained. Consider (122-124) again, shown here in a variation:

(165) a.   Welches von diesen Bildern gefällt dir am besten?
     b.   Welches von diesen Bildern gefällt jedem Kind am besten?
     c.   Welches von diesen Bildern gefällt keinem Kind am besten?
(166) a.   Welches gefällt dir von diesen Bildern am besten?
     b.   Welches gefällt jedem Kind von diesen Bildern am besten?
     c. # Welches gefällt keinem Kind von diesen Bildern am besten?

This is the paradigm that every adult speaker of German knows. The learner, however, can deduce from possible positive evidence such as (166a) that the partitive Wh may be split into the interrogative part and the partitive part. It thus receives a cue that some Wh-phrases check a Top-feature as well as a Wh-feature, and that such phrases may be split up.[127] (166b)

---

[125] I refer to the discussion in sections 3.3.3 and 4.2.3 above.
[126] Variation among German dialects with respect to allowing *warum* and *wie* in a position other (= lower) than CP (as found by Haider (1996)) may be deduced from these facts, too; in order to positively support this hypothesis, though, more relevant data needs to be collected from speakers of that dialect.
[127] I assume that the topic-feature is universal to partitives, and if a learner comes across a partitive and correctly analyses it as such, it knows that [Top] is the relevant feature. Without a thorough discussion of language acquisition, I refer the reader to the works of Crain and Thornton (1998) or Lightfoot (1998), for example, as well as the references cited there. The general assumption here is something like the Strong Continuity Hypothesis, i.e. that we can assume maximal knowledge for minimal input.

gives the learner a cue that some quantifiers may intervene between the Wh-part and the Top-part of the partitive Wh. Knowledge of the behaviour of quantifiers will make available the rest of the paradigm; for example, as soon as the child learns that weak quantifiers may not be topicalized, (166c) can be deduced. It then derives the constructions in (166) correctly by checking the topic-feature prior to checking the Wh-feature.[128]

As the behaviour of partitive Wh-elements in German is practically the same for D-linked and bare Wh-phrases in multiple interrogatives, the learner will assign the syntax of partitives to all Wh-phrases; in other words, it will assume a more intricate structure for all Wh-phrases based on the pattern observed from partitives.[129]

On the other hand, the learner of English will be exposed to different syntactic behaviour between partitives and other Wh-phrases. One unambiguously piece of positive evidence is the contrast observed in (139), repeated here as (167):[130]

(167) a. *   How much money$_i$ do you wonder who to give t$_i$ to?
      b. ?   How much of the money$_i$ do you wonder who to give t$_i$ to?

Suppose the learner detects that partitive Wh-phrases in English differ in their syntactic behaviour from other Wh-phrases; suppose also that the learner correctly identifies this difference with the obligatory movement through TopP. The result is that the learner has now a cue that there are Wh-phrases with more intricate structures that require topicalization. Analogously to possible variation among German dialects, the English learner might then get cues that allow it to extend this structure to other Wh-phrases. Just as the German learners may vary—depending on evidence—whether they should assign the partitive structures to D-linked Wh-phrases and then (in the ideal, standard case) also to bare Wh-phrases (see fn. 130), the English learner may get evidence to assign the partitive structure to D-linked Wh-phrases.[131]

4.4.3 Wh-Topics and Mis-analyses in Child Language

While learners accrue linguistic knowledge in a rapid way, "mistakes" may be made on the way when the learner tests out freshly accumulated data. It is thus predicted that with the intricate behaviour of questions, learners produce mis-analyses to some extent. If this approach and the general picture of language acquisition (such as presented in Crain and Thornton 1998 or Lightfoot 1998) are on the right track, irregularities can be predicted; furthermore, it can be—to some extent—predicted what kinds of mistakes will occur and what kinds will not.

The study on D-linked Wh-elements by Thornton (1995) gives us some nice data that receive a straightforward account in the present framework. Her findings from experiments show that children acquiring English make some interesting mistakes with D-linked Wh-phrases.[132]

---

[128] The intricacies involved here predict that at the beginning stages, mistakes may be made. These mistakes are still within the boundaries of UG, however, so that a learner would never use strategies that are not allowed; it simply mis-analyses them to some degree, often related to economy conditions. Some of these actual mis-analyses will be shown below from English.

[129] This view can then be extended to variation among dialects. I predict that if there are any restrictions in multiple interrogative contexts, the most liberal behaviour will be that of partitives, followed by "heavy" (= D-linked) Wh-phrases and bare Wh-phrases being most restricted. If this is so, the learner will implement these differences by not assigning the deeper (= partitive) structure to bare Wh-phrases first, and then possibly, extend it to "heavy" ones.

[130] Admittedly, I am not an expert on child language. While (167b) may be an unlikely candidate of positive evidence that a child may be exposed to, there are numerous other examples that show different syntactic behaviour of partitives, in "simpler" constructions.

[131] I am not aware of English dialects that show the same syntactic behaviour of partitive Wh-phrases not only with D-linked Wh-phrases, but also with bare ones. If the latter variety does exist, the learner of that particular dialect will thus assign the partitive structures to these elements too.

[132] I cannot go into any more details of Thornton's paper but have to pick the relevant data selectively. Her findings pretty convincingly argue for the use of "referentiality" as a productive property of certain Wh-phrases by some learners; put in current terminology, she shows that children employ TopP in their constructions of questions with D-linked Wh-phrases. Thornton labels this projection Ref(erential)P and actually assumes it to precede CP; this, however, is a sign of the times: both, the label and the position are adopted from Stowell and Beghelli 1994 in the GB-framework. In minimalist terms—especially the framework proposed here—RefP can

Consider (168) from Thornton 1995:163:

(168) a.   Which Smurf the big man couldn't pick up?
      b.   Which mouse what the cat didn't see?

Granted that the children who uttered these questions know of Wh-topics pretty much in line with the current proposal, both constructions receive straightforward derivations which—"wrong" as they may be—are the results of not getting it right immediately.
    Allow the following to be the possible structural representations for (168):

(169) a. $[_{CP} C^0 [_{TopP}$ [which Smurf]$_i$ Top^0 $[_{IP}$ the big man couldn't pick up t$_i$]]]
                    [Wh]   [Wh] ~~[Top]~~   ~~[Top]~~
      b. $[_{CP}$ [which mouse]$_i$ C^0 $[_{TopP}$ what$_i$ Top^0 $[_{IP}$ the cat didn't see t$_i$]]]
                    ~~[Wh]~~   ~~[Wh]~~   ~~[Top]~~ ~~[Top]~~

Most notably, in (169a) the inflected auxiliary fails to be inverted. Inversion (or *do*-support) is obligatory in English questions. However, it is not obligatory (and thus not possible) in topicalized constructions.

(170) a.   Who did you see?
      b. * Who you saw?
(171) a.   This man, I saw
      b. * This man did I see

The structure of (168a) at Spell-Out is thus TopP, not CP. The child correctly moves *which Smurf* to SpecTopP to check the Top-feature. It then does not move further to CP, presumably, this movement will occur at LF where the Wh-feature will be checked. Hence, inversion does not take place. The reasoning here might be that while the child is aware that [Top] and [Wh] need to be checked, it fails to do both overtly, presumably for economy reasons: covert movement is "cheaper" than overt movement and the move from TopP to CP is vacuously anyway. This might be the trigger for the child to understand the second move as not obligatory for the overt syntax.
    In the other construction, however, the child checks both features overtly.[133] First, *which mouse* moves to TopP, then it moves to CP. If we assume that the child moves the Wh to TopP because it assigns this D-linked Wh the partitive structure, the overt realization of *what* in SpecTopP might be the result of a misconception of splitting partitives: the child spells out the wrong part of the phrase and moves the "remainder" to CP. Crucially, this remainder contains the real Wh-element so that Wh-checking may take place. Note that English partitive interrogatives may not be split; however, German partitives may be split and this is thus certainly an option made available by UG.[134] With the partitive structure being very abstract, to say the least, it is quite natural that the child makes a mistake with the spelled out material left behind in TopP.
    In sum, evidence from English learners support the view that partitive interrogatives exist at some stages in English; it furthermore supports the analysis that such elements contain a topic-feature. The examples above also show conclusively that TopP must be situated below CP: this way, we can account for the lack of inversion.
    The reader will observe from the preceding discussion that the acquisition data used only involve one Wh. The development of multiple interrogatives takes place rather late; one

---

[133] Note that the two questions were uttered by two different children.
[134] Also recall (90b) where a Wh is split even in English, repeated below for convenience.

(i) Hey, hey LBJ, how many did cigs you smoke today?

    Whether the splitting is the result of some metric/poetic requirement, it is not unintelligible for speakers of English. The deeper structure to be assumed for *how many…cigs* in this case is certainly that of a partitive. (Compare with the many German examples provided here.)

possible explanation may be the intricate syntax involved and the learner needs to have time to master D-linking as it is—without even considering topicalization in much detail. Interestingly, deVilliers et al. (1996) show that when German learners start to employ Multiple Wh they use those constructions that conform with standard conceptions of the Superiority Condition, i.e. Multiple Wh-constructions in which Wh-subjects precede Wh-objects. More data from early German is needed on all these interrogative contexts to explain this; but one conjecture may be that the learner has difficulties in employing the topics correctly and thus refrains from using them. Positive evidence for Multiple Wh cannot easily be found, judging by the rarity in which they occur in adult language. This sort of "unnaturalness" might also be yet another reason for variation of judgements among speakers, especially when the constructions get more intricate.

5 Conclusion[135]

The goal of this paper was to investigate the nature of multiple interrogatives; in particular I went at length to propose a syntactic account within the minimalist framework (Chomsky 1995) for constructions that are restricted by the discourse—in severity as well as in variation. The initial thesis that Wh-movement may be influenced by movement processes prior to Wh-movement turns out to be borne out. Unfortunately, the exposition of this thesis took a long time and needed much space. Fortunately, many different issues could be investigated.

With evidence from a number of languages, I set out to consider movement processes prior to Wh-movement in German. The obvious candidate is scrambling, or A-movement in general, following the general path that Boskovic (1993) walked for Spanish (and Hebrew), Takahashi (1993) for Japanese and Hornstein (1995) for English. I myself also went there for German (Grohmann 1997b). There is, however, an end to this path. Movement for the purpose of already-known features prior to Wh-movement is not enough.

It became clear that discourse plays a crucial role. Pesetsky (1987) accounted for this with a theory of D(iscourse)-linking. I took the rather unclear details of D-linking as the starting point to re-consider the facts and suggest an elaboration. Finding that German Multiple Wh underlie much stricter discourse restrictions than English, I propose that in German Multiple Wh-constructions both Wh-elements are obligatorily D-linked in a special sense: they must conform to D(iscourse) R(estricted) Q(uantification). In general terms, this means that the referent sets of both Wh-elements must have been introduced in the discourse in order to formulate a felicitous double question. Technically, I propose that DRQ is instantiated by movement to the left periphery (Rizzi 1997) of both Wh-elements. Thus, DRQ fits out a Wh with a feature [Top], to be checked at TopP.

The structure assumed here for partitives helps to visualize this abstract and at first glance ad hoc feature: internally to the DP-structure of a partitive, [Top] is present, namely on the partited DP (the DP following the partitioner). German Wh-elements have thus to be assigned the more intricate structure of partitive interrogatives in Multiple Wh, even if they are only bare Wh-phrases. Once both Wh-elements have checked [Top], the higher Wh moves to CP to check [Wh], in accordance with the MLC.

This approach may carry over to a lesser degree to English D-linked Wh-phrases. Especially in instances where partitive Wh-phrases differ from simpler ones, this approach makes the right predictions. One such instance is extraction out of Wh-islands (inter alia Kiss 1993, Comorovski 1996). The result of a deeper discussion of Wh-extraction out of Wh-islands is that D-linked Wh-phrases may be extracted, when they can be assigned the (abstract) partitive structure; or in other words: Wh-topics may move out of Wh-islands. The

---

[135] I would like to mention that an extension of the current work with respect to some technical and theoretical details will be made in Grohmann, in progress. In particular, the role of AgrP will be evaluated as well as application of checking operations applying to "prolific domains." The general gist, though, will remain, namely that Multiple Wh in German involves multiple Wh-topicalization.

landing site is matrix CP and the jumping site is a high TopP, situated above the embedded CP.

The initially counterintuitive notion of a Wh-topic deserves special consideration. Wu (1996) conclusively shows that there are instances in Chinese Wh-formation where a Wh is fronted; this cannot be the result of checking [Wh]. Wu shows that it is in fact topicalization. In many instances, the notion of a Wh-topic has been explored, employed and elaborated. It designates those Wh-elements that have an underlyingly deeper structure, equipped with a topic-feature (partitive).

In the course of the proposal, some apparent counterexamples and other difficulties have been pointed out, addressed and sometimes even solved. One particularly nasty issue is that of (Wh-)subjects. On the grounds of thematic properties intrinsic to subjects, I show that in some problematic cases, the subject topicalizes also. Moreover, this topicalization is a different process than normal topicalization, as it is more an information-structurally driven property than some morphological feature that motivates this move. While capable of handling the data, this approach throws up a number of other problems. These, as many others, need to be subjected to scrutiny in future work.

With respect to variation, suggestions were advanced that parameterize English D-linking and German DRQ to some degree, also allowing for further variation. Important considerations come from child language (which, even in English, seems to wonderfully go hand in hand with the present proposal; see Thornton 1995) and learnability issues.

Unfortunately, the number of issues followed up here required a density that might blur some of the more important points. Fortunately, I am confident that these points are relevant indeed. A unique example of Wh-extraction out of Wh-islands in German has been observed and assigned a straightforward explanation within the given framework. Many more interesting correlations have been found, too—one of which is that the input of scrambling in obviating German Superiority effects results most naturally as a corollary of my proposal, rather than the driving force; that the latter is not warranted for is apparent: scrambling is not a unified phenomenon with respect to one particular property of the moved elements. It can be, however, topicalization, namely in those instances in German where objects are scrambled above the subject.

This extensive study proposes to take discourse restrictions on Multiple Wh-constructions across languages seriously into syntactic considerations. The suggestion made here is that such restrictions may be due to the nature of structurally more complex Wh-phrases; this structure, in turn, forces movement prior to Wh-movement which is understood as topicalization to license the contextual factors. While this partially extensive study captures a wide range of data and (at first glance unrelated) phenomena, it has always been pointed out that more work needs to be done.

References

Abraham, Werner, Samuel D. Epstein, Höskuldur Thráinsson and C. Jan-Wouter Zwart, eds. 1996. *Minimal Ideas: Syntactic Studies in the Minimalist Framework*. Amsterdam: John Benjamins.

Abraham, Werner and Kleanthes K. Grohmann, eds. 1997. *Groninger Arbeiten zur germanistischen Linguistik* 40. Rijksuniversiteit Groningen: Germanistisch Instituut.

Alexiadou, Artemis. 1997. *Adverb Placement: A Case Study in Antisymmetric Syntax*. Amsterdam: John Benjamins.

Anagnostopoulou, Elena, Henk van Riemsdijk and Frans Zwarts, eds. 1997. *Materials on Left Dislocation*. Amsterdam: John Benjamins.

Aoun, Joseph, Norbert Hornstein and Dominique Sportiche. 1981. Some Aspects of Wide Scope Quantification. *Journal of Linguistic Research* 1: 69-95.

Aoun, Joseph and Y.-H. Audrey Li. 1993. *Syntax of Scope*. Cambridge, MA: MIT Press.

d'Avis, Franz-Josef. 1996. On Wh-Islands in German. In Lutz and Pafel, 89-120.

Baker, C. Lee. 1970. Notes on the Description of English Questions: The Role of an Abstract Question Morpheme. *Foundations of Language* 6: 197-217.

Beck, Sigrid. 1996. Quantified Structures as Barriers for LF-Movement. *Natural Language Semantics* 4: 1-56.

Bolinger, Dwight. 1978. Asking More Than One Thing at a Time. In H. Hiz, ed. *Questions*. Dordrecht: Reidel, 107-150.

Boskovic, Zeljko. 1993. On Certain Violations of the Superiority Condition, AgrO, and Economy of Derivation. Ms., University of Connecticut, Storrs.

Boskovic, Zeljko. 1997. On Certain Violations of the Superiority Condition, AgrO, and Economy of Derivation. *Journal of Linguistics* 33: 227-254.

Boskovic, Zeljko. 1998. Wh-Phrases and Wh-Movement in Slavic. Ms., University of Connecticut, Storrs.

Boskovic, Zeljko. In Press. Sometimes in SpecCP, Sometimes In-Situ. In Martin et al.

Castillo, Juan Carlos. 1998. The Syntax of Container-Content Relations. In E. Murguia, A. Pires and L. Quintana, eds. *University of Maryland Working Papers in Linguistics* 6. University of Maryland, College Park: Department of Linguistics, 32-60.

Chierchia, Gennaro. 1991. Functional Wh and Weak Crossover. In *Proceedings of WCCFL 10*. Stanford: CSLI, 75-90.

Chomsky, Noam. 1973. Conditions on Transformations. In S.R. Anderson and P. Kiparsky, eds. *A Festschrift for Morris Halle*. New York: Holt, Reinhart and Winston, 232-286.

Chomsky, Noam. 1980. On Binding. *Linguistic Inquiry* 11: 1-46.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, Noam. To Appear. Minimalist Inquiries: The Framework. In R. Martin et al.

Cinque, Guglielmo. 1990. *Types of A'-Dependencies*. Cambridge, MA: MIT Press.

Cinque, Guglielmo. In Press. *Adverbs and Functional Heads: A Cross-Linguistic Perspective*. Oxford: Oxford University Press.

Cinque, Guglielmo and Giuliana Giusti, eds. 1995. *Advances in Roumanian Linguistics*. Amsterdam: John Benjamins.

Collins, Chris. 1997. *Local Economy*. Cambridge, MA: MIT Press.

Comorovski, Ileana. 1989. *Discourse and the Syntax of Multiple Constituent Questions*. PhD Diss., Cornell University, Ithaca.

Comorovski, Ileana. 1996. *Interrogative Phrases and the Syntax-Semantics Interface*. Dordrecht: Kluwer.

Corver, Norbert and Henk van Riemsdijk, eds. 1994. *Studies in Scrambling*. Berlin: Mouton de Gruyter.

Crain, Stephen and Rosalind Thornton. 1998. *Investigations in Universal Grammar: A Guide to Experiments on the Acquisition of Syntax and Semantics*. Cambridge, MA: MIT Press.

Culicover, Peter. 1996. On Distinguishing A-bar Movements. *Linguistic Inquiry* 27: 445-463.

Diesing, Molly. 1992. *Indefinites*. Cambridge, MA: MIT Press.

Dobrovie-Sorin, Carmen. 1993. *The Syntax of Rumanian: Comparative Issues in Romance*. Berlin: Mouton de Gruyter.

Dobrovie-Sorin, Carmen. 1997. Review of Comorovski (1996). *GLOT International* 2.8: 15-16.

Engdahl, Elisabet. 1980. *The Syntax and Semantics of Questions in Swedish*. PhD Diss., University of Massachusetts, Amherst.

Epstein, Samuel D. and Norbert Hornstein, eds. Forthcoming. *Working Minimalism*. Cambridge, MA: MIT Press.

Ernst, Thomas. In Press. Scope Based Adjunct Licensing. In *Proceedings of NELS 28*. University of Massachusetts, Amherst: GLSA.

Erteschik-Shir, Nomi. 1998. *The Dynamics of Focus Structure*. Cambridge: Cambridge University Press.

Fanselow, Gisbert. 1997. Minimal Links. Handout, Universität Potsdam.

Grewendorf, Günter and Joachim Sabel. 1997. Wh-Scrambling in the Minimalist Framework. In C. Wilder et al., 58-80.

Grohmann, Kleanthes K. 1996. *Some Remarks on the West Germanic* Mittelfeld*: Scrambling, Case and Pronominalization*. BA Honours Diss., University of Wales, Bangor.

Grohmann, Kleanthes K. 1997a. On Left Dislocation. In Abraham and Grohmann, 1-33.

Grohmann, Kleanthes K. 1997b. German Superiority. In Abraham and Grohmann, 97-107.

Grohmann, Kleanthes K. 1997c. German Superiority: A Discussion of the Cross-Linguistic Phenomenon and an Approach to Solve the German Paradox. Ms., University of Maryland, College Park.

Grohmann, Kleanthes K. 1997d. Pronouns and the Left Periphery of West Germanic Embedded Clauses. In W. Abraham and E. van Gelderen, eds. *German: Syntactic Problems—Problematic Syntax*. Tübingen: Niemeyer, 163-189.

Grohmann, Kleanthes K. 1998a. Some Thoughts on Left Dislocated Constructions. Talk presented at *Colloquium Series*, Rijksuniversiteit Groningen (January 9).

Grohmann, Kleanthes K. 1998b. Speculations on the Syntax and Semantic of German Multiple Interrogatives. In E. Murguia, A. Pires and L. Quintana, eds. *University of Maryland Working Papers in Linguistics* 6. University of Maryland, College Park: Department of Linguistics, 71-101.

Grohmann, Kleanthes K. 1998c. Null Aux Constructions and Minimalism. Ms., University of Maryland, College Park.

Grohmann, Kleanthes K. To Appear. More Extractions with More Questions: DRQ, Multiple Wh and Wh-Islands. *University of Maryland Working Papers in Linguistics* 8.

Grohmann, Kleanthes K. In Progress. *Prolific Peripheries: A Radical View from the Left*. PhD Diss., University of Maryland, College Park.

Haeberli, Eric. 1995. Morphological Case, *pro* and Word Order. Ms., Université de Genève.

Haftka, Brigitta. 1995. Syntactic Positions for Topic and Contrastive Focus in the German Middlefield. In I. Kohlhof, S. Winkler and H. B. Drubig, eds. *Proceedings of the Göttingen Focus Workshop*. Stuttgart/Tübingen: Sonderforschungsbereich 340, 137-158.

Haiden, Martin. 1995. *On the Semantic Evaluation of NP-Syntax*. MA Diss., Universität Wien.

Haider, Hubert. 1996. Towards a Superior Account of Superiority. Ms., Universität Stuttgart.

Heim, Irene. 1982. *The Semantics of Definite and Indefinite Noun Phrases*. PhD Diss., University of Massachusetts, Amherst.

Hendrick, Randolph and Michael Rochemont. 1982. Complementation, Multiple Wh and Echo Questions. Ms., University of North Carolina, Chapel Hill and University of California, Irvine.

Higginbotham, James and Robert May. 1981. Questions, Quantifiers, and Crossing. *The Linguistic Review* 1: 41-79.

Hinterhölzl, Roland. 1997. A VO-Based Approach to Verb Raising. In *Proceedings of NELS 27*. University of Massachusetts, Amherst: GLSA, 187-201.

de Hoop, Helen. 1998. Partitivity. *GLOT International* 3.2: 3-10.

Hornstein, Norbert 1995. *Logical Form: From GB to Minimalism*. Oxford: Blackwell.

Hornstein, Norbert. 1997. The Minimalist Program. Ms., University of Maryland, College Park.

Hornstein, Norbert and Amy Weinberg. 1987. Superiority and Generalized Binding. In *Proceedings of NELS 17*. University of Massachusetts, Amherst: GLSA, 311-326.

Huang, C.-T. James. 1982. *Logical Relations in Chinese and the Theory of Grammar*. PhD Diss., Massachusetts Institute of Technology, Cambridge.

Karttunen, Lauri. 1977. Syntax and Semantics of Questions. *Linguistics and Philosophy* 1: 3-44.

Kayne, Richard. S. 1983. Connectedness. *Linguistic Inquiry* 14: 223-249.

É. Kiss, Katalin. 1993. Wh-Movement and Specificity. *Natural Language and Linguistic Theory* 11: 85-120.

Kitahara, Hisatsugu. 1997. *Elementary Operations and Optimal Derivations*. Cambridge, MA: MIT Press.

Kuno, Susumo. 1982. The Focus of the Question and the Focus of the Answer. *Papers from the Parasession on Nondeclarative Sentences*. University of Chicago: Chicago Linguistic Society, 134-157.

Kuno, Susumu and Ken-Ichi Takami. 1993. *Grammar and Discourse Principles: Functional Syntax and GB Theory*. Chicago: University of Chicago Press.

Kuroda, Sige-Yuki. 1988. Whether We Agree or Not: A Comparative Syntax of English and Japanese. *Lingvisticae Investigationes* 12: 1-47.

Laenzlinger, Christopher. 1996. Adverb Syntax and Phrase Structure. In A.-M. Di Sciullo, ed. *Configurations: Essays on Structure and Interpretation*. Somerville, MA: Cascadilla Press, 99-127.

Lasnik, Howard and Mamoru Saito. 1992. *Move α: Conditions on Its Application and Output*. Cambridge, MA: MIT Press.

Lattewitz, Karin. 1997. *Adjacency in Dutch and German*. PhD Diss., Rijksuniversiteit Groningen.

Lewis, David. 1975. Adverbs of Quantification. In E. Keenan, ed. *Formal Semantics of Natural Language*. Cambridge: Cambridge University Press, 3-15.

Lightfoot, David W. 1991. *How to Set Parameters*. Cambridge, MA: MIT Press.

Lightfoot, David W. 1998. *The Development of Language: Acquisition, Change and Evolution*. Oxford: Blackwell.

Lutz, Uli and Jürgen Pafel, eds. 1996. *On Extraction and Extraposition in German*. Amsterdam: John Benjamins.

Mahajan, Anoop. 1990. *The A/A-bar Distinction and Movement Theory*. PhD Diss., Massachusetts Institute of Technology, Cambridge.

Maling, Joan. 1978. An Asymmetry with Respect to Wh-Islands. *Linguistic Inquiry* 9: 75-89.

Martin, Roger, David Michaels and Juan Uriagereka, eds. In Press. *Step by Step: Essays on Minimalism in Honor of Howard Lasnik*. Cambridge, MA: MIT Press.

Milsark, Gary L. 1977. Toward an Explanation of Certain Peculiarities of the Existential Construction in English. *Linguistic Analysis* 3: 1-29.

Miyagawa, Shigeru. 1987. *Wa* and the Wh-Phrase. In J. Hinds, S. K. Maynard and S. Iwasaki, eds. *Perspectives on Topicalization: The Case of Japanese 'wa'*. Amsterdam: John Benjamins, 185-217.

Moltmann, Friedrike. 1991. Scrambling in German and the Specificity Effect. Ms., Massachusetts Institute of Technology, Cambridge.

Müller, Gereon. 1997. Optional Movement and the Interaction of Economy Constraints. In C. Wilder, H.-M. Gärtner and M. Bierwisch, eds. *The Role of Economy Principles in Linguistic Theory*. Berlin: Akademie Verlag, 115-145.

Müller, Gereon and Wolfgang Sternefeld. 1993. Improper Movement and Unambiguous Binding. *Linguistic Inquiry* 24: 461-507.

Muromatsu, Keiko. 1998. *The Syntax of Classifiers*. PhD Diss., University of Maryland, College Park.

Neidle, Carol, Judy Kegel, Benjamin Bahan, Debra Aarons and Dawn MacLaughlin. 1997. Rightward Wh-Movement in American Sign Language. In D. Beerman, D. LeBlanc and H. van Riemsdijk, eds. *Rightward Movement*. Amsterdam: John Benjamins, 247-278.

Pesetsky, David. 1982. *Paths and Categories*. PhD Diss., Massachusetts Institute of Technology, Cambridge.

Pesetsky, David. 1987. *Wh-in situ*: Movement and Unselective Binding. In A.G.B. ter Meulen and E. Reuland, eds. *The Representation of (In)definiteness*. Cambridge, MA: MIT Press, 98-129.

Pesetsky, David. 1998. Lifestyles of the *Which* and Famous: English is Bulgarian. Talk presented at *Colloquium Series*, University of Maryland, College Park (March 6).

Pollock, Jean-Yves. 1989. Verb Movement, UG and the Structure of IP. *Linguistic Inquiry* 20: 365-424.

Postma, Gertjan. 1995. *Zero Semantics: A Study of the Syntactic Conception of Quantificational Meaning*. The Hague: Holland Academic Graphics.

Reinhart, Tanya. 1995. *Interface Strategies*. Rijksuniversiteit Utrecht: OTS Working Papers.

Richards, Norvin. 1997. *What Moves Where When in Which Language?* PhD Diss., Massachusetts Institute of Technology, Cambridge.

van Riemsdijk, Henk. 1997. *Left Dislocation*. In Anagnostopoulou et al., 1-10.

Rizzi, Luigi. 1978. Violations of the Wh-Island Constraint in Italian and the Subjacency Condition. In C. Dubisson, D.W. Lightfoot and Y.C. Morin, eds. *Montreal Working Papers in Linguistics* 11. McGill, Montreal and Quebec Universities: CL.

Rizzi, Luigi. 1997. The Fine Structure of the Left Periphery. In L. Haegeman, ed. *Elements of Grammar: A Handbook of Generative Syntax*. Dordrecht: Kluwer, 281-337.

Rudin, Catherine. 1985. *Aspects of Bulgarian Syntax: Complementizers and Wh-Constructions*. Columbus, OH: Slavica.

Rudin, Catherine. 1988. On Multiple Questions and Multiple Wh-Fronting. *Natural Language and Linguistic Theory* 6: 445-501.

Stowell, Tim and Filippo Beghelli. 1994. The Direction of Quantifier Movement. Paper presented at *GLOW 17*, Universität Wien (April).

Takahashi, Daiko. 1993. Movement of Wh-Phrases in Japanese. *Natural Language and Linguistic Theory* 11: 655-678.

Thornton, Rosalind. 1995. Referentiality and Wh-Movement in Child English: Juvenile *D-Link*uency. *Language Acquisition* 4: 137-175.

Tsai, W.-T. Dylan. 1994. On Nominal Islands and LF Extraction in Chinese. *Natural Language and Linguistic Theory* 12: 121-175.

Ura, Hiroyuki. 1994. *Varieties of Raising and the Feature-Based Bare Phrase Structure Theory*. Massachusetts Institute of Technology, Cambridge: MITWPL.

Uriagereka, Juan. 1988. *On Government*. PhD Diss., University of Connecticut, Storrs.

Uriagereka, Juan. 1993. Specificity and the Name Constraint. In C.A. Mason, S.M. Powers and C. Schmitt, eds. *University of Maryland Working Papers in Linguistics* 1. University of Maryland, College Park: Department of Linguistics, 121-143.

Uriagereka, Juan. In Press. *A Minimalist Dialogue*. Cambridge, MA: MIT Press.

Vallduví, Enric. 1992. *The Informational Component*. New York: Garland.

deVilliers, Jill, Thomas Roeper and Jürgen Weissenborn. 1996. Superiority: Syntax or Semantics? Talk given at *What the Child Has to Say about Linguistic Theory*, Rijksuniversiteit Utrecht (May).

Watanabe, Akira. 1992. *Wh*-In-Situ, *Subjacency and Chain Formation*. Massachusetts Institute of Technology, Cambridge: MITWPL.

Wilder, Chris, Hans-Martin Gärtner and Manfred Bierwisch, eds. 1997. *Economy Principles in Linguistic Theory*. Berlin: Akademie Verlag.

Wiltschko, Martina. 1997. D-Linking, Scrambling and Superiority in German. In W. Abraham, ed. *Groninger Arbeiten zur germanistischen Linguistik* 41. Rijksuniversiteit Groningen: Germanistisch Instituut, 107-142.

Wu, Jianxin. 1996. Wh-Topic, Wh-Focus and Wh-*In situ*. In J.C. Castillo, V. Miglio and J. Musolino, eds. *University of Maryland Working Papers in Linguistics* 4. University of Maryland, College Park: Department of Linguistics, 173-192.

Zwart, C. Jan-Wouter. 1993. *Dutch Syntax: A Minimalist Approach*. PhD Diss., Rijksuniversiteit Groningen.

Zwart, C. Jan-Wouter. 1996. "Shortest Move" Vs. "Fewest Steps". In Abraham et al., 305-327.

Zwart, C. Jan-Wouter. 1997a. *Morphosyntax of Verb Movement: A Minimalist Approach to the Syntax of Dutch*. Dordrecht: Kluwer.

Zwart, C. Jan-Wouter. 1997b. Transitive Expletive Constructions and the Evidence Supporting the Multiple Specifier Hypothesis. In W. Abraham and E. van Gelderen, eds. *German: Syntactic Problems—Problematic Syntax*. Tübingen: Niemeyer, 105-134.

Kleanthes K. Grohmann
Department of Linguistics
1401 Marie Mount Hall
University of Maryland
College Park, MD 20742
United States of America