Syntactic Inquiries into Discourse
Restrictions on Multiple Interrogatives

Kleanthes Kostas Grohmann
University of Maryland, College Park
grohmann@wam.umd.edu

Generals Paper
Defense Draft
LING 895

April 1998

Advisory Committee:
Prof. Norbert Hornstein (main advisor)
Prof. David Lightfoot (committee member)
Prof. Juan Uriagereka (committee member)
# Table of Contents

Abstract iii
Acknowledgements iv
1 Introduction 1

2 Superiority Effects and “Operations Prior to Wh-Movement” 3
   2.1 Multiple Wh and Superiority Issues 3
      2.1.1 The Superiority Condition and Issues Regarding Multiple Wh 3
      2.1.2 Past Approaches to Multiple Wh 5
      2.1.3 A Minimalist Conception of Multiple Wh 7
      2.1.4 Multiple Wh in German, Extended Data and More on Superiority Effects 8
   2.2 Overt A-Movement Prior to Wh-Checking 11
      2.2.1 Hornstein (1995): Raising and Multiple Wh in English 11
      2.2.2 Takahashi (1993): Scrambling and Multiple Wh in Japanese 12
      2.2.3 Boskovic (1993): AgrOP and Multiple Wh in Spanish (and Hebrew) 14
   2.3 A Possible Application to German 15
      2.3.1 The Agr-Complex in German and Scrambling Below the Subject 15
      2.3.2 The CP-Domain: Pre-Subject Scrambling as Topicalization 16
      2.3.3 Grohmann (1997b): Scrambling and Multiple Wh in German 18

3 DRQ as “Operations Instead of Wh-Movement” 20
   3.1 D-Linking I Not Enough 20
      3.1.1 Context in English and D-Linking 21
      3.1.2 Context in German and DRQ 23
      3.1.3 More Data on Contexts in German 26
      3.1.4 Crucial Differences 30
   3.2 DRQ: Multiple Wh Is Topicalization 33
      3.2.1 Outset: The Problem 33
      3.2.2 DRQ and Wh-Topics 35
      3.2.3 The Rough Syntax 37
   3.3 Extensions of Wh-Topics 42
      3.3.1 Subjects and Multiple Wh: Some Consequences 42
      3.3.2 Scrambling and Multiple Wh: In Response to Wiltschko (1997) 44
      3.3.3 Wh-Adjuncts in Multiple Wh-Constructions: Some Speculations 48
      3.3.4 Multiple Wh and Quantifier Interaction: A Reaction to Beck (1996) 52
4 Cross-Linguistic Variation, Learnability and More Support for Wh-Topics

4.1 The Availability of Wh-Topics in UG

4.1.1 Chinese Wh-Topics
4.1.2 Japanese Wh and the Topic Marker wa
4.1.3 Wh-Topics in General

4.2 Wh-Islands and Wh-Topics

4.2.1 Some Basic Data
4.2.2 Slavic and the Position of TopP
4.2.3 English and German: Some Apparent Problems
4.2.4 Results

4.3 Cross-Linguistic Differences Among Multiple Wh

4.3.1 Chinese and German
4.3.2 Spanish and Hebrew
4.3.3 English and the Comparison

4.4 Child Language and Learnability

4.4.1 Cues for Wh-Adjuncts
4.4.2 Cues for Wh-Topics
4.4.3 Wh-Topics and Mis-analyses in Child Language

5 Conclusion

References

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 underlying 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.
Acknowledgements

This study constitutes the last step towards being admitted to doctoral candidacy. For getting where I am now, I would like to thank Norbert Hornstein, my main advisor, for being there for me when I thought I needed him (and not being there when he didn’t think so), discussing ideas and lending support throughout. Without his clever suggestions when I felt not so clever, I would have never finished this paper. Thanks to his presence, I also face the chance that I will not let this study rest as it is for a long time, either. Welcome new work!

Likewise, Juan Uriagereka and David Lightfoot deserve my gratitude for their support. Being involved as part of the advisory committee, they read and discussed earlier versions and ideas readily which resulted in re-thinking of some issues on my part.

I am happy to thank Werner Abraham, Ian Roberts and Anna Roussou for having been there for me from the beginning and hopefully continuing to do so. Mark Arnold, Aitziber Atutxa, Larisa Avram, Laura Benua, Jonathan Bobaljik, Zeljko Boskovic, Kerstin Bücher, Juan Carlos Castillo, Matthew Chamberlain, Stephen Crain, Lisa Dix, John Drury, Nomi Erteschik-Shir, Rikardo Etxepare, Gisbert Fanselow, Kai von Fintel, Danny Fox, Koldo Garai, Anastasia Giannakidou, Simon Greenhalgh, Christoph Grönegress, Martin Heckmanns, Elena Herburger, Stijn Hoppenbrouwers, Idan Landau, Howard Lasnik, Winfried Lechner, Akemi Masuya, Julien Musolino, Javier Ormazabal, David Pesetsky, Susan Powers, Norvin Richards, Tom Roeper, Itziar San Martin, Matthew Seaman, Caro Struijke, Rozz Thornton, Andreas Wadle, Gert Webelhuth, Amy Weinberg, Norbert Wienke, Martina Wiltschko, Jianxin Wu, Susi Wurmbrand, Yago Zens, Jan-Wouter Zwart deserve special thanks for discussions, data and dear friendship as well as encouragement along the way, as do most likely a number of people I couldn’t think of at 6:51am; I also received plenty of all this from my dear Jennifer Graham. Special thanks here too.

Lastly, I would like to thank for the opportunities to present versions of the present paper at various occasions and especially the audiences from the Colloquium Series at the Rijksuniversiteit Groningen (January 13, 1997), the Annual Student Conference at the University of Maryland in College Park (February 28, 1997), the Spring Meeting of the Linguistics Association of Great Britain in Edinburgh (April 10-11, 1997) and the Comparative Germanic Syntax Workshop 13 at Cornell University (July 18-19, 1997).

Also, without the support from my fellow students—intellectually as well as socially—would I not have been able to enjoy the past two years the way I did? Kathi Faulktingham’s efficient help also made the weird life in the States not as hard as it could have been.

Private thank-you’s go to my entire family, especially Athena, Waltrud, Katharina (and Matt), Alexander and Willy Grohmann as well as Miss Emma, for their great support (emotionally, financially and what-not-ly). A special dedication goes to my sweetie. Thank you, Jennifer, for all that you are! Vielen Dank, Ihr Lieben!

I’m also grateful to have Beny as a special friend. He surely made my life riche!

The usual disclaimers apply (= none of these people can be held responsible for the content of this paper, the interpretation of their suggestions is entirely my fault, no private secretary can be accounted for typos, nobody but me can be blamed for sleep deficit etc.).
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 the past year and a half or so (Grohmann 1997b, 1997c, in press), 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. 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. 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.

---

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)), 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- versus post-Spell Out” but rather to movement that targets a projection below the Wh-projection.
The syntactic framework adopted here is the Minimalist Program in its (more or less) well understood current manifestation. 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 versus German, as illustrated in (1) and (2):

(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.

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 Technical implementations—when relevant—will be introduced and clarified in the text.

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 edite by Abraham et al. (1996), Abraham and Grohmann (1997), Wilder et al. (1997), Epstein and Hornstein (forthcoming), Martin et al. (in press) as well as studies by Collins (1997), Hornstein (1997), Kitahara (1997), Zwart (1997a), Chomsky (1998), 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 give 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.
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. It will be concluded that these inquiries are preliminary indeed; yet, they offer plenty of research material for the future.

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. 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

---

\(^6\) 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” versus “complex” Wh-phrases will be illustrated shortly.
(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
       \ldots X \ldots \left[ a \ldots Z \ldots \right| WYV \ldots
       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 dominating A
dommates 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.

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.\(^8\)

\(^8\) 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)
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, a question such as (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.\(^9\) (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):

(9)  a.  \[CP [who , whom ] [IP t [vp t [kiss t ]]]\]
     b.  \[CP [whom ] [who t [IP t [vp t [kiss t ]]]\]

but also in other languages (Boskovic 1993, Takahashi 1993 and 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.\(^9\)

\(^9\) It has also been observed (Comorovski 1996) that each Wh should denote more than one referent, i.e. the following is not an 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
Assuming Wh-movement at LF, the trace of \textit{whom} is lexically governed by the verb \textit{kiss} in (25a), while the trace of \textit{who} is antecedent governed by \textit{who} (made possible through COMP-indexing). On the other hand, the trace of \textit{who} in the LF-representation (25b) is not properly governed: it adjoins to \textit{whom} which percolates its index to COMP (not \textit{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):\(^{10}\)

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

In the grammatical construction, both Wh-traces are properly governed at LF: \textit{persuade} lexically governs the trace of \textit{who} and \textit{buy} the trace of \textit{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 somehow under the ECP, namely when they need to be antecedent governed.

Based on this and other empirical shortcomings, Pesetsky (1987) proposes to evaluate well-formedness of Wh-movement on the basis of the Nested Dependency Condition (NDC).\(^{11}\) 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:\(^{12}\)

(11) a. What did you wonder who (PRO) to see t in t?  
   \[ \text{[\text{who (PRO) to see}] \text{[t in t]}} \]

   b. * Who did you wonder what (PRO) to see t in t?  
   \[ \text{[\text{what (PRO) to see}] \text{[who in t]}} \]

Here, the path of the grammatical Wh-trace dependency is contained in the other, while in the ungrammatical case they overlap (“cross”).

\(^{10}\) Hornstein and Weiberg (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).

\(^{11}\) Or, as in Pesetsky 1982, the Path Containment Condition.

\(^{12}\) 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 \textit{which}-phrases, which shall be not our concern here, though:

(i) a. Which movie, did you wonder which actor, (PRO) to see t\textsubscript{i} in t\textsubscript{j}?  
   b. * Which actor, did you wonder which movie, (PRO) to see t\textsubscript{j} in t\textsubscript{i}?
Furthermore, 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, 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 (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):\(^{13}\)

\[(12) \text{Minimal Link Condition} \]
\[
K \text{ attracts } \alpha \text{ only if there is no } \beta, \beta \text{ closer to } K \text{ than } \alpha, \text{ such that } K \text{ attracts } \beta.
\]

\[(13) \text{Closeness} \]
\[
\beta \text{ is closer to } H(K) \text{ than } \alpha \text{ iff } \beta \text{ c-commands } \alpha, \text{ and } \beta \text{ is not in the minimal domain of } CH, \text{ where } CH \text{ is the chain headed by } \gamma, \text{ and } \gamma \text{ is adjoined to } H(K).
\]

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 \(\alpha\) containing a Wh-feature (what) because \(\beta\) (who) also contains the relevant Wh-feature, is closer and thus intervenes.\(^{14}\) This is exemplified in (14) where the intervening, closer element (“\(\beta\)” from above) is marked in boldface:

\(^{13}\) Note that Chomsky (1998) 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.

\(^{14}\) 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, 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. For reasons of space, I cannot treat either here.
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 versus *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 Effects

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, also, exceptional conditions do not render some constructions ungrammatical in English.

---

15 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.
One of these exceptional conditions is the following.\(^ {16}\) If the bare Wh-phrases (who, what, whom 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 in this paper; 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. 17), 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?  
b. How did Martin kiss whom?

(23) a. Wen hat Martin warum geküßt?  
b. Warum hat Martin wen geküßt?

\(^ {16}\) 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.
(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.\footnote{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).}

(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) does not exist either (cf. (10), here repeated as (29)).

\footnote{(i) Er hat Maria erst aus keinem besonderen Grund sanft auf die Wange geküßt, sie dann aus Zorn ins Ohr gebissen und ihr schließlich aus Zuneigung einen Zungenkuss gegeben
   \textit{he has Mary first out no particular reason gently on the cheek kissed, her then out anger 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.}
(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.\(^\text{18}\)

I have presented data from English and German Multiple Wh-constructions. The set of data so far suggests certain asymmetries between the two languages as well as asymmetries arising among several constructions in English. The remainder of the paper will deal with the interpretation of Multiple Wh and the syntactic processes involved culminating in an alternative proposal in the narrow sense and to some degree its application from a more broader, cross-linguistic perspective.

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).

\(^\text{18}\) 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.
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.\(^{19}\) 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). The observation to take with us from this example is that NP-movement prior to Wh-movement may have an effect on the application of the MLC and hence the relevance of the Superiority Condition.

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

---

\(^{19}\) For present purposes, it is immaterial whether this position exists, or what the status of vacuous movement is.
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). 20

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

(35) a. Dare-ga nani-o tabeta no?
\[\textit{who-NOM what-ACC ate \textit{Q}}\]
‘Who ate what?’

b. Nani-o dare-ga tabeta no?

(36) a. John-ga dare-ni [Mary-ga nani-o tabeta to] itta no?
\[\textit{John-NOM who-DAT Mary-NOM what-ACC ate COMP said Q}\]
‘Who did John tell that Mary ate what?’

b. ?? Nani-o John-ga dare-ni [Mary-ga \textit{t} tabeta to] itta no?

c. Pizza-o John-ga dare-ni [Mary-ga \textit{t} 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 \textit{in situ} and the Wh-feature thus weak.

---

20 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, the following (Akemi Masuya, personal communication) is unexpected:

(i) a. Dare-ga Mary-ni nani-o ageta no?
\[\textit{Who-NOM Mary-DAT what-ACC gave \textit{Q}}\]
‘Who gave Mary what?’

b. Dare-ga nani-o Mary-ni ageta no?

c. * Nani-o Mary-ni darega ageta no?
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.\(^{21}\) The upshot of his discussion is that the presence of an AgrOP may have further consequences.

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):

\[(37)\]
\[\begin{align*}
a. & \text{Mi kana ma?} \\
& \text{who bought what} \\
& \text{‘Who bought what?’} \\
\end{align*}\]
\[\begin{align*}
b. & \text{Ma kana mi?} \\
\end{align*}\]

\[(38)\]
\[\begin{align*}
a. & \text{Quién dijo qué?} \\
& \text{who said what} \\
& \text{‘Who said what?’} \\
\end{align*}\]
\[\begin{align*}
b. & \text{Qué dijo quién?} \\
\end{align*}\]

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 \textit{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 to check the relevant features, while a Wh-object moves through AgrOP.\(^{22}\)

Boskovic (1993:6) offers the following derivations for (37b) and (38b), respectively:

---

\(^{21}\) Note that in the published version, 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 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.

\(^{22}\) 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 1998 and other works), this type of operation is feasible, if not even desired.
(39)  a. \[\text{CP } ma\_kana\_k \{\text{Agr}OP \_t \{\text{VP } \_t \_t \text{mi}\}\}\]
    b. \[\text{CP } que\_\_dijo\_\_t \{\text{Agr}OP \_t \{\text{VP } \_t \_t \text{quién}\}\}\]

It is the lower Wh-element that moves. As it has to check its D-features (such as Case) on its way (see fn. 22), 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.  

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 1993, 1997a and Grohmann 1996, 1997d for overview and literature). 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:

(40)  a. Der Martin hat der Maria das Buch nicht \{\text{VP } \_t \_t gegeben \_t \_t\} 
      \text{‘Martin didn’t give the book to Mary’}
    b. Der Martin hat der Maria ein (bestimmtes) Buch nicht \{\text{VP } gegeben\}

(41)  a. * Der Martin hat der Maria nicht \{\text{VP } das Buch gegeben\}
    b. * Der Martin hat der Maria nicht \{\text{VP ein (bestimmtes) Buch gegeben}\}

(42)  a. Der Martin hat der Maria nicht \{\text{VP (irgend)ein Buch/was gegeben}\}
    b. * Der Martin hat der Maria (irgend)ein Buch/was nicht \{\text{VP gegeben}\}

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

\[\text{I will not discuss further whether this only applies to definite arguments (Diesing 1992) or specific ones (Moltmann 1991), though the latter seems to be on the right track (viz. (40b)); furthermore, as we will see below, specificity is also involved in Multiple Wh. (Haiden (1995) presents good evidence that all arguments leave VP.)}\]
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 versus following it.\(^{25}\) (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.\(^{26,27}\) This yields the following structural configurations:

\[
\begin{align*}
(43) \quad a. \quad & \left[ \text{AgrSP} \right. \\
\quad & \left. \text{Der Martin} \right] \text{ hat } \left[ \text{AgrIOP} \right. \\
\quad & \left. \text{der Maria} \right] \text{ nicht } \left[ \text{AgrDOP} \right. \\
\quad & \left. \text{das Buch} \right] \text{ gegeben } \left[ \text{VP} \right. \\
\quad & \left. \text{t} \right] \text{ t } \text{ t } \text{ t } \text{ t } \\
\text{b.} \quad & \left[ \text{AgrSP} \right. \\
\quad & \left. \text{Der Martin} \right] \text{ hat } \left[ \text{AgrDOP} \right. \\
\quad & \left. \text{das Buch} \right] \text{ [AgrIOP} \left. \text{der Maria} \right] \text{ nicht } \left[ \text{VP} \right. \\
\quad & \left. \text{t} \right] \text{ gegeben } \left[ \text{t} \right. \text{ t } \text{ t } \text{ t } \text{ t } \\
\end{align*}
\]

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

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 preparation). This relies crucially on the assumption that the subject in

\[\text{(i) Der Martin hat der Maria kein Buch/nichts gegeben}\]

\[\text{'Martin didn't give Mary a book/anything'}\]

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).

\(^{25}\) As a matter of fact, indefinites under negation morphologically amalgamate *nicht* and *ein*:

\[\text{(i) Der Martin hat der Maria kein Buch/nichts gegeben}\]

\[\text{'Martin has the Mary \_ book/\_ nothing given'}\]

\[\text{‘Martin didn't give Mary a book/anything’}\]

\(^{26}\) 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 preparation).

\(^{27}\) 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. A rich body of literature exists about the nature of scrambling and the role of base-generation versus 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.
German sits in a unique position in unmarked clauses (SpecAgrSP in the present configurations), an assumption that seems to be borne out (Grohmann, in preparation).

The subject is arguably in its assigned position, SpecAgrSP. 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 ever since Pollock 1989. 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. 31), the following are the major projections including their intrinsic features, immediately relevant for present purposes:


All arguments leave VP (where they are θ-marked) and move to the specifiers of the respective Agr-projections (for Case-checking); 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.

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 other adverbs) for instances of pre-subject scrambling:

(45) a. [TopP Das Buch hat AgrSP der Martin AgrDOP t AgrIOP der Maria gegeben]]]
    b. [TopP Der Maria hat TopP das Buch AgrSP der Martin AgrDOP t AgrIOP t gegeben]]]
    c. [TopP Das Buch hat TopP der Maria AgrSP der Martin AgrIOP t AgrDOP t gegeben]]]

---

28 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). I will adopt the existence of AgrP for expository purposes.

29 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.

30 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 projection later.

31 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.
With freely ordered AgrOPs, both objects satisfy the MLC in movement out of VP. 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. It is also shown that the same process takes place in embedded contexts (see Grohmann 1996, 1997d).

(46) a. \[
\text{[TopP Gestern} \_ \text{hat [AgrSP der Martin} \_ \text{das Buch der Maria gegeben]]}
\]
\[
\text{yesterday has the Martin the book the Mary given}
\]
‘Yesterday, Martin gave the book to Mary’

b. \[
\text{Er sagte, [ForceP dass [TopP die Maria} \_ \text{[AgrSP der Martin} \_ \text{geliebt hat]]]]}
\]
\[
\text{he said that the Mary the Martin loved has}
\]
‘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. 33); in (46b) the fronted object moves there.

### 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.

---

32 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).

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

34 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 versus complementizers might be.
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 examples from embedded Multiple Wh come from Grohmann 1997b:101:

(47) a. Ich frage mich \[ CP \text{ wer} C^0 [AgrSP \text{ ti} AgrOP \text{ was gekauft hat}] \]
   I ask myself who what bought has
   ‘I wonder who bought what’

b. Ich frage mich \[ CP \text{ was} C^0 [TopP ti AgrSP \text{ wer} AgrOP \text{ ti} gekauft hat] \]

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, \textit{was} targets SpecTopP to check off a Top-feature. In this position, above the subject, only \textit{was} can satisfy the MLC in the following Wh-movement.

One prediction, then, 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 when endowed with an additional Wh-feature. The consequence is that prior to Wh-movement, the particular element is situated in the specifier of either TopP or AgrP.\(^\text{36}\)

(48) Ich frage mich…

a. \[ \ldots [CP \text{ wem} [TopP \text{ ti} AgrSP \text{ wer} AgrOP \text{ ti} AgrOP \text{ das Buch gegeben hat}] ] ]\]

b. \[ \ldots [CP \text{ wem} [TopP \text{ ti} AgrSP \text{ der Hans} AgrOP \text{ ti} AgrOP \text{ was gegeben hat}] ] ]\]

c. \[ \ldots [CP \text{ wem} [TopP \text{ ti} TopP \text{ was} AgrSP \text{ der Hans} AgrOP \text{ ti} AgrOP \text{ gegeben hat}] ] ]\]

d. \[ \ldots [CP \text{ wem} [TopP \text{ ti} TopP \text{ das Buch} AgrSP \text{ wer} AgrOP \text{ ti} AgrOP \text{ gegeben hat}] ] ]\]

\(^{35}\) 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.

\(^{36}\) 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.
(49) Ich frage mich…
   a. … \[CP \text{was}_1 \text{TopP}_1 \text{AgrSP}_1 \text{wer}_1 \text{AgrOP}_1 \text{der Maria gegeben hat}]\]
   b. … \[CP \text{was}_1 \text{TopP}_1 \text{AgrSP}_1 \text{der Hans}_1 \text{AgrOP}_1 \text{wem gegeben hat}]\]
   c. … \[CP \text{was}_1 \text{TopP}_1 \text{TopP}_2 \text{wem}_2 \text{AgrSP}_2 \text{der Hans}_2 \text{AgrOP}_2 \text{t}_2 \text{gegeben hat}]\]
   d. … \[CP \text{was}_1 \text{TopP}_1 \text{TopP}_2 \text{der Maria}_2 \text{AgrSP}_2 \text{wer}_2 \text{AgrOP}_2 \text{t}_2 \text{gegeben hat}]\]

These simplified representations (Grohmann 1997:102)—indicating only the relevant projections and traces—show the general approach, namely 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 (henceforth DRQ) 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.37

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:

(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

37 The proposal advanced here, based on the following discussion, differs in some details from Grohmann, in press.
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 (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 been 
established in the discourse.\textsuperscript{38} 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 \textit{which N(s)} and \textit{which of the Ns} are inherently D-linked.\textsuperscript{39} This accounts for the 
grammaticality of (50-51) on the one hand and the grammaticality difference in (1a) versus (1b) on 
the other; \textit{who}, \textit{what} etc. are not—and in general, cannot be\textsuperscript{40}—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-\textit{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 \textit{what} did \textit{who} break? 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 \textit{where what} goes!

\textsuperscript{38} Pesetsky (1987:107f.) formulates it as “\textit{[w]hen a speaker asks a question like \textit{Which book did you read?}, the} 
range of felicitous answers is limited by a set of books both speaker and hearer have in mind.” Comorovski (1996:2) 
takes this as the basis for her elaboration when calling \textit{“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.”

\textsuperscript{39} It is a widely held view that D-linking and partitivity are alike in their semantic (and ultimately syntactic) 
properties. I will not explicitly discuss this here but refer the interested reader to the literature (see e.g., Kiss 1993, 

\textsuperscript{40} D-linking of bare Wh-phrases is possible under heavy stress, for instance, or in even more restricted context. See 
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).

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.

(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. This carries over to bare Wh:

---

41 The other example Pesetsky (op. cit.: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 what did who (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.

42 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 (op. cit.:85), however, presents the following judgements to show that neither Bolinger nor Pesetsky were correct with their accounts:

(i)  a. ? Which book did how many people buy?
    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.
(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.\(^{43}\)

(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?

To summarize our discussion of English multiple interrogatives, context may obviate Superiority effects by D-linking Wh-phrases; to which extend it may is yet unclear. After a thorough discussion of the situation in German, we will be able to describe English Multiple Wh more accurately which will also lead us to a more satisfactory account of the facts.

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.\(^{44}\)

\(^{43}\) 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*).

\(^{44}\) 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-
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:

(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.\(^45\) 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.\(^46\)

Compare Situation I with another context, given in (60):

(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).
(60) Situation II: Same author, same books but this time he informs his boy-friend 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); 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 complete sets of referents are given.

(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 an exhaustive list of 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 exhaustive sets of referents.

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. 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.

The other contexts given above for English concern (58) (and (i) in fn. 45), 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?

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-arguments. (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 that day 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.”

---

48 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, die Tassen und die Schalen zerbrochen haben; aber…
   … I also know also that Peter, Paul and Mary the plates, the cups and the bowls broken have; but…

49 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.
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 acceptability/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 (standardly assumed ever 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.\(^{50}\)

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?

\(^{50}\) To come back to the remark from fn. 9 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.51

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:

---

51 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. The fact that topicalization in embedded contexts is not as well-formed as in matrix clauses (at least in English) may bear on this.
(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.\textsuperscript{52}

(70) Situation: Martin’s aunt informs her friend of Martin’s 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?

(71) Situation: His aunt informs her friend of Martin’s 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.

\textsuperscript{52} It is irrelevant here whether it is an adjunct, an argument or a quasi-argument.
(72) Situation: Someone informs someone else of Norbert’s lending out *The Minimalist Program*.
“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 Norbert’s lending out of the book includes his 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. 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.

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.

---

53 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.
54 See, for example, Kiss 1993 who uses this test to distinguish properties of Multiple Wh in English versus
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.)
55 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
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:

(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 this is crucial 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.\(^{56}\)

(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 sets has been established, nothing goes (DRQ). Unlike English, German allows this special kind of D-linking for all Wh-elements.

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 Wh-topics, and (iii) consider the different syntactic realization of semantic and/or pragmatic factors across languages.

\(^{56}\) 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.
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 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 unselectively? It would be nice if we could pin-point the special discourse-properties down 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). 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?

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.
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 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—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 evidenced 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. \([_{CP} \text{Which student}_{i} \left[_{IP} t_{j} \left[_{VP} t_{j} \text{read which book}\right]\right]]\) |
    b. \([_{CP} \text{Which book}_{j} \left[_{IP} \text{which student}_{i} \left[_{VP} t_{j} \text{read } t_{j}\right]\right]]\) |

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”;

58 Kuno (op. cit.: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) and discussion).
this special case of D-linking is, I argue, topicalization which allows movement for feature-checking purposes, yet keeping up the idea that 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 for English versus German.

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. But this is not enough; specificity alone cannot distinguish successfully the difference of D-linked Wh in English and DRQed Wh in German.

Specificity of certain Wh-elements implies different a 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 internal structure of a partitive DP:\footnote{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 (in press) and Muromatsu (in progress).}

\begin{enumerate}
\item \[\text{DP many Q [NP students]}\]
\item \[\text{DP many Q [Part of those [NP students]]}\]
\item \[\text{DP which Wh [NP student(s)]}\]
\item \[\text{DP which Wh [Part of those [NP students]]}\]
\end{enumerate}

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:
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.\(^60\) 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 first, 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.

\(^{60}\) 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.
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 going on, let me briefly mention three cases in support of this view, one of which we have seen already (with 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 to 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. 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. 44 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 presumably inside the CP-domain (following Rizzi (1997)) and the appropriate structural configuration looks as follows:

---

61 Please bear with me. This section presupposes the existence of Wh-topics and their presence in German Multiple Wh. I will lay out their syntax here, then expand on its problems and consequences. In section 3.3.2, I will show that one objection to this approach can be refuted in its own terms (Wiltschko 1997). Section 4.1.1 will contain a convincing discussion that Wh-topics exist (Wu 1996). The presentations following will support the crucial role that topicalization plays in a range of constructions.

The upshot is that while the direct positional evidence for both Wh-elements to sit in TopP may not be as strong as one could hope for, consequences of this approach are. I hope to return to the former in future research, while lending even more support to the latter issue as well. Please remember that the present lines are inquiries in the strictest sense.
The topic-feature which is part of the feature-bundle of XP (the topic phrase) must be checked off against the topic-feature of the head; this is indicated in (83) by strikethrough typeface. As shown by Rizzi, languages may employ more than one topic per sentence (thus having multiple topic projections, indicated as 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 discussion and the collection of papers in Anagnostopoulou et al. 1997).

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):

\[(85) \quad \text{[Wh [Top]]}\]

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 C0.

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

---

62 For present purposes, further information such as the instantiation of topic-comment structure, or even focus-background will be irrelevant here; for proposals, see among many others Vallduví 1990.
As already laid out in section 2.3.3, “scrambling” results in argument-movement to the specifiers of AgrP. This 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]. With three possible candidates (der Martin, wem, was) this can only concern either Wh: both are marked with a Wh- and a Top-feature. 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 enough. But the connection to scrambling arises as a corollary: scrambling over the subject is independently taken to be topicalization (following Grohmann (1996, 1997d)).

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. \[
\begin{array}{c}
\text{[CP Wem hat [TopP \{TopP was \{AgrSP der Martin \{AgrDOP \{AgrIOP \text{gegeben}\}\}\}\}\}\]}]
\end{array}
\]

b. \[
\begin{array}{c}
\text{[CP Was hat [TopP \{TopP wem \{AgrSP der Martin \{AgrDOP \{AgrIOP \text{gegeben}\}\}\}\}\]}]
\end{array}
\]

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 are freely ordered. Once the subject has moved to AgrSP, the functional head Top⁰ is introduced in the derivation which attracts the closest element bearing a Top-feature (MLC), which in the case of (87a) is *was*; the Wh thus moves to SpecTopP. Next, another Top⁰ 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 D-linking variant of Multiple Wh in German.

---

64 On the other hand, all arguments leave VP and thus “scramble”—a further argument why scrambling alone cannot be enough; I will return to this below.
Wh-elements in German Multiple Wh are Wh-topics and undergo a derivation in accordance to 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.65 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) \[
\begin{array}{cc}
\text{CP} & \\
\text{DP} & \text{C'} \\
\{\text{Wh}\} & \\
\text{C^0} & \text{TopP} \\
\{\text{Wh}\} & \\
\text{t/t_i} & \text{Top'} \\
\{\text{Top}\} & \\
\text{Top^0} & \text{IP} \\
\{\text{Top}\} & \\
\text{t_i/DP_i} & \text{I'} \\
\text{I^0} & \text{VP} \\
\text{t_j} & \text{V'} \\
\text{V^0} & \text{DP/t_j} \\
\end{array}
\]

On analogy with the previous derivation, (89) represents the derivations for (88).66 The partitive Wh needs to check both, the Wh-feature and the Top-feature. The generalization, then, is that all

65 This is just one way of interpreting the facts, i.e. that the properties of the clausal type (multiple interrogation, either as D-linking in some contexts or as DRQ in all contexts) directly influences the properties of the Wh-elements. Let us stick to this picture of the interrogative world for expository purposes.

66 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 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.₆⁷

(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. 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 is the Wh-phrase from the lowest AgrP that moves to the highest TopP. This is what we have seen in (87) already.₆₈

---

₆⁷ 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:-)

₆₈ Recall that AgrOP is assumed to be ordered freely. The MLC, then, straightforwardly forces the higher object to move to the lower TopP (where the lower Top⁰ is introduced first, attracting the element in the higher AgrOP).
(91) a. \[
\text{CP } \text{Wen}\hat{\text{j}} \text{ hat } [\text{TopP } t_j [\text{TopP } \text{wem}\hat{\text{i}} [\text{AgrSP } \text{der Martin } [\text{AgrIOP } t_i [\text{AgrDOP } t_j \text{ vorgestellt}]]]]]?
\]
\text{‘Whom did Martin introduce to who yesterday?’}

b. \[
\text{CP } \text{Wen}\hat{\text{j}} \text{ hat } [\text{TopP } t_j [\text{TopP } \text{wem}\hat{\text{i}} [\text{AgrSP } \text{der Martin } [\text{AgrDOP } t_i [\text{AgrIOP } t_j \text{ vorgestellt}]]]]]?
\]

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).

Here, 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 (23a), repeated as (92a), and an alternative to (91) as in (92b):

(92) a. \[
\text{Wen hat Martin warum geküßt?}
\]

b. \[
\text{Wen hat Martin wem vorgestellt?}
\]

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); if the subject is not tied to one structural position (e.g., AgrSP), then it is at least with respect to its relation to the other arguments (Grohmann, in preparation). But in cases such as (92), it is plausible that the subject moves further, once the relevant features are checked by AgrS\textsuperscript{0}.

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: following Zwart (1997a and earlier work), 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

---

\textsuperscript{69} 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 1997 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.
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).

However, I propose that this movement is independent from movement to TopP of other elements; the two types do not interfere. If they would, a basic Multiple Wh such as (2a), repeated as (93), could not be derived:

\[(93) \left[\text{CP} \, \text{Wer} \, \text{i} \, \text{hat} \, \left[\text{TopP} \, \text{t} \, \text{i} \, \text{TopP} \, \text{wen} \, \left[\text{AgrSP} \, \text{t} \, \text{i} \, \text{AgrOP} \, \text{t} \, \text{geküßt}\right]\right]\right]\]

If both moves would underlie basically the same operation in (93), wer had to move first in order to satisfy the MLC and the NDC. In this case, (93) could never be derived. I thus propose that the MLC is relevant to attract wen to SpecTopP. After this operation, the subject raises in order to check its thematically dominant feature at a higher position.

Given these premises, then, (92) receives the following representation:

\[(94) \begin{align*} 
\text{a. } & \left[\text{CP} \, \text{Wen} \, \text{i} \, \text{hat} \, \left[\text{TopP} \, \text{t} \, \text{i} \, \text{TopP} \, \text{Martin} \, \text{t} \, \text{TopP} \, \text{warum}\,(?) \, \left[\text{AgrSP} \, \text{t} \, \text{k} \, \text{AgrOP} \, \text{t} \, \text{geküßt}\right]\right]\right] \\
\text{b. } & \left[\text{CP} \, \text{Wen} \, \text{j} \, \text{hat} \, \left[\text{TopP} \, \text{t} \, \text{j} \, \text{TopP} \, \text{Martin} \, \text{t} \, \text{TopP} \, \text{wem}\text{\textemdash} \left[\text{AgrSP} \, \text{t} \, \text{k} \, \text{AgrDOP} \, \text{t} \, \text{vorgestellt}\right]\right]\right] 
\end{align*}\]

In summary, 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 or not.

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:

\footnote{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.

\footnote{The reason for the bracketed ‘?’-index on the adjunct \textit{warum} is that I can only remain agnostic about the role that adjuncts play in a derivation (cf. fn. 33): are they generated where they appear, maybe even inserted once all other movement has applied? Or are they generated in a specific position and undergo movement?}
(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, for example), 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 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: 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.73

72 See fn. 61 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) and the range of data this approach covers.

73 Why non-interrogative was can only be non-specific, see Haiden 1995 or Postma 1995, among many others.
(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 (“existential”) *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.

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?

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

   b. Welcher Schüler hat was richtig lösen können?

---

74 Diesing (1992) proposes VP as the invariant domain for existential interpretation, while Postma (1995) suggests to link it to an existential operator.
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, 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.76

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-

75 Presumably only with partitive Wh-elements, though this is not crucial.
76 See also fnn. 61 and 72.
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 (which is not relevant for this study).\footnote{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\textsuperscript{0} attracts is *wer zum Teufel and *wer, respectively.}

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

(105) a. * Wer zum Teufel hat wen zum Teufel gesehen?
   \textit{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?

Aggressively non-D-linked Wh-elements must move to SpecCP overtly; German C\textsuperscript{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 \textit{where} and \textit{when} versus \textit{why} and \textit{how}.\footnote{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 \textit{why/how} are still pretty much a mystery.} Only the former may be preceded by an overt preposition such as \textit{from where, since when}. We can schematize this as follows with the null preposition represented as \textit{p} (from Kitahara 1997:101):

(106) a. [\textit{pp p where/when}]
   
   b. * [\textit{pp p why/how}]

Huang uses this fact to propose that \textit{where/when} pattern like \textit{why/how} in the syntax, while they behave like \textit{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.79

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 versus 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 versus where/when do not act uniformly due to the fundamental differences in the nature of movement. This might be a clue that why and how are generated in a specific position, rather moving there.80,81

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

(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

79 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.
80 Reinhart (1995 and earlier work) argues for this and identifies the position uniformly as SpecCP.
81 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 possibility further.
We will see an instance of TopP preceding CP in Bulgarian further below.
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 (adverbal) 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. 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, left dislocation (and hanging topics) presently.

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

(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 die 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

---

82 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.

83 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.
the introduction—van Riemsdijk (1997)—for the validity of this approach), 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.

The following paradigm and the corresponding simplified representations show this:

(112) a. Diesen Mann, hat Maria geküßt
   this-ACC man, has Mary kissed
   ‘This man, Mary kissed’

b. Diesen Mann, den hat Maria geküßt
   this-ACC man, the-ACC has Mary kissed
   ‘This man, Mary kissed him’

c. Dieser Mann, den hat Maria 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.84

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).

(113) a. This man, why did Mary kiss him?

b. * This man, why did Mary kiss?

---

84 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.
(114) a. Diesen Mann, warum hat Maria den geküßt?\(^85\)
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-created left dislocation. These show also that a left-peripheral position blocks movement.\(^86\)

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 thus that hypothesis from (107) can be retained and these restriction account for the differences observed by Kuno and Takami.\(^87\)

### 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 (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.\(^88\)

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

---

\(^85\) With the Wh in CP, V2 has not to be saturated in any special form. 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.

\(^86\) 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, as pointed out to me by Norbert Hornstein:

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

\(^87\) 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.

\(^88\) 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.
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:

(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.

---

89 I cannot indulge here in a deeper discussion of the exact reason for this restriction. Only some quantifiers can be topicalized (maybe along the weak/strong-dichotomy of quantifiers following Milsark (1977) and much subsequent work), and for the present purpose it suffices to note that some quantifiers such as nobody cannot topicalized.
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?

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 of the world 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 Karl may freely topicalize, the quantified object niemanden 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 in between topicalize in the left periphery.

On the other hand, comparable constructions to (115) with a “strong” quantifier (cf. fn. 89) in between the two Wh-elements are predicted to be fine: such a quantified expression should be able to topicalize.90

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

90 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.
(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,\(^{91}\) in these cases, the subject is a strong 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 weak quantified subject cannot do so (122a), unlike a non-quantified one (122b).

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

(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?

---

\(^{91}\) 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.
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 AgrP just fine. In (123b) TopP above the subject is filled, forcing illegitimate movement of niemand to TopP.

As predicted, (122) is grammatical with a strong quantifier as the subject, and so is (123).  

\[(124)\]
\[
a. \text{Wen hat jede Studentin von den Musikern getroffen?} \\
b. \text{Was hat jeder für Bücher gelesen?}
\]

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

The approach as just outlined may 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 for Wh-Topics

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.

\[92\] For the sake of completeness, the version of (123b) without a quantified subject is also grammatical:

\[(i) \text{Was hat Luise für Bücher gelesen?}\]
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. This is considered the standard approach to Chinese interrogatives ever since Huang 1982.

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 as being related. The general idea is very much in line with the present 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 shei mai le ti?
   what who buy-ASP
   *What did who buy?’

b. [Shenme dongxi] ni quan shei bu yao mai ti?
   what thing you persuade who not buy
   *What did you persuade who not to buy?’

---

93 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.
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.\footnote{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 (op. cit.: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.} 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):

\begin{equation}
\text{(127) Situation: Zhangsan went to the supermarket to do some shopping.}
\begin{align*}
a. \quad & \text{Zhangsan mai-le shenme?} \\
& \text{Zhangsan buy ASP what} \\
& \text{‘What did Zhangsan buy?’}
\end{align*}
\end{equation}

\begin{equation}
b. \quad & \text{Shenme Zhangsan mai-le?}
\end{equation}

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.)

\begin{equation}
\text{(128) Situation: Zhangsan went to the supermarket and bought items from a mutually known list.}
\text{Shenme Zhangsan mai-le?}
\end{equation}

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. 94 on
interpretation). This might again be related to the general absence of morphological features in Chinese driving overt movement.\textsuperscript{95}

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, does his, mother like t\textsubscript{i}?
    b. John, his, mother likes t\textsubscript{i}.

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

(130) a. Shei, ta, de muqing hen xihuan t\textsubscript{i}?
    \textit{who he DE mother very like}
    \textquoteleft*Who does his mother like?	extquoteright
    b. * Ta, de muqing hen xihuan shei,?

Wu (1996:179) shows that the same contrast can be observed in Chinese: if the Wh stays \textit{in situ}, weak crossover effects can be found (130b), while they 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: (where irrelevant details are left out)

\textsuperscript{95} While the Wh-subject is unlikely to be topicalized in (127a), it is interesting to note that if a non-interrogative adverb is used, it should appear after the subject, not in between the two Wh-elements (Jianxin Wu, personal communication). It might thus be argued that nothing may intervene between the topicalized Wh and the subject—lending further support to the inherent topichood of the subject.
Overt movement to TopP satisfies [Top], subsequent movement to CP will occur at LF. Note the crucial difference to German: 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. 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, as well as the Japanese context-restriction shown below, indicate that the present approach hits upon something novel that is worth pursuing, even from a wider cross-linguistic perspective.

### 4.1.2 Japanese Wh and the Topic Marker \textit{wa}

As shown by Miyagawa (1987), the Japanese topic marker \textit{wa} may under certain circumstances also appear in interrogative contexts. This has not been widely noticed, for exactly the reason that

\textit{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.}
these circumstances are very restricted. But it shows that Wh and topicalization are not mutually exclusive by nature.\footnote{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).}

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:

\begin{enumerate}
\item \textit{Dare-wa kita no?} \quad who-TOP came Q
\end{enumerate}

‘Who came?’

\begin{enumerate}
\item \textit{Dare-wa kîte, dare-wa konakatta no?} \quad who-TOP come-GER who-TOP didn’t-come Q
\end{enumerate}

‘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 of this is, again, that a connection between Wh-formation 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. Further research will have to take discourse restrictions into account as I have shown for German.

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.\footnote{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 the interesting observations from ASL to my attention.} 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.
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’

While a bare Wh may not be extracted out of a Wh-island (133a), 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.99 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?

---

99 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).
(134) a. Shenme dongxi Zhangsan xiang zhidao Lisi mai mei mai ti?
   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 ti 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.\(^{100}\)

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

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 for by Rudin (1985) and is mentioned in Richards 1997:111f. who cites the following:

(135) a. \[
\text{[TopP Ivan vcera [CP kakvo kupi]]?}
\quad \text{Ivan yesterday what bought}
\]
   ‘What did Ivan buy yesterday?’

b. \[
\text{[TopP Vcera Ivan [CP kakvo kupi]]?}
\]

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

(136) a. * \[
\text{[TopP Koj vcera [CP kakvo kupi]]?}
\quad \text{who yesterday what bought}
\]
   ‘Who bought what yesterday?’

b. \[
\text{[TopP Koja zena vcera [CP koja kniga kupi]]?}
\quad \text{which woman yesterday which book bought}
\]
   ‘Which woman bought which book yesterday?’

---

\(^{100}\) 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.
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):

(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 \text{ Wh}_1 [... [IP \text{ ... [CP \text{ Wh}_3 \text{ ... [IP \text{ ... [t_1]}]}]}]\]

   b. \([\text{TopP \text{ Wh}_1 [... [IP \text{ ... ([TopP \text{ t_1} ... [CP \text{ Wh}_3 \text{ ... ([TopP \text{ t_1} ... [IP \text{ ... [t_1]}]}]}]}]}]\]

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 features and movement allows for multiple CP-specifiers where every Wh-phrase is tucked in after the first

---

101 The seminal work on Slavic Multiple Wh is Rudin 1988.
102 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).
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

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, do you wonder who to give $t$ to?
   b. ? How much of the money, do you wonder who to give $t$ 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:

(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 d’Avis 1996 for discussion and references).

(141) a. * Was weißt du, wer gelesen hat?
   \[\text{what know you who read has}\]
   ‘What do you know who read?’
   b. * Welches Buch weißt du welcher Student gelesen hat?

---

103 Space does not allow for a more accurate discussion of when Wh-extraction is allowed in English.
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 TopP of the matrix clause. 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. 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 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, 1998 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:

---

104 I have not looked at the possibility of the ordering [TopP Top0 [CP C0…]] in Swedish or in Chinese rather than only allowing [CP C0 [TopP Top0 …]]; both might be feasible, however, given the following discussion.
(143) a. Welche Bücher weißt du, die welche Studenten gekauft haben?

‘Which books do you know that which students bought them?’

b. * Welche Bücher weißt du, welche Studenten gekauft haben?

c. * Welche Bücher weißt du, die haben welche Studenten gekauft?

d. * Diese Bücher weiß ich, die 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.¹⁰⁵

¹⁰⁵ 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.
The overall construction 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, \textit{welche Studenten}.\(^\text{106}\) 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 \textit{welche Studenten} has already checked [Top], \textit{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

\(^{106}\) 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) \textit{Welche Studenten weißt du, die welche Bücher gekauft haben?}
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.\textsuperscript{107}

We have now evidence that even German allows Wh-extraction out of Wh-islands.\textsuperscript{108} 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 turn 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.\textsuperscript{109}

\textsuperscript{107} 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.

\textsuperscript{108} 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.

\textsuperscript{109} 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:

(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.

---

110 The ungrammaticality of (150) also shows that [Top] must be checked prior to [Wh].
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.\footnote{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.}

As the data have shown—and especially the discussion of German—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\textsubscript{0} that forces further movement.

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

\begin{align}
(152) \ a. \ & \text{[CP Wh}_i \ [\ldots \text{IP} \ldots \text{[TopP t}_i \ldots \text{CP Wh} \ldots \text{IP} \ldots t_i \ldots]]] \\
& \text{b.} ^* \ [\text{CP Wh}_i \ [\ldots \text{IP} \ldots \text{CP Wh} \ldots \text{IP} \ldots t_i \ldots]]] \\
& \text{c.} ^* \ [\text{TopP Wh}_i \ [\ldots \text{IP} \ldots \text{TopP t}_i \ldots \text{CP Wh} \ldots \text{IP} \ldots t_i \ldots]]] \\
& \text{d.} ^* \ [\text{CP Wh}_i \ [\ldots \text{IP} \ldots \text{CP Wh} \ldots \text{TopP t}_i \ldots \text{IP} \ldots t_i \ldots]]] \\
& \text{e.} ^* \ [\text{TopP Wh}_i \ [\ldots \text{IP} \ldots \text{CP Wh} \ldots \text{TopP t}_i \ldots \text{IP} \ldots t_i \ldots]]] \\
\end{align}

The only acceptable 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.\footnote{This project allows only for a very limited and crude look into discourse restrictions across languages.} 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.\textsuperscript{113}

Consider thus the following situations (verified by Jianxin Wu, personal communication):\textsuperscript{114}

(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:

\textsuperscript{113} 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.

\textsuperscript{114} Here and in the remainder of the section, the questions in all languages are translated as \textit{Who bought what?} in the a- and c-examples, and \textit{What did who buy} in the b- and d-examples, similar in essence to (59).
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. 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.¹¹⁵

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

(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?

¹¹⁵ In this respect, they pattern much like English.
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 here in that they allow the following:

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

Interestingly, if only the reference to the Wh-subject has been introduced, a Multiple Wh is well-formed if the Wh-object is fronted.

4.3.3 English and the Comparison
The two classes as presented above both differ from English in several respects. 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,

---

116 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 are).
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 known, 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

<table>
<thead>
<tr>
<th>Given:</th>
<th>Wh-SU… Wh-OB</th>
<th>Wh-OB… Wh-SU</th>
</tr>
</thead>
<tbody>
<tr>
<td>Referent of Wh-SU</td>
<td>Class III</td>
<td>Class II</td>
</tr>
<tr>
<td>Referent of Wh-OB</td>
<td>Class II, Class III</td>
<td>*</td>
</tr>
<tr>
<td>Referent of both Wh</td>
<td>Class I, Class II, Class III</td>
<td>Class I, Class II</td>
</tr>
</tbody>
</table>

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. We have already seen some differences or conditions above, and the following section will evaluate possible accounts for variation based on child language and learnability issues.

4.4 Child Language and Learnability

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.

---

117 This section contains speculations and hence theory-driven considerations of the issue. I have not conducted any experiments or collected data from child language other than noted. This, again, suggests plenty of research material.
Under a cue-based approach to learnability (Lightfoot, forthcoming), 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:

(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. 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 German learner is that all these constructions can be possible evidence, as they are grammatical.

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

---

118 I refer to the discussion in sections 3.3.3 and 4.2.3 above.

119 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.
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.\textsuperscript{120} (166b) 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.\textsuperscript{121}

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-

---

\textsuperscript{120} 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 or Lightfoot, 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.

\textsuperscript{121} 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.
phrases; in other words, it will assume a more intricate structure for all Wh-phrases based on the pattern observed from partitives.\textsuperscript{122}

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):\textsuperscript{123}

\begin{quote}
(167) a. * How much money, do you wonder who to give t\textsubscript{i} to?
   b. ? How much of the money, do you wonder who to give t\textsubscript{i} to?
\end{quote}

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. 123), the English learner may get evidence to assign the partitive structure to D-linked Wh-phrases.\textsuperscript{124}

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 Lightfoot or Crain and Thornton) 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.

\textsuperscript{122} 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.

\textsuperscript{123} 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.

\textsuperscript{124} 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.
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.\footnote{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 be understood as TopP, and its position may as well be below CP; this conception is also the result from class discussions (Rozz Thornton, Spring 1998, UMCP) for which I am grateful to Larisa Avram, Elpida Bairaktari, Stephen Crain and Rozz Thornton.}

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. \[
\begin{array}{c}
[CP \ C^0 \ [TopP \ [\text{which Smurf}]_i, \ Top^0 \ [\text{the big man couldn’t pick up} \ t_i]])
\end{array}
\]

b. \[
\begin{array}{c}
[CP \ [\text{which mouse}]_i, \ C^0 \ [TopP \ [\text{what the cat didn’t see} \ t_i]])
\end{array}
\]

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 \textit{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.

\[125\]
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. 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. 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 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 native speakers, especially when the constructions get more and more intricate.

---

126 Note that the two questions were uttered by two different children.
127 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.)
5 Conclusion

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

---

128 Given the many open questions raised and pointed out throughout this study, this might actually be a misnomer; call it “preliminary summary.”
words: Wh-topics may move out of Wh-islands. The 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 nice 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.

129 (A personal note: this sentence sums up the past 82 pages most accurately.)
References


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 

at Colloquium Series, Rijksuniversiteit Groningen (January 9).

Grohmann, Kleanthes K. In Press. 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.

Grohmann, Kleanthes K. In Preparation. A Fixed Subject Position in German. Ms. (Squib), 
University of Maryland, College Park.


Haftka, Brigitta. 1995. Syntactic Positions for Topic and Contrastive Focus in the German 
Focus Workshop. SFB 340.


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.


Dissertation, Massachusetts Institute of Technology, Cambridge.


11: 85-120.

MIT Press.


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


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.


