Lowering and Cyclicity: Attraction by X from Spec XP

Norvin Richards

MIT

Among the many ungrammatical examples that we want our theory to rule out are "lowering" examples of the type in (1):

(1) *She told __ [who John left]

Assuming that examples like (1) are to be ruled out in the syntax, we could imagine a number of means of ruling them out explicitly, including conditions on operations like movement or attraction, or on the representations created by such operations. We might, for instance, state that heads may only attract objects in their c-command domain, or that chains created by movement must involve a head which c-commands its tail.

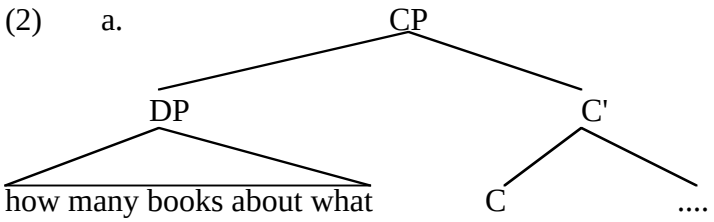
On the other hand, many examples of lowering, including the one in (1), will be ruled out for us by current approaches to cyclicity (cf. Chomsky 1995, 1998, 1999 and much other work). In these approaches, the tree is assembled from the bottom up and each attractor must do all the attracting it can as soon as it is introduced into the structure. This kind of cyclicity will rule out (1): the embedded interrogative C must do its attracting of wh-phrases at a point in the derivation at which the wh-phrase *who* does not yet exist, and the derivation crashes.

There is at least one case in which a cyclicity-based approach to the ill-formedness of (1) might differ in its predictions from an approach which posits an

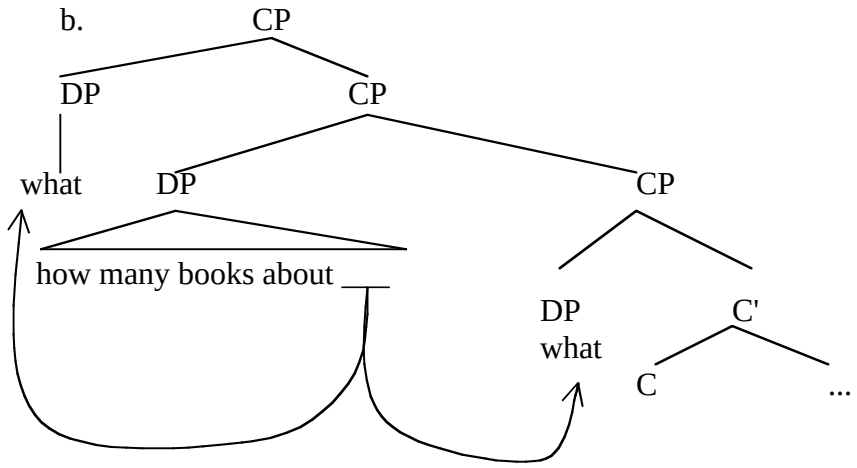
* Many thanks to Roumyana Pancheva and Marina Todorova for their Bulgarian judgments. All Bulgarian data here are from one or another or both of them, except where noted. Thanks too to David Pesetsky and to the audience at NELS 32 for helpful discussion. Any errors are my own responsibility.

explicit ban on lowering. The relevant case has to do with Attract relations between heads and objects embedded in their specifiers, as in the derivation in (2) (in (2b), to avoid prejudging the issue, I have drawn movement arrows in both of the two conceivable directions for this kind of movement):

(2) a.



b.



A theory that banned lowering--for example, by requiring heads to c-command objects that they Attract--might rule out Attract relations like the one in (2b). On the other hand, a theory based on cyclicity might allow them, depending on the exact nature of the theory of cyclicity being posited. If, for example, the requirement is that a head cannot perform Attract relations unless its maximal projection is undominated (as in Chomsky 1995), and if this is the only requirement banning lowering, then attraction by a head out of its own specifier ought to be perfectly permissible. We will see some evidence that this is the case, drawn from facts about wh-movement in Bulgarian.

Bulgarian does have multiple questions whose word order is consistent with a derivation of the type in (2):

(3) a. [Ot kakvo] [kolko gord __] beše Ivan ?
of what how proud was Ivan
'How proud of what was Ivan?'

b. [Ot koi strani] [po kolko studenta __] predstavi na Ivan ?
from which countries DIST how-many students you-introduced to Ivan
'How many students from which countries did you introduce to Ivan?'

In what follows I will try to show that Bulgarian questions like the ones in (3) (which I will refer to here as *Russian doll questions*) do have a derivation like the one in (2); for instance, that a question like (3a) involves movement of the entire wh-phrase *kolko gord*

ot kakvo 'how proud of what' to Spec CP, followed by movement of the wh-phrase *ot kakvo* from its position inside that moved wh-phrase to another Spec of CP. Of course, there are a number of other imaginable derivations for this word order. I will try to show that the wh-phrase which begins the derivation embedded in another wh-phrase (henceforth, the *embedded wh-phrase*) is required to move from its base position via wh-movement; that this movement does begin from a position inside another wh-phrase (and not, for example, from an extraposed position); that wh-movement is to an additional specifier of C (and not, for example, to a high structural position within the embedding wh-phrase); and, finally, that movement of the embedded wh-phrase takes place after the wh-phrase containing it has moved to Spec CP.

1. What type of movement is involved?

It appears that the embedded wh-phrase is undergoing wh-movement; that is, that it is undergoing obligatory overt movement of a kind which is only possible for wh-phrases. We can see this most clearly in structures where the base position of the embedded wh-phrase has phonologically overt material on both sides of it in the wh-phrase of which it is a part. In (4) below, for example, the complement PP *po matematika* 'of mathematics' must be to the right of the head noun *studenti* 'students', and to the left of the adjunct PP *ot Bulgaria* 'from Bulgaria':

(4) a. Vidja studenti **[po matematika]** [ot Bulgaria]
you-saw students of mathematics from Bulgaria
'You saw students of mathematics from Bulgaria'
b.* Vidja [**po matematika**] studenti [ot Bulgaria]
c.* Vidja studenti [ot Bulgaria] **[po matematika]**

On the other hand, if we consider the corresonding Russian doll question, the facts change dramatically. The embedded wh-phrase can no longer be in its base position, as in (5a); it must be either on the left of the embedding wh-phrase (5b) or on the right (5c).

(5) a. *Kolko studenti **[po kakvo]** [ot Bulgaria] vidja?
how-many students of what from Bulgaria you-saw
'How many students of what from Bulgaria did you see?'
b. **[po kakvo]** kolko studenti [ot Bulgaria] vidja?
c. Kolko studenti [ot Bulgaria] **[po kakvo]** vidja?

Thus, the embedded wh-phrase appears to be undergoing obligatory overt wh-movement.

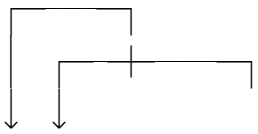
2. Where does wh-movement begin?

Even granted that wh-movement of the embedded wh-phrase is involved, however, we might wonder whether such movement actually begins from the base position of the wh-phrase, as I have proposed above. Someone might suggest, for instance, that the first step in the derivation of Russian doll questions is some kind of scrambling or extraposition of the embedded wh-phrase out of the wh-phrase containing it, after which multiple question formation can take place in the usual way.

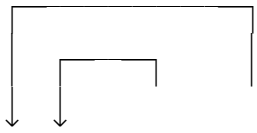
In fact, the data in (4-5) give us good reasons to reject an approach of this kind. To begin with, Bulgarian extraposition appears to be more constrained than its English counterpart; as we saw in (4), extraposition of the kind that would be needed is not ordinarily possible. Even if we were to posit a type of extraposition which only affects wh-phrases, it is not clear that the problem would then be solved. Suppose that there is an extraposition operation that applies to the embedded wh-phrase. This operation will presumably move the embedded wh-phrase to a position c-commanding the larger wh-phrase in which it was embedded. From what we know independently about Bulgarian wh-movement, we would expect such a configuration to yield the word order in (5b) (repeated as (7b)), but not the order in (5c) (repeated as (7c)):

(7) a. *Kolko studenti **[po kakvo]** [ot Bulgaria] vidja?
how-many students of what from Bulgaria you-saw
'How many students of what from Bulgaria did you see?'
b. **[po kakvo]** kolko studenti [ot Bulgaria] vidja?
c. Kolko studenti [ot Bulgaria] **[po kakvo]** vidja?

Work on multiple wh-movement in Bulgarian (cf. Rudin 1988, Bošković 1997, 1999, Richards 1997, 2001) has established that the order of the moved wh-phrases reflects the base c-command relations between them; if the base position of wh-phrase α c-commands that of wh-phrase β, then α precedes β. When the wh-phrases are a subject and an object, for instance, the subject must precede the object (Rudin 1988, 472-473):



(8) a. Koj kogo __ vižda __?
who whom sees
'Who sees whom?'



b.* Kogo koj __ vižda __?

Thus, if multiple questions like the ones in (7) involved extraposition of the embedded wh-phrase as a first step, followed by independent wh-movement of the two wh-phrases, we would expect the embedded wh-phrase to be required to precede the wh-phrase in which it was embedded. This order is possible, as we have seen, but it is not required.

There is one domain in which the constraints on ordering of multiple wh-phrases in Bulgarian is relaxed, namely that of D-linked multiple questions:

(9) a. Koj profesor koja kniga e vidjal?
which professor which book AUX seen
'Which professor saw which book?'

b. ?Koja kniga koj profesor e vidjal?

We might wonder, then, whether the questions under investigation here show some optionality of ordering because they are D-linked. D-linking is difficult enough to define precisely that it is probably impossible to completely rule out this hypothesis, but it seems unlikely. Consider, for example, the questions in (10), where the embedded wh-phrase *ot kakvo* 'of what' wh-moves either to the left or (I assume, string-vacuously; recall from (7) above that movement appears to be obligatory) to the right:

(10) a. [Ot kakvo] [kolko gord __] beše Ivan __ ?
of what how proud was Ivan
'How proud of what was Ivan?'
b. [Kolko gord __][ot kakvo] beše Ivan __ ?

However, neither of the wh-phrases in (10) seem to be D-linked; they are not ordinarily able to be freely reordered with other wh-phrases, as (11-12) show:

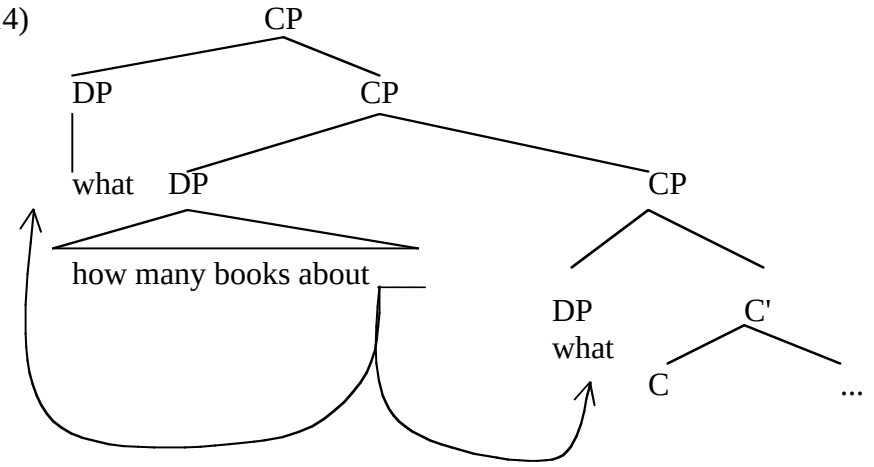
(11) a. Koj [ot kakvo] beše gord?
who of what was proud
b. *[Ot kakvo] koj beše gord?
(12) a. Koj [kolko gord ot tova] beše?
who how proud of this was
b.*[kolko gord ot tova] koj beše?

Suppose we grant, then, that the optionality exhibited in pairs like the one in (10) has nothing to do with D-linking. What does it follow from? This is a question I hope to focus on more in future work, but at this point I can offer some speculations. Suppose, first of all, that the account of Bulgarian wh-phrase ordering in Richards (1997, 2001) is essentially correct. That theory claims that each instance of wh-movement must "tuck in", landing in a specifier below all existing specifiers. I tried to get this to follow from general principles of locality; in particular, I suggested that all the participants in the movement relation (namely, the attractor, the attractee, and the new copy of the attractee created by movement) were subject to a requirement that they be maximally close to each other (which I referred to as Shortest). "Tucking in" followed from the requirement that the head of the created chain be as close as possible to the attractor--and, redundantly, from the requirement that the head and tail of the new movement chain be maximally close together. Depending on how "closeness" is defined, we could make use of this redundancy here. Suppose we define it in the following way:

(13) a. path
The path between α and β is the non-null set of nodes x such that α c-commands x and x dominates β (cf. Pesetsky 1982)
b. Shortest
The relation between α and β obeys Shortest iff there is a path π between α and β such that for any γ, $\gamma \neq \beta$, π is a subset of the path π' between α and γ.

These definitions will have the desired result of making "tucking in" optional just in the case of extraction of an embedded wh-phrase. Consider again the relevant structure, repeated here as (14):

(14)



Recall that Shortest is taken to constrain both the relation between the head of the movement chain and the attractor and the relation between the head and tail of the movement chain. In ordinary multiple-wh questions, these requirements both (with some redundancy) require tucking in. Here, however, the redundancy disappears. The Shortest condition on the distance between the attractor and the head of the chain clearly favors tucking in; if movement is to a specifier below the already occupied one, then the path between *what* and C contains only the C' node, which is a subset of any other paths which are under consideration. The Shortest condition on the distance between the head and tail of the chain, however, has the opposite effect. If *what* tucks in, then there is no path consisting of nodes that *what* c-commands and which dominate its trace; Shortest is therefore disobeyed. Since the two constraints are in conflict, either option is possible.

In this section, we have seen two reasons to believe that wh-movement of the embedded wh-phrase actually does take place out of the wh-phrase in which it is embedded, rather than being fed by extraposition of the embedded wh-phrase. First, extraposition of the relevant wh-phrases appears to be impossible. Secondly, even if it were possible, a theory that posited extraposition would make the wrong predictions about the ordering of the two wh-phrases; I have shown that the kind of derivation posited in this paper can be made to make the right predictions.

3. Where does wh-movement land?

Having determined that wh-movement of the embedded wh-phrase does in fact begin inside the embedding wh-phrase, let us consider the question of where this wh-movement lands. I have been claiming that wh-movement is to an additional specifier of CP, but we might entertain the hypothesis that movement is in fact to some position within the larger, containing wh-phrase. This hypothesis might be particularly attractive in the framework of Grewendorf (2001), in which wh-phrases may move not only by substitution to Spec CP but by adjunction to other wh-phrases.

There do seem to be reasons, however, to believe that the two wh-phrases are in distinct specifiers by the end of the derivation. For instance, if we embed a wh-question of the type under discussion here in another clause with a [+wh] specifier, it turns out that either of the two wh-phrases may move into the higher clause:

(15) a. [ot koi strani] se opitvash da razberesh
from which countries you-try to find-out
[[kolko studenti __] e ubil Ivan __]?
how-many students AUX killed Ivan
'From which countries are you trying to find out
[how many students __] Ivan killed?'
b. [Kolko studenti __] se opitvash da razberesh [[ot koi strani] e ubil Ivan __]?
'[How many students __] are you trying to find out
[from which countries] Ivan killed?'

(15a), of course, is also well-formed in English. The interesting case is (15b); here the embedded wh-phrase has remained behind in the lower clause, while the wh-phrase of which it was a part has moved up into the higher clause. This would seem to be inconsistent with a theory in which the embedded wh-phrase is moving to a position within the wh-phrase in which it is contained; the two wh-phrases need to be separable from each other.

4. When does wh-movement take place?

So far I have tried to show that the embedded wh-phrase is undergoing obligatory wh-movement, from a position inside another wh-phrase, to another specifier of CP (besides the one occupied by the other wh-phrase). Now I will address questions about the timing of the wh-movements; does the embedded wh-phrase undergo wh-movement before the wh-phrase in which it is embedded does, or after? Or is there some third option?

These are difficult questions to answer, and the answers I give here will be somewhat theory-dependent. Consider again examples like (7c) above, repeated as (16):

(16) [Kolko studenti __ ot Bulgaria] **[po kakvo]** vidja?
how-many students from Bulgaria of what you-saw
'How many students of what from Bulgaria did you see?'

The theory of ordering of wh-phrases in Bulgarian offered in Richards (1997, 2001) suggests that each wh-phrase "tucks in" to a specifier below any specifiers that exist when movement takes place; this is taken to follow from general principles of locality, which require wh-movement to land in a position as close to the attracting head, and to the tail of the chain, as possible. If this theory is on the right track, then the embedded wh-phrase *po kakvo* 'of what' must be landing in Spec CP after the wh-phrase in which it is embedded does.

Moreover, suppose we assume that Bulgarian has some version of the A-over-A condition--presumably to be made to follow, again, from general principles of locality, assuming that when one phrase dominates another the dominating phrase is closer to an attractor c-commanding them both than the dominated phrase is. If this assumption is

warranted, then the attractor should not be able to attract the embedded wh-phrase before it attracts the wh-phrase in which it is embedded; that is, it should attract the embedding wh-phrase first.

On some reasonably innocent assumptions, then, we seem driven to the conclusion that, of the two wh-phrases in the question in (16), attraction of the embedded wh-phrase happens second (because of the A-over-A condition), and that the embedded wh-phrase is the second to land in a specifier of CP (because of tucking in). To put it another way, wh-movement of the embedded wh-phrase begins after wh-movement of the wh-phrase in which it is embedded begins, and it ends after wh-movement of the embedding wh-phrase ends.

If wh-movement is instantaneous, then we may be finished; wh-movement of the embedded wh-phrase happens after the wh-phrase in which it is embedded has already moved to Spec CP, and thus must involve attraction by the head C out of its own specifier. However, it is at least conceivable that movement operations could be interspersed with each other. If wh-movement is successive-cyclic in the way envisioned in Chomsky (1998, 1999), for instance, then it would certainly be possible in principle for an operation to target the embedded wh-phrase after wh-movement of the embedding wh-phrase has already begun but before it reaches its landing site.

Such interleaving of wh-movement operations will have to be constrained, however. Consider (17), a straightforward instance of Superiority:

(17) a. What did you give __ to whom?
b. *To whom did you give what __?

As I mentioned immediately above, we could certainly imagine allowing movement of *to whom* in (17a) to begin taking place after movement of *what* had begun but before *what* had reached its landing site (perhaps while *what* was adjoined to the vP phase, for instance). This kind of "tandem movement" will have to be subject to constraints, however, which prevent the generation of examples like (17b). Locality (in this case, Superiority, presumably an instance of Shortest Attract) cannot simply be satisfied by a higher wh-phrase being the first to begin its journey up the tree; the lower wh-phrase will have to be unable to pass it.

Similar reasoning ought to apply in the Bulgarian case. If the A-over-A condition (again, hopefully an incarnation of Shortest Attract) bans attraction of the embedded wh-phrase while the embedding wh-phrase remains in situ, then this ban should be in force for the entire duration of the embedding wh-phrase's transition to Spec CP. Only once the embedding wh-phrase has landed in Spec CP should it be possible for the embedded wh-phrase to be extracted. In other words, it looks as though wh-extraction of the embedded wh-phrase must take place after the wh-phrase in which it is embedded has reached Spec CP; to put it yet another way, C must be able to attract the embedded wh-phrase out of its own specifier[1].

[1] One disturbing alternative derivation remains. What if C is not attracting the embedded wh-phrase out if its specifier, but rather attracting the "original copy" of the embedded wh-phrase out of the trace of wh-movement of the embedding wh-phrase? We could then stick to the assumption that heads may only attract out of their c-command domain (though we would need to assume that material inside a trace is available for later syntactic operations). In principle, there should be differences in empirical predictions

5. Conclusion

I have tried to argue in this paper that in principle, a head may attract material out of its own specifier, and that this material then moves to another specifier of the same head. If true, this conclusion removes the last obstacle to a simplification of the theory: we no longer need to state a "domain" which the head is allowed to search in order to determine what it should attract. Lowering is prevented by an essentially tautological condition, following from cyclicity; the attractor must attract an object which is present in the structure at the point in the derivation at which attraction takes place (perhaps, following Chomsky (1995), the portion of the derivation during which the attractor's maximal projection is undominated).

Appendix: a possible further application

Much of the preceding discussion has centered on contrasts like the one in (18):

(18) a. **[po kakvo]** [kolko studenti ___ ot Bulgaria] vidja?
of what how-many students from Bulgaria you-saw
'How many students of what from Bulgaria did you see?'
b.* Vidja [**po matematika**] [studenti __ ot Bulgaria]
you-saw of mathematics students from Bulgaria
'You saw students of mathematics from Bulgaria'

The contrast in (18) was meant to convince us, among other things, of the claim that the PP *po kakvo* 'of what' is really undergoing wh-movement in (18a). Here we have a case where a phrase (the PP complement of *studenti* 'students') ordinarily cannot move to the periphery of the phrase containing it (the DP, in this case), but can do so (in fact, is required to do so) when the DP undergoes movement to Spec CP and the PP is itself a wh-phrase. We should be looking, then, for other cases in which something cannot move to the periphery of a phrase unless that phrase has itself undergone movement.

The literature on pied-piping might offer us a case of the relevant kind. Van Riemsdijk (1984) and Aissen (1996) both discuss types of pied-piping which typically require movement of an operator within the pied-piped phrase to its left periphery:

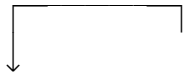
Basque (Ortiz de Urbina 1989, 248)

(19) **Nor** [joango d -ela ___] esan du Jon -ek?
who go AUX Q say AUX John-ERG
'Who has John said will go?

---

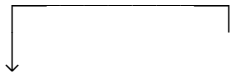
between this kind of derivation and the one proposed in the text. If the embedded wh-phrase is being extracted from the trace, for instance, then it is crossing all the material between C and the trace; if it is being extracted from the specifier, it does not cross that material (though the wh-phrase in which it is embedded does). I will try to exploit this difference in predictions in future work.

German (van Riemsdijk 1984, 165)



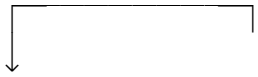
(20) den Wagen, **den** [zu kaufen __]er sich schon lange vorgenommen hatte
the car which to buy he self already long planned had
'the car which he had planned to buy for a long time'

Imbabura Quechua (Hermon 1984, 152)



(21) **ima**-ta [wawa ___ miku-chun-taj] Maria muna-n
what ACC child eat -FIN -Q Maria want -TNS.AGR
'What does Mary want (that) the child eat?'

Tzotzil (Aissen 1996, 457)



(22) **buch'u** [x-ch'amal __] icham ?
who 3 child died
'Whose child died?'

One kind of approach to examples like (22), for instance, would posit movement of *buch'u* 'who' to a high position within the moved DP, along with wh-movement of that DP into Spec CP; we would then develop a theory of pied-piping that required *buch'u* 'who' to move within DP in order to pied-pipe CP (see Aissen 1996 for such a theory).

The theory of the Bulgarian facts developed above, however, raises another possible account of the facts in (19-22); they might involve wh-movement, first of a DP or CP containing a wh-phrase, then of the wh-phrase itself to a distinct specifier. As it happens, a theory of movement that would yield these results might be constructable from available theoretical materials.

Chomsky (1999) suggests that a strong phase becomes inaccessible to the computation as a result of Spell-Out, which takes the material inside the phase and sends it to PF and LF. Suppose we were to accept that the interior of a strong phase becomes inaccessible to the computational system at some point, but reject the idea that this is a result of an irrevocable Spell-out operation. Instead, we might make use of one of Chomsky's other observations about phases, namely that they seem to be the kinds of objects that can undergo syntactic movement. The inaccessibility of objects inside a strong phase might be an instance of Shortest Attract. Attracting heads are seeking the closest available object that can move and contains an instance of the feature being attracted; objects buried inside phases are therefore unmoveable, since the containing phase is closer to the attracting head and is therefore a preferred option. We could formalize this idea via the version of Shortest Attract in (23):

(23) a. A head X must attract a syntactic object α that contains an instance of the feature being attracted, such that there is no β also containing the attracted feature which is closer to X than α is.
b. α is closer to X than β just in case the lowest node dominating α dominates β, and the reverse is not true.

The version of Shortest Attract in (23) defines closeness in terms of a version of c-command that lacks the clause stating that α and β are not in a c-command relation if one dominates the other. This version has the virtue of covering both Superiority and A-over-A condition effects; α will be closer to a c-commanding head than β if α c-commands β, and also if α dominates β. If we wish to maintain the claim that the highest specifier of a phase is an escape hatch for extraction from the phase, we can do so by defining domination along the lines in Kayne (1994), which yield the result that the specifier of X is not taken to be dominated by XP. The specifier of XP and XP itself will then both be dominated by the same lowest node, and will be equally close to attracting heads, as desired.

Furthermore, there will have to be a general condition excluding objects that have already been attracted from consideration; this is generally assumed, though seldom made explicit. In a Superiority configuration, for instance, attraction of the highest wh-phrase will have to make it possible to attract lower wh-phrases afterwards, even though the highest wh-phrase is still higher than the other wh-phrases (and thus "closer" in the sense in (23b)) after it has undergone wh-movement to Spec CP.

Something like this set of assumptions might account for the data in (19-22). The derivation for (22) might proceed as follows; first, the tree is built up to the interrogative C, as in (24):

(24) **C** icham [x-ch'amal **buch'u**]
died 3 child who

Assume that DP is a phase, and hence, following Chomsky, a movable syntactic object. Wh-movement of *buch'u* 'who' is therefore impossible, since the closest moveable object containing the wh-feature sought by C is the DP *x-ch'amal buch'u* 'whose child'[2]. The DP can thus be attracted by C:

(25) [x-ch'amal **buch'u**] **C** icham ___ ?
3 child who died

Now that the DP phase has been attracted, Shortest Attract has been obeyed; C attracts *buch'u* 'who' into a higher specifier, just as in the Russian doll questions from Bulgarian:

(25) [**buch'u**] [x-ch'amal __] **C** icham ___?
who 3 child died

---

[2] Of course, it might be possible for *buch'u* 'who' to move to the edge of the DP phase while the DP is still being constructed, and thus escape from DP (this option apparently is available; cf. Aissen 1996 for details). The derivation under consideration here is one in which it does not do this.

Many questions remain, of course. We would like to know, for one thing, why the Tzotzil, German, Basque, and Quechua versions of the derivation discussed for Bulgarian cannot involve tucking in of the extracted wh-phrase:

(25) *[x-ch'amal __] [**buch'u**] **C** icham ___?
3 child who died

I will have to leave questions like these for future work.

References

Aissen, Judith. 1996. Pied-piping, abstract agreement, and functional projections in Tzotzil. *Natural Language and Linguistic Theory* 14.447-491.

Bošković, Željko. 1997.On certain violations of the Superiority Condition, AgrO, and economy of derivation. *Journal of Linguistics* 33.227-254.

Bošković, Željko. 1999. On multiple feature checking: multiple wh-fronting and multiple head movement. Samuel David Epstein and Norbert Hornstein (eds.) *Working Minimalism*, MIT Press, Cambridge.

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

Chomsky, Noam. 1998. Minimalist inquiries: the framework. ms., MIT.

Chomsky, Noam. 1999. Derivation by phase. ms., MIT.

Grewendorf, Günther. 2001. Multiple wh-fronting. *Linguistic Inquiry* 32.87-122.

Hermon, Gabriella. 1984. *Syntactic modularity*. Foris, Dordrecht.

Kayne, Richard. 1994. *The antisymmetry of syntax.* MIT Press, Cambridge.

Ortiz de Urbina, Jon. 1989. *Parameters in the grammar of Basque,* Foris, Dordrecht.

Richards, Norvin. 1997. *What moves where when in which language?* Doctoral dissertation, MIT.

Richards, Norvin. 2001. *Movement in language: interactions and architectures*. Oxford University Press.

Riemsdijk, Henk van. 1984. On pied-piped infinitives in German relative clauses. Jindrich Toman (ed.) *Studies in German grammar*, Foris, Dordrecht.

Department of Linguistics and Philosophy
E39-245
MIT
Cambridge, MA 02139

norvin@mit.edu