Featural Cyclicity and the Ordering of Multiple Specifiers

Norvin Richards, MIT

norvin@mit.edu

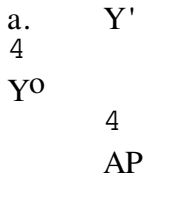
DRAFT: comments welcome

Chomsky (1995) suggests that we derive cyclicity by assuming something like the following principle:

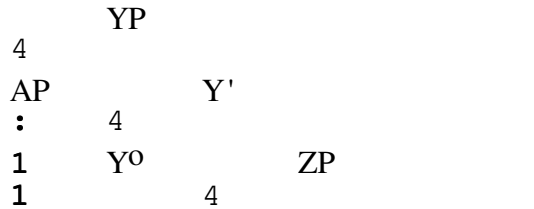
1. A strong feature must be checked as soon as possible after being introduced into the derivation, before any new structure is merged.

Together with the assumptions that (1) Merge always expands the tree and (2) overt movement can only take place in response to a strong feature, this principle derives cyclicity in a pleasingly minimalist way. Consider the derivations in 2-3:

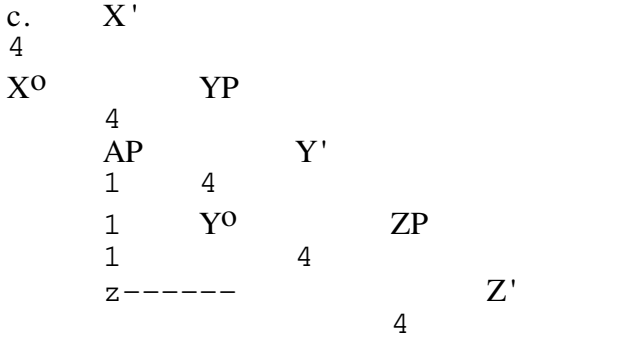
2. a.


Y'
Yo
ZP
AP
Z'
Zo

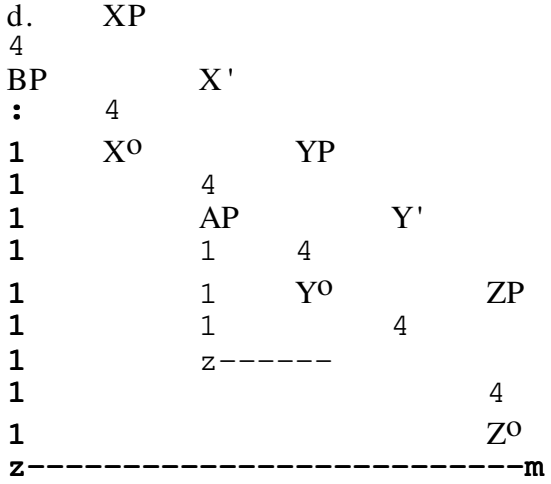
b.


YP
AP
Y'
Yo
ZP
Z'
Zo
BP

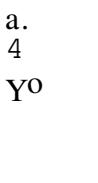
c.


X'
Xo
YP
AP
Y'
Yo
ZP
Z'
Zo
BP

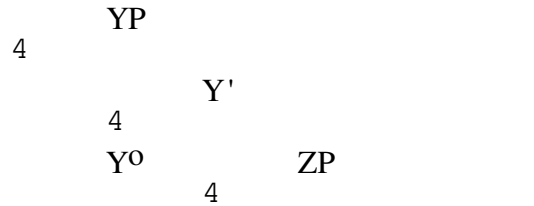
d.


XP
BP
X'
Xo
YP
AP
Y'
Yo
ZP
Z'
Zo

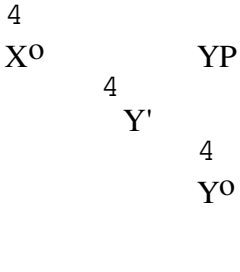
3. a.


Y'
Yo
ZP
AP
Z'
Zo
BP

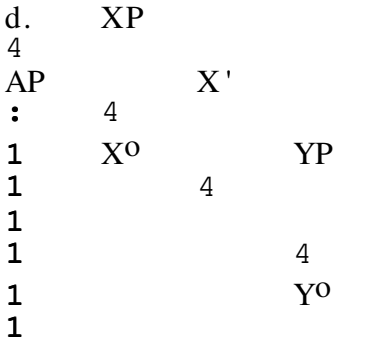
b.


YP
Y'
Yo
ZP
AP
Z'
Zo
BP

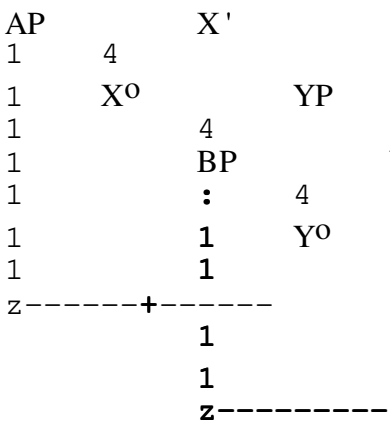
c.


X'
Xo
YP
Y'
Yo
ZP
AP
Z'
Zo
BP

d.


XP
AP
X'
Xo
YP
Y'
Yo
ZP
Z'
Zo
BP

e.

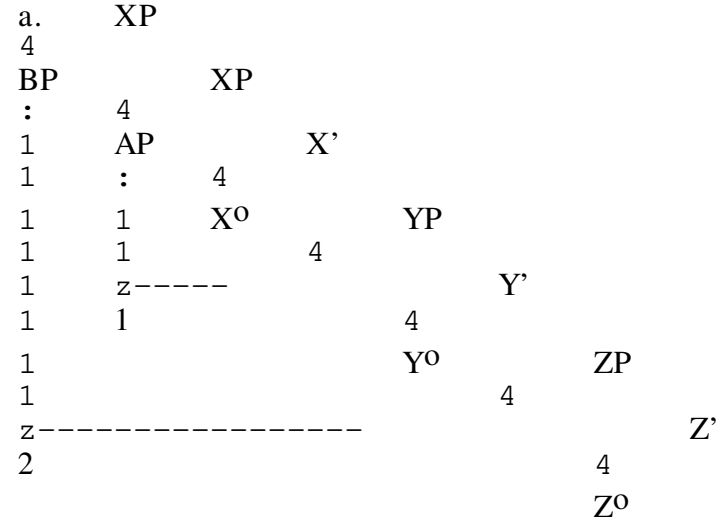

XP
AP
X'
Xo
YP
BP
Y'
Yo
ZP
Z'
Zo
BP

2 and 3 both involve a ZP with specifier AP and complement BP, to which are added the heads Y^o and X^o, both of which bear a strong feature which might in principle attract either AP or BP. 2 is the well-formed, Cyclicity-obeying derivation; in 3, by contrast, Cyclicity is disobeyed, as BP moves to Spec YP after XP has already been projected (in step e). The principle in 1 correctly distinguishes between the two derivations. In the well-formed derivation in 2, the strong features introduced in the heads Y^o and X^o are checked off in the steps immediately after the features are introduced, as 1 requires. In 3, on the other hand, the strong feature introduced in Y^o in step a. is not checked off until step e. Cyclicity thus correctly rules out the derivation.

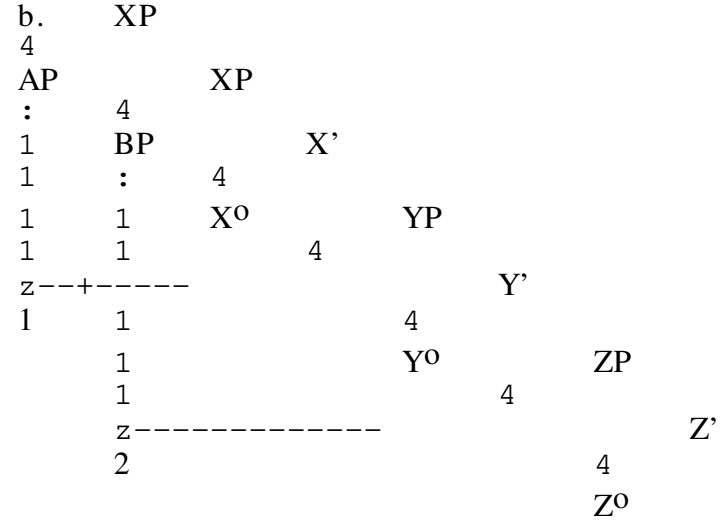
This way of deriving cylicity avoids problems raised by head-movement for Chomsky's (1993) definition of cyclicity. Chomsky (1993) suggested that all operations must necessarily expand the tree. This requirement successfully distinguishes between the derivations in 2-3, but it is always violated by head-movement, which apparently never expands the tree. The definition of cyclicity in terms of strong features, on the other hand, can be satisfied by head-movement, as long as the head-movement is checking a strong feature.

The two versions of cyclicity might also make different predictions in cases of movement to multiple specifiers of a single head. Suppose a head is generated with two strong features and attracts two XPs, as in 4:

4. a.



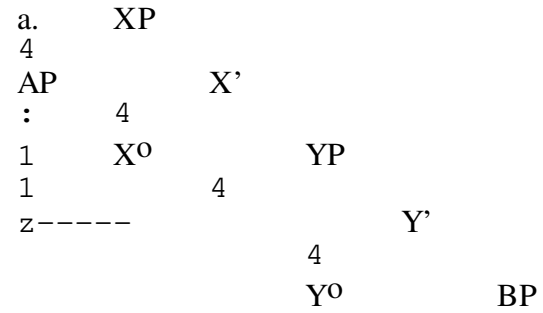
b.



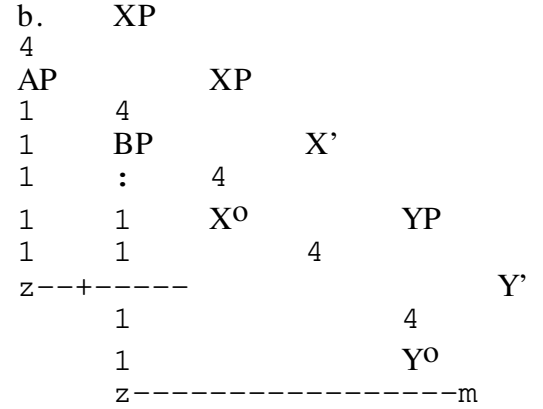
In 4.a the two specifiers are treated just like specifiers of two separate heads; the first movement is to a specifier which is lower than the landing site of the second move. In 4.b, on the other hand, the two paths cross, and the second move lands closer to the head than the first move. Chomsky's (1993) derivation of cyclicity from a requirement that every operation expand the tree would rule out the derivation in 4.b; the second move here does not expand the tree. Chomsky's more recent version of cyclicity, on the other hand, fails to distinguish between the two derivations. As long as both XPs are moving to check a strong feature, either derivation ought in principle to be possible.

In fact, depending on what version of Shortest Move we assume, we might expect the derivation in 4.b to be preferred over that in 4.a. Consider the derivation in 4.b again, step by step:

5. a.



b.



In 5.a, we first move AP into a specifier of XP. In 5.b, we move BP, and must decide where to move it to; does it go to a specifier outside AP, or one inside AP? Featural cyclicity, as we have seen, allows both options. On the other hand, if the specifier inside the one containing AP is closer, in the relevant sense, to the base position of BP than the one outside AP is, then Shortest Move will require us to move to the lower specifier, as shown in 5.b.

A feature-based notion of cyclicity, then, along with a certain conception of Shortest Move, predicts that multiple specifiers of a single head will be treated very differently from specifiers of multiple heads. Paths to such specifiers ought to cross, rather than nesting, thus maintaining the base c-command relations (and, in principle, the base order) among XPs which move to them. In this paper I will argue that this is a correct

result; paths to multiple specifiers of a single head do indeed cross, all other things being equal.

A non-trivial question which I will not address at any length in this paper is precisely what version of Shortest Move we will need to get this result. The theory developed here is compatible with any of a number of technical implementations of this notion. The basic idea would be that feature-checking is subject to locality requirements which force the landing site of the moving element to be as close to the attracting head as possible. Something like this is presumably necessary in any case, in order to prevent feature-checking from simply being able to occur with no movement at all, or via movement to a specifier of a higher head. The locality condition in question will have to differ from the condition of Shortest Attract, however, in that it will view multiple specifiers as being at different heights, while Shortest Attract fails to distinguish between multiple specifiers of a single head; all are equally available for attraction.

1. Multiple wh-movement

Suppose we consider the case of multiple wh-movement, one case in which movement to multiple specifiers is arguably involved. As is well known, multiple wh-movement is subject in some languages to a restriction on the order of movement; the highest wh-word must be moved first:

6. a. Who t bought what?

b.?? What did who buy t?

This is plausibly viewed as an effect of Shortest Move; the C^{o} which attracts the wh-words prefers to attract *who* rather than *what*, since *who* moves a shorter distance. Of course, movement of *what* must then follow anyway, but on the assumption that the grammar cannot look ahead in the derivation, this is irrelevant to the choice of which wh-word to move first.

In certain other languages, Superiority phenomena seem to take on a rather different form. Our standard assumptions seem to make precisely the wrong predictions for wh-

movement in certain multiple overt wh-movement languages. Rudin (1988) shows that such languages can be divided into two groups; those which impose no ordering on multiple fronted wh-words (Serbo-Croatian, Polish), and those which do (Bulgarian, Romanian). For those languages which do impose such an ordering, the order essentially preserves the base c-command order (Bulgarian, from Rudin 1988):

7. a. Koj kogo e vidjal?
who whom AUX seen
'Who saw whom?'

b. *Kogo koj e vidjal?

Here we must apparently give up one of our standard assumptions. If we wish to maintain the idea that movement always expands the tree, creating a specifier higher than all the existing structure, we must apparently conclude here that a kind of "Anti-Superiority" is at work; the lower of the two wh-words (*kogo* 'whom' in 7.a) must move first. Another possibility, of course, would be to say that the order of wh-movements in this case is just as in English; *koj* 'who' moves first, followed by movement of *kogo* 'what' to a lower specifier. On this account, the paths of these multiply fronted wh-words must obligatorily cross, rather than nest.

In Japanese, we find a similar phenomenon dubbed "Anti-Superiority" by Saito (1982):

8. a. Taroo -ga nani -o naze katta no?
Taroo NOM what ACC why bought Q
'Why did Taroo buy what?'

b. ?Taroo -ga naze nani -o katta no?
Taroo NOM why what ACC bought Q
'What did Taroo buy why?'

Saito points out that the contrast between 8.a and 8.b can be given an explanation based on the ECP (Chomsky 1981). On such an account, movement of *naze* 'why' must precede

movement of *nani-o* 'what-ACC', in order to secure antecedent-government from Comp; the trace of *nani-o* can then be lexically governed by the verb, and the structure is well-formed. This is referred to as "Anti-Superiority" because it apparently requires us to assume that the lower of the two wh-words is moved first; 8.a, where *nani-o* 'what' precedes and presumably c-commands *naze* 'why', is the well-formed structure, in which, on Saito's theory, *naze* is able to move first. Again, this is the opposite of the order we expect on traditional versions of Cyclicity; the wh-movement paths appear to be crossing, rather than nesting, in just this case.

Thus far, the generalization appears to be that we get Anti-Superiority effects when multiple wh-words are moved to a single landing site on a single level (at LF, in Japanese, or in the overt syntax, in Bulgarian), and Superiority effects when wh-words are moved on different levels (as in English). In fact, there is some further evidence for this generalization from Japanese.

Takahashi (1993) notes that Japanese long-distance scrambling of wh-words exhibits a Superiority effect; scrambling of a wh-word over another wh-word is impossible (from Takahashi (1993, 664)):

9. a. John-ga [Bill -ga dare -ni [Mary -ga nani -o tabeta to]
John NOM Bill NOM who DAT Mary NOM what ACC ate that
itta to] omotteiru no?
said that thinks Q
'Who does John think that Bill told that Mary ate what?'

b. **Dare-ni** John-ga [Bill-ga **t** [Mary-ga nani-o tabeta to] itta to] omotteiru no?

c. ***Nani-o** John-ga [Bill-ga dare-ni [Mary-ga t tabeta to] itta to] omotteiru no?

This is what our generalization leads us to expect; here, a single wh-word is being attracted in the overt syntax, so it must obey Superiority, just as in English. Multiple long-distance scrambling of wh-words, on the other hand, apparently exhibits Anti-Superiority effects again, just as we expect (Takako Aikawa, p.c.); the paths must cross:

10. a. John -ga [Tanaka-sensee -ga dare -ni nani -o yomaseta to] itta no?
John NOM Tanaka teacher NOM who DAT what ACC read-CAUS that said Q
'Who did John say Professor Tanaka made read what?'

b. Dare$_i$ -ni nani$_j$ -o John -ga [Tanaka-sensee-ga t_i t_j yomaseta to] itta no?
who DAT what ACC John NOM Tanaka teacher NOM read-CAUS that said Q

c. *Nani$_j$ -o dare$_i$ -ni John -ga [Tanaka-sensee-ga t_i t_j yomaseta to] itta no?
what ACC who DAT John NOM Tanaka teacher NOM read-CAUS that said Q

Thus, it is not simply the case that Japanese is an "anti-superiority language"; the choice between Superiority and Anti-Superiority is apparently a property of particular constructions, rather than of languages.

We have seen, then, that in all the cases in which the ordering of multiple wh-movements to a single landing site can be observed, the paths must apparently cross. In cases where a single wh-movement to a single landing site takes place, of course, the highest available wh-word must move. As was observed above, this is precisely the prediction of a featural theory of cyclicity, together with a certain definition of Shortest Move. Given these assumptions, we expect the highest wh-word to move first, followed by movement of the next highest wh-word to a specifier below the landing site of the first movement. Such a derivation does seem to give us the correct word order.

In fact, there is also evidence that the derivation proceeds in the order predicted, with the higher wh-word moving first, just as in English. Such evidence will necessarily have to involve a phenomenon which is sensitive to the order of syntactic operations. The data in 11, as I have argued elsewhere, exemplify such a phenomenon (Bulgarian, from Roumyana Izvorski, Ani Petkova, and Roumyana Slabakova):

11. a. *Koja kniga$_i$ otrece senatorat [malvata ce pravitelstvoto iska da zabrani t$_i$]?
which book denied the-senator the-rumor that the-government wanted to ban
'Which book did the senator deny the rumor that the government wanted to ban?'

b. ? Koj senator koja kniga$_i$ otrece [malvata ce pravitelstvoto
which senator which book denied the-rumor that the-government
iska da zabrani t$_i$]?
wanted to ban
'Which senator denied the rumor that the government wanted to ban which book?'

11 shows that a Complex NP Constraint violation in Bulgarian can be redeemed by moving another wh-phrase into another specifier of the same C^o. In Richards (1995, 1996), I suggested that this fact should be dealt with by the same theory that accounts for the well-known contrast between 12.a and the reading of 12.b in which *what* has wide scope:

12. a. *What do you wonder who bought?
b. Who wonders who bought what?

The generalization, across both languages, seems to be that a wh-movement can disobey Subjacency if a Subjacency-obeying movement to the same position has already occurred. Note that this phenomenon is apparently sensitive to the order of operations; the Subjacency-obeying movement must precede the Subjacency-violating movement, as 13 shows:

13. a. Who$_i$ t$_i$ persuaded the man who bought which car$_j$ to sell the hubcaps?
b. *Which$_i$ car did John persuade the man who bought t$_i$ to sell which hubcaps?

In 13.a we see another case of a well-formed movement licensing a later ill-formed movement; in this case, the later movement is that of *which car* out of the relative clause to the matrix Comp, and the well-formed move is that of *who*. In 13.b, we have attempted to license an ill-formed movement with a well-formed movement again, but in this case the well-formed movement is that of *which hubcaps*, which follows the ill-formed movement[1]

[1] Here I assume, as is standard, that covert movements follow overt movements in the derivation.

of *which car*. Apparently, then, the licensing move must precede the move which it licenses. In English, of course, this will mean that the licensing move will be overt and the licensed move covert, since only the first wh-movement to a given position is overt in English. However, the Bulgarian facts, which parallel the English facts but involve only overt movements, suggest that the overt/covert distinction has nothing to do with this phenomenon.

We have seen, then, that we can use Subjacency-amelioration as a diagnostic for the order of movements; if one movement is to license another, it must occur first. Consider the Bulgarian facts again:

14. a. *Koja kniga$_i$ otrece senatorat [malvata ce pravitelstvoto iska da zabrani t$_i$]?
which book denied the-senator the-rumor that the-government wanted to ban
'Which book did the senator deny the rumor that the government wanted to ban?'

b. ? Koj senator koja kniga$_i$ otrece [malvata ce pravitelstvoto
which senator which book denied the-rumor that the-government
iska da zabrani t$_i$]?
wanted to ban
'Which senator denied the rumor that the government wanted to ban which book?'

In the well-formed 14.b, the Subjacency-obeying wh-phrase *koj senator* 'which senator' linearly precedes the Subjacency-violating wh-phrase *koja kniga* 'which book'. If the chain of reasoning outlined above is correct, this means that *koja kniga* must have moved to Spec CP first, followed by movement of *koja kniga* to a lower Spec CP[2]. This is precisely what the featural notion of cyclicity predicts.

2. Multiple A-specifiers

We have now seen that the predictions of a theory incorporating featural cyclicity and a certain notion of Shortest Move seem to be borne out in cases of multiple wh-movement; we can successfully explain ordering restrictions on wh-movement in a variety

[2] footnote about Danny's suggestion for another possible significance for the test...

of languages using fairly simple assumptions. Given such simple assumptions, however, we expect to find this pattern arising quite generally in all cases involving movement to multiple specifiers of a single head. Is it possible to find evidence for this theory from the domain of A-movement, for example?

15 shows that Japanese local scrambling, which might in principle be thought to involve movement to multiple A-specifiers, does not appear to exhibit any Superiority effects. Two arguments can be scrambled over the subject in either order:

15. a. Taroo-ga gakusei -ni hutari hon -o ageta
Taroo NOM student DAT two book ACC gave
'Taroo gave books to two students'

b. Gakusei- ni hutari hon -o Taroo -ga ageta
student DAT two book ACC Taroo NOM gave

c. Hon -o gakusei -ni hutari Taroo -ga ageta
book ACC student DAT two Taroo NOM gave

Of course, there may be any of a number of confounding factors preventing us from observing the effects of Superiority here; there may, for example, be more than one attracting head involved. Developing a complete theory of scrambling is well beyond the scope of this paper. In the next sections I will simply try to show that Superiority effects do indeed hold in a certain well-defined subset of A-scrambling cases, which suggests that the general account of Superiority developed above may be on the right track. It is well-known that local scrambling interacts with such semantic properties as definiteness, specificity, and quantifier scope. In the next few sections, I will try to show that when these effects are controlled for, local scrambling does obey Superiority, as defined here. Sections 2.1 and 2.2 will show that local scrambling which has no semantic effects obeys Superiority. In section 2.3 I will give some evidence suggesting that local scrambling which violates Superiority does in fact involve multiple attractors.

2.1 Part 1: idiom chunks

One area we might want to consider is the scrambling of idiom chunks. Idiom chunks are presumably non-referential and therefore impervious to any effects scrambling might have on discourse properties. If it is these effects which are responsible for the apparent failure of A-scrambling to conform to Superiority, then, we might expect to get a clearer picture by looking at the scrambling of idiom chunks. Miyagawa (1994) notes that scrambling of idiom chunks in Japanese is confined to short-distance scrambling; long-distance scrambling, as in 16.b, is impossible:

16. a. Kosi-o John -ga t orosita
hip ACC John NOM lowered
'John sat down'
b. ?*Kosi -o Mary -ga [John -ga t orosita to] itta
hip ACC Mary NOM John NOM lowered that said
'Mary said that John sat down'

Idiom chunks, then, may apparently only be A-scrambled. It is interesting to note, therefore, that multiple idiom chunks do indeed appear to obey Superiority in their short-distance scrambling. If scrambling only affects one idiom chunk, it must raise the higher of the two (Shigeru Miyagawa, Kazuko Yatsushiro, p.c.):

17. a. Taroo -ga hi -ni abura -o sosoida
Taroo NOM fire DAT oil ACC poured
'Taroo made things worse'
b. Hi -ni Taroo -ga t abura-o sosoida
fire DAT Taroo NOM oil ACC poured
c. *Abura-o Taroo-ga hi-ni t sosoida
oil ACC Taroo NOM fire DAT poured

Both idiom chunks may also be scrambled, but the two paths must cross, in conformance with Superiority:

18. a. Hi$_i$ -ni abura$_j$-o Taroo -ga t$_i$ t$_j$ sosoida
fire DAT oil ACC Taroo NOM poured
'Taroo made things worse'
b. *Abura$_j$-o hi$_i$ -ni Taroo -ga t$_i$ t$_j$ sosoida
oil ACC fire DAT Taroo NOM poured

Superiority, then, does not affect only A-bar movement. Furthermore, as we saw with A-bar movement, multiple A-movements to a single attractor must apparently cross and not nest, just as the account developed here predicts.

2.2 Part 2: quantifier scope

Japanese A-scrambling typically has effects on quantifier scope, creating scope ambiguities which are unavailable in the base order (Japanese, from Kuroda (1971))

19. a. Dareka -ga daremo -o hihansita
someone NOM everyone ACC criticized
'Someone criticized everyone' (∃>∀, *∀>∃)
b. Daremo -o dareka -ga t hihansita
everone ACC someone NOM criticized
'Someone criticized everyone' (∃>∀, ∀>∃)

However, Kazuko Yatsushiro (1996, and p.c.) has noted that there is in fact a type of local scrambling of quantifiers which has no effect on scope relations. This is precisely the scrambling which obeys Superiority, as defined in this paper (Yatsushiro 1996 and p.c.):

20. a. John-ga dareka -ni daremo -o syookaisita
John NOM someone DAT everyone ACC introduced
'John introduced everyone to someone'
unambiguous: someone>everyone
b. Dareka -ni John -ga t daremo -o syookaisita
someone DAT John NOM everyone ACC introduced
unambiguous: someone>everyone
c. Dareka$_i$ -ni daremo$_j$ -o John -ga t$_i$ t$_j$ syookaisita
someone DAT everyone ACC John NOM introduced
unambiguous: someone>everyone

Thus, when scrambling affects the higher of two quantifiers, as in 20.b, or affects both quantifiers but preserves the underlying c-command relation between them, as in 20.c, the sentence remains unambiguous. When Superiority is disobeyed, by contrast, ambiguity is created:

21. a. Daremo -o John -ga dareka -ni t syookaisita
everyone ACC John NOM someone DAT introduced
ambiguous: someone>everyone, everyone>someone
b. Daremo$_j$ -o dareka$_i$ -ni John -ga t$_i$ t$_j$ syookaisita
everyone ACC someone DAT John NOM introduced
ambiguous: someone>everyone, everyone>someone

The contrast between 20.c and 21.b, in particular, is a striking one. On the account developed here, the ordering in 21.b can only be the result of attraction by multiple attractors, a conclusion for which more evidence will be adduced in section 2.3. The ordering in 20.c, on the other hand, can be produced by movement to multiple specifiers of a single head. It could also, in principle, be a result of movement to specifiers of multiple heads (say, A-movement of *daremo-o* 'everyone' followed by A-bar-movement of *dareka-ni* 'someone' into a higher maximal projection). The lack of ambiguity in 20.c suggests that this second option is in fact excluded. A number of accounts might be given for this; one possibility is that speakers assume only as many attractors as are necessary to account for a particular word order[3]. The availability of a grammatical structure for the word order in 20.c which involves only a single attracting head, then, rules out structures with more than one such attractor.

We have seen that a certain class of A-scrambling movements do in fact appear to obey Superiority: namely, scrambling of idiom chunks and local quantifier scrambling which fails to create scope ambiguity. This lends some support to the idea that A-scrambling does in principle obey Superiority, and that what create the appearance of disobedience of Superiority are processes which are sensitive to such phenomena as referentiality and quantifier scope.

2.3 Part 3: scrambling generally

[WARNING: this section is based largely on rather shaky intuitions, and should not be trusted too far]

In the previous sections we have seen certain cases of semantically vacuous local scrambling which apparently must obey Superiority. In this section I will argue further that scrambling with crossing paths is (or can be) the result of multiple attraction by a single attractor, while scrambling with nesting paths must involve multiple attractors.

---

[3] Footnote about 'die Mutter liebt die Tochter' and 'Kakakain ng leon ng tigre'.

The argument is based on a kind of relativization in Japanese which violates islands, discussed (among others) by Kuno 1973, Hasegawa 1984, and Ochi 1996:

22. [[e_i e_j kiteiru] huku$_j$ -ga yogoreteiru] kodomo$_i$
wear clothes NOM dirty child
'the child$_i$ that the clothes$_j$ that t_i is wearing t_j are dirty'

Here relativization of the operator associated with *kodomo* 'child' seems to take place from inside a relative clause modifying *fuku* 'clothes', in violation of the wh-island condition and the CED. Such relativization, as Hasegawa (1984) notes, is strongly constrained. Hasegawa observes that the relative clause out of which relativization takes place (which I will call, for purposes of exposition, the "contained" relative clause) must modify the subject of the relative clause by which it is contained. Thus, 23 contrasts with 22:

23. *[[Mary-ga [e_i e_j kiteiru] huku$_j$ -o tukutta] kodomo$_i$
Mary NOM wear clothes ACC made child
'the child$_i$ that Mary made the clothes$_j$ that t_i is wearing t_j'

In fact, the conditions on the positioning of the contained relative clause are somewhat freer than this. In particular, there is a contrast between 24.a-b, as well as between 25.a-b (Takako Aikawa, p.c.):

24. a.*[Taroo-ga [e_i e_j kaita] hon$_j$ -o katta] sakka$_i$
Taroo NOM wrote book ACC bought author
'the author$_i$ that Taroo bought the book$_j$ that t_i wrote t_j'
b. ?[Taroo-ga [e_i e_j kaita] hon$_j$ -ga suki na] sakka$_i$
Taroo NOM wrote book NOM likes author
'the author$_i$ that Taroo likes the book$_j$ that t_i wrote t_j'

25. a.*[Taroo-ga [e_i e_j kaita] hon$_j$ -o katta] sakka$_i$
Taroo NOM wrote book ACC bought author
'the author$_i$ that Taroo bought the book$_j$ that t_i wrote t_j'
b.?[[e_i e_j kaita] hon$_j$ -o Taroo-ga katta] sakka$_i$
wrote book ACC Taroo NOM bought author
'the author$_i$ that the book$_j$ that t_i wrote t_j, Taroo bought '

In 24, we can see that the contained relative clause can modify the object of a double-*ga* predicate; such predicates have recently (cf. Ura 1996) been analyzed as having the multiple *ga*-marked elements in multiple specifiers of a single head. Furthermore, 25 shows that scrambling can affect the acceptability of contained relative clauses; a contained relative clause can modify an direct object which has been scrambled to the front of its clause.

I will not try to develop a full theory here of the nature of this kind of relativization, but the contrast in 24 suggests that we can use contained relative clauses as a test for structure. If a non-initial argument is capable of hosting a contained relative clause, we may conclude that it and the preceding argument(s) are in multiple specifiers, as in 24.b, rather than in separate maximal projections, as in 24.a.

Having established this, let us turn to the distribution of contained relatives in double object constructions. On the theory proposed here, scrambling of both double objects which preserves the base c-command relation between the scrambled arguments may involve scrambling to multiple specifiers of a single head. Scrambling which alters the base c-command relation, on the other hand, must necessarily involve movement to distinct maximal projections. The prediction, then, is that scrambling which preserves the basic IO-DO order will allow contained relative clauses to modify either object, since the objects, like the *ga*-marked nominals in 24.b, are in multiple specifiers of a single head. Scrambling which creates the DO-IO order, on the other hand, will only allow a contained relative clause on the DO, since the DO and IO must be in specifiers of different heads, like the subject and object in 24.a.

These predictions seem to be fulfilled (Takako Aikawa, p.c.):

IO > DO

contained relative on DO:

26. a. ? [Hanako-ni [t_i t_j kaita] hon$_j$ -o Taroo -ga ageta] sakka$_i$
Hanako DAT wrote book ACC Taroo NOM gave author
'the author$_i$ that Taroo gave Hanako the book$_j$ that t_i wrote t_j'

contained relative on IO:

b. ? [[t_i t_j osieta] gakusee$_j$-ni syoo -o committee-ga ataeta] sensee$_i$
taught student DAT prize ACC committee NOM gave teacher
'the teacher$_i$ that the committee gave the student$_j$ that t_i taught t_j a prize'

DO > IO

contained relative on DO:

27. a. ? [[t_i t_j kaita] hon$_j$ -o Hanako-ni Taroo -ga ageta] sakka$_i$
wrote book ACC Hanako DAT Taroo NOM gave author
'the author$_i$ that Taroo gave Hanako the book$_j$ that t_i wrote t_j'

contained relative on IO:

b. * [syoo -o [t_i t_j osieta] gakusee$_j$-ni committee -ga ataeta] sensee$_i$
prize ACC taught student DAT committee NOM gave teacher
'the teacher$_i$ that the committee gave the student$_j$ that t_i taught t_j a prize'

In the last section, we saw that a single attractor is apparently necessarily involved whenever multiple scrambling preserves the c-command relation between scrambled elements; that is, no more attractors are posited than necessary. This was taken to account for the lack of quantifier scope ambiguity in cases in which double objects are scrambled in a way which preserves their base order. In this section, we have seen that scrambling which does not preserve the base order must necessarily involve multiple attractors. The most straightforward theory about scrambling, then, can apparently be maintained; scrambling is subject to the general conditions on movement discussed in this paper. The apparent violations of Superiority found in scrambling are a result of the availability of multiple attractors. Speakers apparently may and must posit exactly as many attractors as are necessary to account for the word order of a particular sentence.

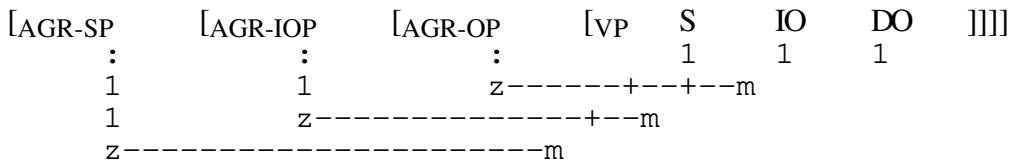
3. Object Shift

In the preceding sections we have seen evidence that nested paths are a result of multiple attraction by multiple attractors, while crossing paths are a result of multiple attraction by a single attractor. To the extent that we find this generalization accurate and

compelling, we are entitled to suspect the work of a single attractor in cases where paths must obligatorily cross.

One case in which crossing paths are standardly invoked is in the movement of arguments from their base positions to their case-checking positions. Some variant of the following structure has traditionally been assumed for such movement:

28.



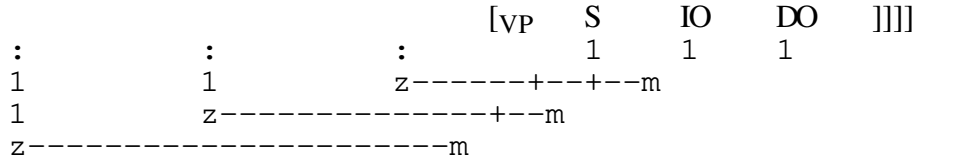
That is, it is typically assumed that case-checking takes place in the specifiers of three distinct heads, and that movements to the specifiers of these heads must obligatorily cross. This assumption is taken to account, for instance, for the data in 29-30 (29 is Dutch, from Neeleman 1994, 419; 30 is Icelandic, from Collins and Thráinsson 1993, 143, 145, 149, 154):

29. a. ...dat Jan **gisteren** de mannen de foto toonde
that Jan yesterday the men the picture showed
b. ...dat Jan de mannen **gisteren** de foto toonde
that Jan the men yesterday the picture showed
c. ...dat Jan de mannen de foto **gisteren** toonde
that Jan the men the picture yesterday showed
'that Jan showed the men the photo yesterday'
d. *...dat Jan de foto **gisteren** de mannen toonde
d.*...dat Jan de foto de mannen **gisteren** toonde

30. a. Ég lána **ekki** Maríu bækurnar
I lend not Maria books
b. Ég lána Maríu **ekki** bækurnar
I lend Maria not books
c. Ég lána Maríu bækurnar **ekki**
I lend Maria books not
'I do not lend Maria the books'
d. * Ég lána bækurnar **ekki** Maríu
e. * Ég lána bækurnar Maríu **ekki**

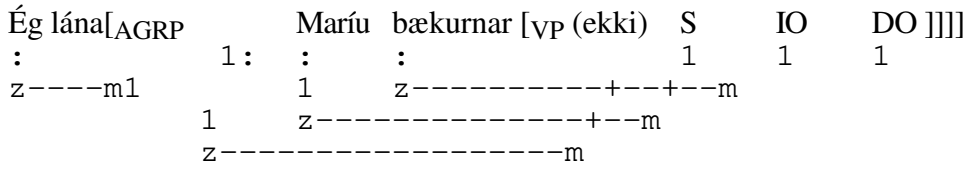
Examples such as these are typically taken as involving shift of the NPs in question over the relevant adverbs. Crucially, in these cases no reordering of the arguments with respect to other arguments is possible; that is, if only one argument shifts, it must be the higher of the two, and if both arguments shift, their paths must cross.

Given the approach developed here, we might take the obligatory crossing of the paths in this case to indicate that only a single attractor is at work. That is, we might alter 28 to 31:

31. 

We might assume, then, that argument shift is always triggered by a single AgrP head, which obligatorily triggers overt raising of some non-null set of arguments (that is, at least the subject, and possibly some of the other arguments), the case features of which are checked by Agr. There are a number of other assumptions which are consistent with these data (for instance, the subject could be base-generated in a higher position, rather than obligatorily raising), but I will attempt to maintain this one for the time being, as it strikes me as the simplest one available.

This is presumably not the whole story. I assume that there is at least one other head which attracts a single XP to its specifier and hosts the finite verb in the overt syntax, which we can think of as the head responsible for the EPP; I will refer to this head as π (cf. Branigan 1992). The structure of a sentence like 30.c, then, would be as in 32:

32. 

One potential advantage of this approach over some of its predecessors is that of simplicity. The movements ascribed to arguments in this picture follow completely straightforwardly from the same mechanisms as those used to regulate A-bar movement; no additional stipulations are necessary, and the number of maximal projections assumed diminishes considerably. Of course, this is not an advantage if independent motivation can be mustered for either the stipulations or the maximal projections.

A second potential advantage will only be an advantage if the facts turn out right, and at this point I have only a tiny fraction of the relevant facts. Consider two groups of languages, one of which I will refer to as "object shift languages", and the other as "non-object-shift languages". The groups will crucially differ from those established by the [±Spec TP] parameter of Jonas and Bobaljik (1993), Bobaljik (1995), Bobaljik and Jonas (1996) in that I will include among the object shift languages languages such as Danish, Swedish, and Norwegian, in which only pronominal objects may shift. This group will therefore include all the Germanic languages other than English, while the non-object shift languages will include English and presumably some non-Germanic languages.

The claim made here is that object shift languages may potentially move subjects, objects, and indirect objects into multiple specifiers of a single head in the overt syntax. Suppose we accept the popular claim (Ura 1996, Rudin 1988, Comorovski 1986) that in such a configuration all the multiple specifiers are equally accessible to attraction; none is structurally higher than the others, at least as far as c-commanding attractors are concerned. Then we expect to find evidence that arguments other than the subject are more accessible to attraction in object shift languages than in non-object shift languages.

There does seem to be further evidence that object shift must sometimes be used as an escape hatch for further movement. Haegeman (1993, 1996) gives some evidence along these lines from West Flemish. In West Flemish, as in Dutch, the word order is rigidly S-IO-DO (Haegeman 1996, 150):

33. a. da Valère Marie dienen boek verzekerst nie toogt
that Valere Marie that book probably not shows
b. da Valère Marie verzekerst dienen boek nie toogt
that Valere Marie probably this book not shows
c. da Valère verzekerst Marie dienen boek nie toogt
that Valere probably Marie this book not shows

Thus, object shift of the direct object is apparently only possible if the indirect object also shifts. This follows straightforwardly from the approach given here, in which object shift involves multiple attraction by a single attractor; as always, the highest available mover must be attracted first.

Interestingly, cliticization of the direct object is also only possible if the indirect object shifts (Haegeman 1996, 160):

34. a. da Valère Marie **ze** misschien gegeven eet
that Valere Marie them perhaps given has
b. da Valère **ze** Marie misschien gegeven eet
c. da **ze** Valère Marie misschien gegeven eet
35. a. *da Valère **ze** misschien Marie gegeven eet
that Valere them perhaps Marie given has
b. *da **ze** Valère misschien Marie gegeven eet

That is, for the direct object clitic to undergo further movement to a higher position, it must first enter into a multiple-specifier relation with the indirect object. Similarly, in Dutch (Sjef Barbiers, p.c.), shift of the indirect object is necessary for the direct object to move to the pre-V2 slot[4]:

36. a. De foto heeft Jan de mannen gisteren getoond
the photo has Jan the men yesterday shown
b.??De foto heeft Jan gisteren de mannen getoond

Again, movement of the direct and indirect objects to multiple specifiers of AgrP allows the syntactically lower object to move to a higher position, a familiar situation to devotees of multiple specifiers.

Path Containment Condition effects are another case in which object shift languages might be claimed to show a more "flat" structure than non-object shift languages. PCC effects are claimed to be absent for at least some speakers of Norwegian, Swedish, and Danish (Christensen 1982, Engdahl 1982, Erteschik-Shir 1982):

37. a. Det var Lisa$_i$ læreren spurte hvilken gutt$_j$ vi trodde t$_i$ var sint på t$_j$ (Norw.)
it was Lisa the-teacher asked which boy we thought was mad at
b. Sina föräldrar$_i$ är det lätt att glömma hur mycket$_j$ man är skyldig t$_i$ t$_j$ (Swedish)
self's parents is it easy to forget how much one owes
c. Marie$_i$ ved jeg ikke hvem$_j$ Peter tror t$_i$ elsker t$_j$ (Danish)
Marie know I not whom Peter thinks loves

---

[4] There is apparently a similar effect with wh-movement, but the contrast is not as strong, for reasons I do not understand (Sjef Barbiers, p.c.):
i. Wat heeft Jan de mannen gisteren getoond?
what has Jan the men yesterday shown
ii. ? Wat heeft Jan gisteren de mannen getoond?

Furthermore, it has apparently been claimed (Christensen, p.c. to David Pesetsky, p.c.) that such crossing paths are only available for D-linked phrases; that is, for elements which might in principle undergo object shift. The theory sketched here makes a number of predictions, which I have not yet had a chance to test. One prediction is that the sentences in 39 will become bad again if a clause boundary intervenes between the extraction sites; such a boundary would presumably prevent the arguments from becoming specifiers of a single head. This prediction is borne out in Norwegian, at least (Øystein Vangsnes, p.c.):

38. a. ? Hvilken oppdagelsesreisende$_i$ spurte laereren deg
which explorer asked teacher you
hvilket kontinent$_j$ t$_i$ oppdaget t$_j$?
which continent discovered
b. * Hvilken oppdagelsesreisende$_i$ spurte laereren deg
which explorer asked teacher you
hvilket kontinent$_j$ t$_i$ trodde at han kunne oppdage t$_j$?
which continent thought that he could discover

We might expect to find similar facts with respect to Superiority, and in German we apparently do (Fanselow 1997, Grohmann 1997, Kim and Sternefeld 1997, Wiltischko 1997); this is reminiscent of the claim I defended in Richards (1996) to the effect that languages with overt scrambling lack local Superiority effects.

4. Cliticization

Another case in which multiple movements appear to obligatorily cross is that of cliticization in many languages:

39. (Serbo-Croatian, from Franks 1997)

```
                    a----------l
                a--+--------l1
                ?     ?                 11
```

a. Vesna mu je uvek nudi
Vesna him-DAT it-ACC always offers

(Tagalog)

```
                    a----------l
                a+--------l1
                ??              11
```

b. Nakita niya ako kahapon
saw s/he me yesterday

In the theory developed here, this would indicate that clitics of this type are all syntactically moving to check their features against a single head. Some evidence that something like

this may in fact give the correct structure for the clitic cluster is given in work in progress by Sandra StjepanoviŒ, who notes the following ellipsis possiblities for the clitic cluster in Serbo-Croatian:

40. Ona mu ga je dala,...
she him-DAT it-ACC AUX-3sg gave

a. ...a i ja **sam** **mu** **ga** [dala]
and also I AUX-1sg him-DAT it-ACC gave
b. ...a i ja **sam** **mu** [ga dala]
c. ...a i ja **sam** [mu ga dala]
d.* ...a i ja **sam** [mu] **ga** [dala]

On the assumption that ellipsis targets constituents, these ellipsis data argue for a structure of the clitic cluster like that in 41:

41. 4
clitic 4
clitic 4
clitic

This, of course, is the structure predicted by the approach given here. Many interesting questions now arise, which I am not in a position to address here. Are the clitics in 39-41 in multiple specifiers? What happens in cases in which clitics attach to a head with specifiers? In part, this will depend on the content of the notion "specifier", a question which is not crucial to the theory developed in this paper (despite its title). In a bare theory of phrase structure, we presumably expect the syntactic position of moved elements to follow from independent principles of syntax. We would not be suprised to find, for instance, that clitics are forced for morphosyntactic reasons to "tuck in" to positions below the specifiers in which fully phrasal categories check features. The only requirement of the theory developed here would be that each movement be to as low a landing site as possible.

5. Conclusion

In this paper I have tried to show that Chomsky's (1995) notion of featural cyclicity is to be preferred over earlier versions of cyclicity. Featural cyclicity predicts--correctly, I have argued--that multiple movements will cross rather than nesting just in case their destinations are multiple specifiers of a single head. I have claimed that this is true for both

A- and A-bar movement, which allows us to derive these facts from a straightforward theory based on featural cyclicity and Shortest Move.

Bibliography

Bobaljik, Jonathan. 1995. Morphosyntax: the syntax of verbal inflection. Doctoral dissertation, MIT.

Bobaljik, Jonathan, and Dianne Jonas. 1996. Subject positions and the roles of TP. Linguistic Inquiry.

Branigan, Philip. 1992. Subjects and complementizers. Doctoral dissertation, MIT.

Chomsky, Noam. 1981. Lectures on government and binding. Foris, Dordrecht.

Chomsky, Noam. 1993. A minimalist program for linguistic theory. 1-52 in The view from building 20, edited by Ken Hale and S. Jay Keyser. MIT Press, Cambridge.

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

Collins, Chris, and Höskuldur Thráinsson. 1993. Object shift in double object constructions and the theory of Case. 131-174 in MITWPL 19: Papers on Case and Agreement II, edited by Colin Phillips. MITWPL, Cambridge, MA.

Comorovski, Ileana. 1986. Multiple wh-movement in Romanian. 17:171-177.

Fanselow, Gisbert. 1997. MLC and Strict Cyclicity. Paper presented at the Workshop on the Minimal Link Condition, Potsdam.

Franks, Steven. 1997. South Slavic clitic placement is *still* syntactic. Paper presented at Penn Linguistics Colloquium.

Grohmann, Kleanthes. 1997. German Superiority. ms., University of Maryland.

Haegeman, Liliane. 1996. The typology of syntactic positions: L-relatedness and the A/A'-position. 141-165 in Minimal ideas, edited by Werner Abraham et al. John Benjamins, Amsterdam.

Hasegawa, Nobuko. 1984. On the so-called "zero pronouns" in Japanese. The Linguistic Review 4: 289-341.

Holmberg, Anders, and Christer Platzack. 1995. The role of inflection in Scandinavian syntax. Oxford University Press, Oxford.

Jonas, Dianne, and Jonathan Bobaljik. 1993. Specs for subjects: the role of TP in Icelandic. 59-98 in MITWPL 18: Papers on case and agreement I, edited by Jonathan Bobaljik and Colin Phillips. MITWPL, Cambridge, MA.

Kim, Shin-Sook, and Wolfgang Sternefeld. 1997. Superiority vs. Cross Over. Paper presented at the Workshop on the Minimal Link Condition, Potsdam.

Kitahara, Hisatsugu. 1994. Target-α: a unified theory of movement and structure-building. Doctoral dissertation, Harvard University.

Kuno, Susumu. 1973. The structure of the Japanese language. MIT Press, Cambridge, MA.

Kuroda, Sige-Yuki. 1971. Remarks on the notion of subject with reference to words like also, even, or only, Part II, In Annual Bulletin 4, 127-152. Logopedics and Phoniatrics Research Insitute, University of Tokyo. (Reprinted in Papers in Japanese Linguistics 11, 157-202.)

Mueller, Gereon. 1996. Optimality Theory in Syntax II: Parallel Movement. ms., University of Tübingen.

Mulders, Iris. 1996. Multiple checking and mirrored specifiers: the structure of CP. ms., Utrecht University.

Mulders, Iris. To appear. Mirrored specifiers. To appear in Linguistics in the Netherlands 1997.

Neeleman, Ad. 1994. Scrambling as a D-structure phenomenon. 387-429 in Studies on Scrambling, edited by Norbert Corver and Henk van Riemsdijk. Mouton de Gruyter, Berlin.

Ochi, Masao. 1996. On the nature of relativization in Japanese. Paper presented at CONSOLE 5.

Pesetsky, David. 1982. Paths and categories. Doctoral dissertation, MIT.
Richards, Norvin. 1995. The principle of minimal compliance. ms., MIT.
Richards, Norvin. 1996. Subjacency forever. To appear in Proceedings of CONSOLE 5.
Rudin, Catherine. 1988. On multiple questions and multiple wh-fronting. Natural Language and Linguistic Theory 6, 445-501.
Saito, Mamoru. 1982. Scrambling, topicalization, and strong crossover. ms., MIT.
Ura, Hiroyuki. 1996. Multiple feature-checking: a theory of grammatical function splitting. Doctoral dissertation, MIT.
Wiltschko, Martina. 1997. Superiority in German. Paper presented at WCCFL 16.
Yatsushiro, Kazuko. 1996. On the Unaccusative Construction and Nominative Case Licensing. ms., University of Connecticut.