Subjacency Forever

Norvin Richards, MIT

Baker (1970) notes that sentences of the following form are ambiguous in English:

1. [Who knows [who bought what]]?

Here *what* can have scope either with the *who* in the embedded clause or with the *who* in the matrix clause. The availability of the latter reading is somewhat surprising in light of the ill-formedness of 2:

2. * [What do you know [who bought]]?

That is, the process whereby *what* gets its scope in 1 is apparently immune to the constraint which rules out 2. One can imagine a number of ways in which the contrast between 2 and the wide-scope reading of 1 might be explained; I will concentrate here on two.

The first, which has been fairly well accepted in much of the literature on this problem, would be to say that the LF movement whereby *what* gets its scope in 1 is immune to Subjacency. Let us refer to this as the "levels approach". One argument for this approach comes from Huang (1982), who notes that in Chinese, a language in which wh-movement is apparently not overt, no wh-island effects are observed. The same is true in Tibetan (3.a from Huang 1982, xx; 3.b from Ngawang Jorden, p.c.)[1]:

Many thanks to the following people for their help with the facts of their languages: Roumyana Izvorski and Roumyana Slabakova (Bulgarian), Hooi-Ling Soh and Wei-Tien Dylan Tsai (Chinese), Takako Aikawa, Shigeru Miyagawa, and xxx Fukuda (Japanese), Michael Brody, Julia Horvath, and Anna Szabolcsi (Hungarian), Snjezana Kordic and Milan Mihaljevic (Serbo-Croatian), and Ngawang Jorden (Tibetan). All of these people have been extremely patient with me, and this paper would have been nowhere near as interesting without their help. For more general comments on the paper I would also like to thank Yoo-Kyung Baek, Alec Marantz, Martha McGinnis, David Pesetsky, and Shigeru Miyagawa. Any faults that I have managed to sneak past all these people are purely my own responsibility.

[1] Both of these examples, irrelevantly for our purposes, also have a reading in which both wh-words have embedded scope.

3. a. Ni xiang-zhidao shei mai-le sheme?
you wonder who bought what
'What do you wonder who bought?'

b. Khong-gyis khyedrang-la [su -s gare gzigs-pa]
he ERG you DAT who ERG what buy that
bka''dri- gnang- pa- red?
question do-HON PST Agr
'What did he ask you who bought?'

Another possible way of accounting for the distinction between 1 and 2, argued for in Brody 1995b and Richards 1996, would be to say that Subjacency need only be obeyed once per wh-comp; that is, in 1, the local movement of *who* satisfies Subjacency, rendering all subsequent wh-movements to that site immune to Subjacency. I will refer to this as the "Subjacency Tax approach". Evidence for this approach comes from the contrasts in 4 (Japanese, from Watanabe 1992) and 5 (Bulgarian, from Roumyana Izvorski and Roumyana Slabakova, p.c.):

4. a.?? John-wa [Mary -ga nani -o katta ka dooka]
John TOP Mary NOM what ACC bought whether
siritagatte-iru no?
know-want Q
'What does John want to know whether Mary bought?

b. John-wa [Mary-ga nani -o katta ka dooka]
John TOP Mary NOM what ACC bought whether
dare-ni tazuneta no?
who DAT asked Q
'Who did John ask whether Mary bought what?'

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

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

4.a and 5.a are both Subjacency violations. In the b. sentences we can see that adding an additional wh-word outside the Subjacency island improves the structure in both cases. These are both languages in which all wh-movement is apparently done on a single syntactic level, so the approach outlined above, in which the relevant factor is whether the movement is overt or covert, would seem to have nothing straightforward to say about these cases[2].

Neither account deals easily with the contrast between Japanese and Chinese, shown again in 6:

[2]One might, as Watanabe (1992) does, postulate a class of movements which are overt but invisible to deal with the Japanese facts in 4; on such an account, Subjacency applies to wh-movement in 4.a because *nani* 'what' actually moves overtly, though invisibly. 4.b would then be well-formed because the wh-word which moves overtly is *dare*; *nani* moves covertly in this case (and therefore invisibly). To deal with the facts in 5, one could presumably make a similar claim in reverse, postulating a class of movements which are visible but covert. This would be the kind of movement which applies to *koja kniga* 'which book' in 5.b; Subjacency does not apply here because the movement is covert (although it can be seen). Maneuvers of this kind will allow us to maintain the generalization that all and only overt movements are subject to Subjacency, but carry a certain risk of rendering that generalization vacuous.

6. a.?? John-wa [Mary -ga nani -o katta ka dooka]
John TOP Mary NOM what ACC bought whether
siritagatte-iru no?
know-want Q
'What does John want to know whether Mary bought?
b. Ni xiang-zhidao shei mai-le sheme?
you wonder who bought what
'What do you wonder who bought?'

Languages which do all their wh-movement covertly seem to be divisible into two classes; a class with wh-island effects (including Japanese and Korean) and a class without them (including Chinese and Tibetan). What property accounts for this distinction?

Rudin (1988) notes that languages which do all their wh-movement overtly are also divisible into a class which exhibits wh-island effects (including Serbo-Croatian and Polish) and a class which lacks them (including Bulgarian and Rumanian):

7. (Serbo-Croatian, from Rudin 1988, 459)
a. *Šta si me pitao ko može da uradi?
what have-2s me asked who can to do
'What have you asked me who can do?'
(Bulgarian, from Rudin 1988, 457)
b. ?Koja ot tezi knigi se čudiš koj znae koj prodava?
which of these books wonder-2s who knows who sells
'Which of these books do you wonder who knows who sells?'

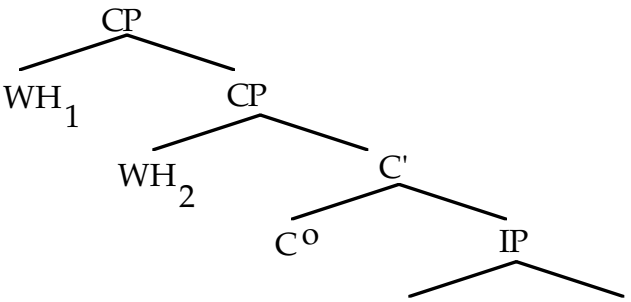
The hypothesis which will be pursued in this paper will be that the difference between the Japanese class and the Chinese class is the same as the difference between the Serbo-Croatian class and the Bulgarian class. We will see that

there are independent reasons for the Chinese class of languages to lack wh-islands, and the levels approach is therefore undermined.

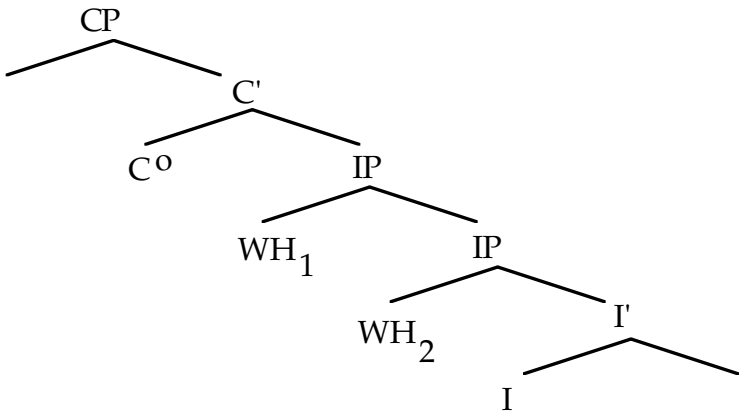
2. CP-Absorption and IP-Absorption

Slightly adapting the proposal of Rudin (1988), I will hypothesize that multiple wh-movement can take place either by movement to multiple specifiers of CP[3], as in 8.a, or by movement to multiple specifiers of an IP projection, as in 8.b.

8. a. "CP-absorption" (Bulgarian, Chinese)



b. "IP-absorption" (Serbo-Croatian, Japanese)



I will refer to the former type as "CP-absorption languages", and to the latter as "IP-absorption languages". Bulgarian and Chinese are CP-absorption languages; Serbo-Croatian and Japanese are IP-absorption languages.

CP-absorption languages have more or less familiar properties; wh-movement is always to a specifier of CP, and is always A-bar movement. IP-

[3] or possibly by multiple adjunction to CP; I have no evidence bearing on this distinction, if indeed such a distinction exists.

absorption languages, on the other hand, have somewhat more exotic properties. Here wh-movement most closely resembles the scrambling found in languages like Hindi (cf. Mahajan 1992) and Japanese (cf. Saito 19xx); local wh-movement has certain properties of A-movement, while long-distance wh-movement uniformly acts like A-bar movement, presumably because A-chains are subject to stricter locality principles. In some IP-absorption languages, a single wh-word apparently moves obligatorily to Spec CP (Serbo-Croatian, for example, appears to be such a language, although Hungarian is not). I will not speculate here on the force driving this move.

In the next section I will give some evidence that Hungarian is an IP-absorption language; this will provide further insight into the nature of IP-absorption. I will then move on to discuss the differences between IP-absorption languages and CP-absorption languages more generally, and try to show that Japanese and Chinese do indeed differ in this regard.

3. Hungarian

Hungarian obligatorily moves all wh-words to a position which is preverbal but follows an overt complementizer (as well as another preverbal position, often occupied by the subject). Compare 9 and 10 (Horvath 1986, 54 and 67):

9. a. Mari az asztalra tette az edényeket
Mary the table-onto put the dishes-ACC
'Mary put the dishes on the table.'
b. *Mari tette az asztalra az edényeket
Mary put the table-onto the dishes-ACC

10. a. Nem tudtuk hogy Mari mit tett az asztalra
not knew-1pl. that Mary what-ACC put the table-onto
'We didn't know what Mary had put on the table'
b. *Nem tudtuk hogy Mari mit az asztalra tett
not knew-1pl. that Mary what-ACC the table-onto put

In 9-10 we can see that the preverbal position, obligatorily occupied by *az asztalra* 'onto the table' in 9, is obligatorily occupied by the wh-word *mit* 'what-ACC' in 10; furthermore, this position is to the right of the complementizer *hogy* 'that'. In multiple interrogation all the wh-words move to this position (from Kiss 1994, 38):

11. Mari kinek mit adott el
Mary who-DAT what-ACC sold PREV
'What did Mary sell to whom?'

Hungarian thus appears to be an IP-absorption language, although it apparently differs from the IP-absorption languages treated in Rudin (1988); in those languages, a single wh-word must apparently always occupy a [+wh] Spec CP, while the other wh-words are adjoined to IP. Hungarian wh-words seem to simply adjoin to IP. I will not attempt to investigate this difference here; I assume it has to do with the strength of the wh-feature on C^{o}.

There is evidence, however, that Hungarian long-distance movement takes place via Spec CP. Local wh-movement in Hungarian triggers

inversion, causing a class of particles which are usually preverbal to follow the verb:

(adapted from Kiss 1994, 21, 37)

12. a. Mari felhívta Jánost
Mary PREV-called John
'Mary called up John'

b. Ki hívta fel Jánost?
who called PREV John
'Who called up John?'

The mechanism driving this is unimportant for our purposes (see Horvath 1985, Kiss 1994, Brody 1995a for discussion). Interestingly, long-distance wh-movement fails to trigger inversion on verbs along the path of movement; in 13, the particles *ki* and *be* remain in their preverbal positions:

13. János melyik fiúnak$_i$ gondolta hogy Péter kijelentette
John which boy-to thought that Peter PREV-reported
hogy a házigazda már bemutatta Marit t_i?
that the host already PREV-showed Mary-ACC
'To which boy did John think Peter declared that the host had already introduced Mary?'

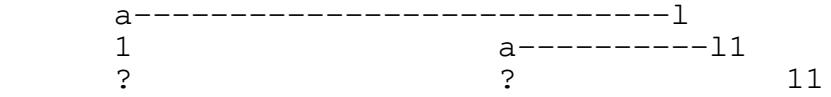
This suggests that long-distance movement does not involve successive-cyclic movement to the landing site used for local wh-movement (as it does in English, for instance). Rather, as Horvath (1985) points out, long-distance wh-movement apparently uses some other landing site as an escape hatch, possibly Spec CP. Note that the final landing site for the long-distance move is apparently still an IP-adjoined position, as the wh-phrase *melyik fiúnak* 'to which boy' follows *János* 'John'.

The distribution of wh-island effects in Hungarian provides further support for this conclusion, as Horvath (1985) notes. Recall that the IP-adjunction languages discussed in Rudin (1988) have an additional requirement, not found in Hungarian, that a single wh-word must move overtly into any +wh Spec CP. These languages exhibit wh-island effects (Serbo-Croatian, from Rudin 1988, 459):

14. *Šta si me pitao ko može da uradi?
what have-2s me asked who can to do
'What have you asked me who can do?'

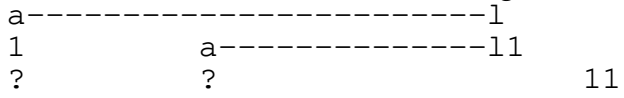
This is to be expected; *ko* 'who' occupies the embedded Spec CP, forming a wh-island.

Interestingly, local wh-movement in Hungarian does not create wh-islands:



15. Mari kinek$_i$ tudta [hogy Péter mit$_j$ küldött t_j t_i?]
Mary who-to knew that Peter what-ACC sent
'To whom did Mary know what Peter had sent?'

Here local movement of *mit* 'what' does not create a wh-island for long movement of *kinek* 'to whom'. Again, this is not surprising if local wh-movement in Hungarian is always IP-adjunction. Note, however, that long wh-movement does create wh-islands in Hungarian:



16. * Mari kinek$_i$ mit$_j$ tudta [hogy Péter küldött t_j t_i?]
Mary who-to what-ACC knew that Peter sent
'To whom did Mary know that Peter had sent what?'

Here long-distance movement of both *mit* and *kinek* is impossible. This suggests, again, that long-distance movement uses Spec CP as an escape hatch even in Hungarian, and that Hungarian, like Serbo-Croatian, projects only a single Spec CP.

4. Diagnostics for CP- and IP-absorption

4.1 Wh-islands

First, let us consider how the posited structures for CP- and IP-absorption languages account for the distribution of wh-island effects. We have seen that some IP-absorption languages make use of Spec CP as a landing site for wh-movement; Serbo-Croatian apparently requires one wh-word to raise to Spec CP. We saw evidence in the last section suggestion that Spec CP is necessarily an escape hatch for long-distance wh-movement, even in IP-absorption languages such as Hungarian.

Now we are in a position to give an account of wh-islands, essentially following Rudin (1988) and Comorovski (1986). Suppose that wh-movement past a filled Spec CP is universally barred, for familiar reasons having to do with considerations of Shortest Move. The only languages which will allow wh-movement out of a question, then, will be ones in which CP can have multiple specifiers, so that wh-movement need never skip a CP projection; there will always be a specifier of CP available as an escape hatch. In IP-absorption languages, on the other hand, it is IP which has multiple specifiers, and CP has only one. A single wh-word which has been forced to move to Spec CP, then, blocks further wh-movement past that specifier position. Thus, IP-absorption languages should exhibit wh-island effects whenever a single element occupies Spec CP, while CP-absorption languages should lack such effects.

4.2 Scrambling

The IP-absorption languages all exhibit a form of local scrambling which fixes weak crossover violations:

(Serbo-Croatian, from Milan Mihaljevic)

17. a. ??Njegov$_i$ susjedi ne vjeruju nijednom politicaru$_i$
his neighbors not trust no politician
'His neighbors trust no politician'

b. Nijednom politicaru$_i$ njegov$_i$ susjedi ne vjeruju t$_i$
no politician his neighbors not trust

(Japanese, from Saito 1992, 73)

18. a. ?*Soitu$_i$-no hahaoya-ga dare$_i$-o aisiteiru no?
guy GEN mother NOM who ACC loves Q
'Who does his mother love?'

b. ?Dare$_i$-o soitu$_i$-no hahaoya-ga aisiteru no?
who ACC guy GEN mother NOM loves Q

(Hungarian, from Kiss 1994, 22)

19. a. *Nem szeret az pro$_i$ anyja mindenkit$_i$
not loves the mother-his everybody-ACC
'His mother does not love everybody'

b. Nem szeret mindenkit$_i$ az pro$_i$ anyja
not loves everybody-ACC the mother-his

CP-absorption languages, on the other hand, apparently lack such a form of scrambling; scrambling is either absent entirely or is A-bar movement:

(Bulgarian, from Roumyana Slabakova)

20. a. *Majka mu obicha vseki chovek
mother his love every person
'His$_i$ mother loves everyone$_i$'

b. *Vseki chovek obicha majka mu
every person love mother his

(Chinese, from Hooi-Ling Soh)

21. a. *Tade mama ai meigeren
his mother love everyone
'His$_i$ mother loves everyone$_i$'

b. *Meigeren tade mama ai
everyone his mother love

(Tibetan)

The pattern, then, seems to be that all and only languages which allow local A-scrambling are IP-absorption languages; this is true regardless of the level on which wh-movement occurs. On the assumption that both IP-absorption and local A-scrambling involve either adjunction to IP or movement into multiple specifiers of IP, this result has a certain intuitive appeal; if a language allows this kind of movement, it uses it both for scrambling and for wh-movement, and if not, neither scrambling nor IP-absorption will be found.

4.3 Superiority

Rudin (1988) notes that in languages like Bulgarian (CP-absorption languages, in this theory), the ordering of fronted wh-phrases is subject to a rigid ordering, which she attributes to Superiority:

22. a. Koj kogo e vidjal?
who whom is seen
'Who saw whom?'
b. *Kogo koj e vidjal?

23. a. Koj kude udari Ivan?
who where hit Ivan
'Who hit Ivan where?'
b. *Kude koj udari Ivan?
where who hit Ivan

In an IP-absorption language like Serbo-Croatian, on the other hand, Superiority effects obtain only for long-distance movement, and not for local movement, as Boskovic (1995) points out:

24. a. Ko je koga vidjeo?
who is whom seen
'Who saw whom?'
b. Koga je ko vidjeo?

25. a. Ko si koga tvrdio da je istukao?
who are whom claimed that is beaten
'Who did you claim beat whom?'
b. *Koga si ko tvrdio da je istukao?
whom are who claimed that is beaten

The sense in which these restrictions on ordering may be attributed to Superiority is not a straightforward one, but I will defer discussing this matter until section 5. For the time being, it is sufficient to note that the differences between Bulgarian and Serbo-Croatian are accounted for by the theory developed here, assuming that Superiority constrains A-bar movement but not A-movement. All Bulgarian wh-movements are A-bar movements,

being movements to Spec CP; in Serbo-Croatian, on the other hand, wh-movement is adjunction to an IP-level projection, and may be an A-movement if it is sufficiently local. The lack of Superiority effects for local movement in Serbo-Croatian (that is, in IP-absorption languages) therefore follows.

Interestingly, a similar asymmetry between local and long-distance movement seems to be present in the LF-moving IP-absorption languages. Japanese Anti-superiority, like the Superiority effects in Serbo-Croatian, is stronger with long-distance movement than it is with local movement (Shigeru Miyagawa, xxx Fukuda, p.c.):

26. a. John -ga nani -o naze katta no?
John NOM what ACC why bought Q
'What did John buy why?'

b. ?John-ga naze nani -o katta no?
John NOM why what-ACC bought Q

27. a. Mary -ga [John -ga nani -o naze katta to] omotteiru no?
Mary NOM John NOM what ACC why bought that thinks Q
'What did Mary think John bought why?'

b. * Mary-ga [John -ga naze nani -o katta to] omotteiru no?
Mary NOM John NOM why what ACC bought that thinks Q

Thus, the Serbo-Croatian and Japanese equivalents of Superiority seem to behave similarly, in that they constrain only long-distance movement; according to the story developed here, this is because only long-distance movement has A'-properties in these languages. The prediction of this account would be that Chinese Superiority, like Bulgarian Superiority, would be equally strong locally and long-distance. Chinese word order is too rigid to test this; no alternatives parallel to those in 26-27 can be constructed. In

Tibetan, on the other hand, another LF-moving CP-absorption language, scrambling is possible, and we find strong local Superiority effects (Ngawang Jorden, p.c.):

28. a. Bkrashis-lags-gyis yag garebyadnas gzigs-gnang-pa- red?
Tashi HON ERG yak why buy-HON-PAST-AGR
'Why did Tashi buy a yak?'

b. Bkrashis-lags -gyis garebyadnas yag gzigs-gnang-pa -red
Tashi HON ERG why yak buy HON PAST AGR

29. a. Bkrashis-lags -gyis gagi garebyadnas gzigs-gnang -pa -red
Tashi HON ERG which why buy HON PAST AGR
'Why did Tashi buy what?'

b. *Bkrashis-lags -gyis garebyadnas gagi gzigs-gnang-pa -red
Tashi HON ERG why which buy HON PAST AGR

28.a-b show that scrambling of *garebyadnas* 'why' over the direct object is possible in principle, but 29.a-b show that it is impossible if the direct object is itself a wh-word. Thus, the Tibetan equivalent of Japanese Anti-superiority strongly constrains local movement, as we expect on the hypothesis that Tibetan is like Bulgarian in that all wh-movement is A-bar movement to a Spec CP position[4].

4.4 Weak Crossover

Another asymmetry between local and long-distance movement in IP-absorption languages appears in the domain of weak crossover. CP-absorption languages like Bulgarian have weak crossover effects of a fairly familiar kind (Roumyana Slabakova, p.c.):

[4] (we also expect that xx.b won't be any better than the equivalent for LD mvmt--I haven't had a chance to check this yet).

30. a. Koj obicha majka si?

who loves mother his

'Who$_i$ loves his$_i$ mother?'

b. *Kogo obicha majka su?

who loves mother his

'Who$_i$ does his$_i$ mother love?'

In IP-absorption languages like Hungarian and Serbo-Croatian, on the other hand, weak crossover effects are found only long-distance, not locally: (Hungarian from Kiss 1987, 208, and Brody 1995a; Serbo-Croatian from Snjezana Kordic):

31. a. Ki szereti az anyjat?

who loves the mother-his-ACC

'Who$_i$ loves his$_i$ mother?'

b. Kit szeret az anyja?

who-ACC loves the mother-his

'Who$_i$ does his$_i$ mother love?'

c. *Kit gondol az anyja hogy Mari szeret?

who-ACC thinks the mother-his that Mary loves

'Who$_i$ does his$_i$ mother think that Mary loves?'

32. a. Tko voli svoju majku?
who loves his-ACC mother-ACC
'Who$_i$ loves his$_i$ mother?'

b. Koga voli njegova majka?
who loves his-NOM mother-NOM
'Who$_i$ does his$_i$ mother love?'

c. *Koga njegova majka misli da Marija voli?
who his-NOM mother-NOM thinks that Maria loves
'Who$_i$ does his$_i$ mother think that Mary loves?'

A surprising fact, given the theory developed here, is the presence of weak crossover effects in IP-absorption languages like Japanese (Saito 1992, 73):

33. ?*Soitu$_i$-no hahaoya-ga dare$_i$-o aisiteru no?
guy GEN mother NOM who ACC love Q
'Who does his mother love?'

It is not clear why *dare* cannot adjoin to IP at LF in a position higher than the pronominal variable it binds, thus obviating the weak crossover violation. One possible answer will be outlined in the next section.

4.4 Wh-movement and QR

IP-absorption, as developed here, is syntactically reminiscent of QR, in that it involves multiple adjunction to IP in order to establish scope relations. In some languages in which IP-adjunction occurs overtly, movement does indeed seem to have effects on scope relations:

(Hungarian, from Kiss 1994, 71)

34. a. Mindenki két lányt is meg táncoltatott
everyone two girl-ACC even PREV danced
'Everybody danced with two (potentially different) girls'

b. Két lányt is mindenki meg táncoltatott
two girl-ACC even everyone PREV danced
'Two girls (the same two girls) were danced with by everybody'

(Japanese, adapted from Sauerland 1996, 21)

35. a. Dareka-ga daremo-o aisiteru
someone-NOM everyone-ACC loves
'Someone loves everyone' (∃>>∀, *∀>>∃)

b. Daremo$_i$-o dareka-ga t_i aisiteru
everyone-ACC someone-NOM loves
'Someone loves everyone' (∃>>∀, ∀>>∃)

Furthermore, IP-adjoined wh-words in Hungarian demonstrably occupy a position which is also used as a landing site by a certain class of quantificational elements; such quantifiers have their scopes determined by movement to this position (see Kiss 1987, 1994 for discussion). In 36 we see that both wh-movement and this form of overt QR trigger inversion of the verb with a preverb, a standard test for occupying this position (from Kiss 1994 (37, 64)):

36. a. Ki hívta fel Jánost?
who called PREV John-ACC
'Who called up John?'

b. János kevés fogást kóstolt meg
John few dish-ACC tasted PREV
'John tasted few dishes'

Japanese and Chinese are both "rigid scope" languages; the scope of quantifiers is apparently entirely determined by their surface position, so that 37.a-b are both unambiguous, with the subject QP taking scope over the object QP.

(Chinese & Japanese, from Aoun and Li 1993, 365)

37. a. (Yaoshi) yige ren piping meigeren...
if one man criticize everyone
'(If) someone criticized everyone...'

b. Dareka -ga daremo -o semeta
someone-NOM everyone-ACC criticized
'Someone criticized everyone'

According to the theory developed here, Chinese and Japanese differ in that Japanese uses the same syntactic mechanism, IP-adjunction, to assign scope to quantifiers and to wh-words, while Chinese uses two different syntactic mechanisms: IP-adjunction and substitution to Spec CP. Interestingly, the "rigid scope" property of quantifiers is extended to wh-words in Japanese, but not in Chinese, as Aoun and Li (1993) point out; the ill-formedness of 38.b may be attributed to the inability of the wh-word to take scope over the other operators in the sentence:

38. a. Meigeren dou maile shenme?
everyone all bought what
'What did everyone buy?'

b. *Daremo -ga nani -o kaimasita ka?
everyone NOM what ACC bought Q
'What did everyone buy?'

Although I have no account of "rigid scope" to offer, it seems clear that the theory developed here makes the difference between Chinese and Japanese

look less surprising; the generalization, apparently, is that LF IP-adjunction in these languages cannot result in a change of scope relations.

Rigid scope might also be responsible for the ill-formedness of 33, repeated as 39:

39. ?*Soitu$_i$-no hahaoya-ga dare$_i$-o aisiteru no?

guy GEN mother NOM who ACC love Q

'Who does his mother love?'

Whatever our eventual account of scope rigidity might be, we might expect it to say that *dare* is unable to bind any variables at LF which it cannot bind in the overt syntax; the surprising ill-formedness of 39 would then follow.

5. Appendix: Superiority and Anti-Superiority

In the course of this paper we have noted a number of effects of the general Superiority/Anti-Superiority type. The claim developed here has been that A-bar movements are generally subject to either Superiority or Anti-Superiority; local wh-movements in IP-absorption languages, not being A-bar movements, are free from these constraints, but all other wh-movements must in principle obey them. The discussion so far has essentially conflated the notions of Superiority and Anti-Superiority. In this section I will speculate briefly about the nature of these phenomena.

The Superiority paradigm is well-known from English; the generalization appears to be that an attractor must attract the highest available element first, and this generalization seems to follow more or less straightforwardly from Shortest Move:

40. a. Who t bought what?

b. ?? What did who buy t?

The Anti-Superiority paradigm was noted in Japanese by Saito (1982) and Watanabe (1992), and is exemplified by the following contrast:

41. 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?'

These facts can be explained, as Saito and Watanabe point out, if Japanese is subject to some version of the ECP which requires adjuncts to move to Spec CP before arguments, and if the order of movement is the reverse of the English pattern; that is, if the lower of the two wh-words moves first. In 41.a, then, *naze* 'why', which is c-commanded by *nani* 'what', raises first, obeying the ECP, while in 41.b *nani* 'what' is c-commanded by *naze* 'why' and therefore moves first, violating the ECP.

Interestingly, the ordering of multiple overtly moved wh-words also seems to obey Anti-Superiority (Bulgarian, from Boskovic 1995, 3):

42. a. Koj kogo e vidjal
who whom is seen
'Who saw whom?'

b. *Kogo koj e vidjal
whom who is seen

On the standard assumption that adjunction to CP ought to proceed leftwards, the order in 42.a can only be produced by moving *kogo* 'whom' before *koj* 'who'; that is, by moving the lower of the wh-words first, as in Japanese, though not in English.

Thus far, then, the evidence seems to suggest that Anti-Superiority is a result of multiple attraction by a given attractor on a single level; multiple overt movement and multiple covert movement both exhibit Anti-

Superiority, while multiple movement which is partly covert and partly overt, as in English, exhibits Superiority effects. Evidence suggesting that this is the correct generalization can in fact be found in Japanese.

Takahashi (1993) notes that long-distance scrambling of wh-words in Japanese obeys Superiority, rather than Anti-Superiority (from Takahashi 1993, 664):

43. 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?

Only the higher of the two wh-words may undergo long-distance scrambling; long-distance scrambling of a single wh-word, in other words, obeys Superiority, rather than Anti-Superiority. This is not simply a fact about long-distance scrambling, since multiple long-distance scrambling to a single position exhibits Anti-Superiority again (Takako Aikawa, p.c.):

44. a. John -ga [dare -ga nani -o katta to] omotteiru no?
John NOM who NOM what ACC bought that thinks Q
'Who does John think bought what?'

b. Dare -ga nani -o John -ga [t t katta to] omotteiru no?
who NOM what ACC John NOM bought that thinks Q

c. *Nani -o dare -ga John -ga [t t katta to] omotteiru no?
what ACC who NOM John NOM bought that thinks Q

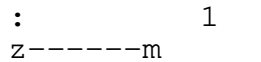
Thus, the distribution of Anti-Superiority and Superiority is not simply a matter of a "Superiority parameter" which makes Bulgarian and Japanese "Anti-Superiority languages" and English a "Superiority language". Rather, the choice between Superiority and Anti-Superiority appears to be a property of individual constructions.

The generalization proposed above, then, seems to hold; Anti-Superiority is a property of multiple attraction to a single position on a single level, while Superiority is a property of multiple attraction on different levels. What could account for this generalization? I will sketch two possible explanations here; I hope to spend the next semester or so trying to choose between these.

5.1 Maintaining c-command relations

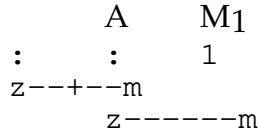
Let us consider the derivations which obey Superiority and Anti-Superiority. In particular, let us focus our attention on the first level on which attraction takes place; on this level, a single wh-word is moved in the Superiority case, while multiple wh-words are moved in the Anti-Superiority case. Schematically, the paths are as follows (here A symbolizes an attractor and M a moved element):

45. a.



(Superiority)

b.



(Anti-Superiority)

One thing these derivations have in common is that they both result in the c-command relation between the moved elements being the same after the

derivational step is taken;[5] M_1 and M_2 c-command each other before and after movement[6]. Of course, this ceases to be true in the Superiority case on the next level of representation, after the second move is taken. This is presumably not a problem, given the reasonable assumption that the grammar lacks look-ahead properties; at the first point at which movement takes place, it preserves c-command relations, although this results in an unavoidable failure to preserve c-command relations on a later level of representation.

Suppose, then, that the generalization underlying both types of Superiority is that an attractor must operate at all times in a way which maximally preserves c-command relations among the elements in the tree. We might see this as a form of Structural Economy, requiring that existing c-command relations be retained whenever possible. This generalization could conceivably be expanded to cover Shortest Move. The shortest possible move would be measured not in terms of nodes crossed, but in terms of how many c-command relations are altered; a longer move will always result in the moved element c-commanding more new positions than it would if it had moved to a closer position.

5.2 "Tucking in"; Shortest Move

Let us consider again an Anti-Superiority case like 42 (repeated as 46):

---

[5] this can presumably be rephrased in a way which makes more sense in a copy theory of movement, but I won't try to do so here.

[6] Here I assume, contra much interesting work (McGinnis, Kayne?...) that M_1 can c-command M_2 if M_1 and M_2 are in multiple specifiers.

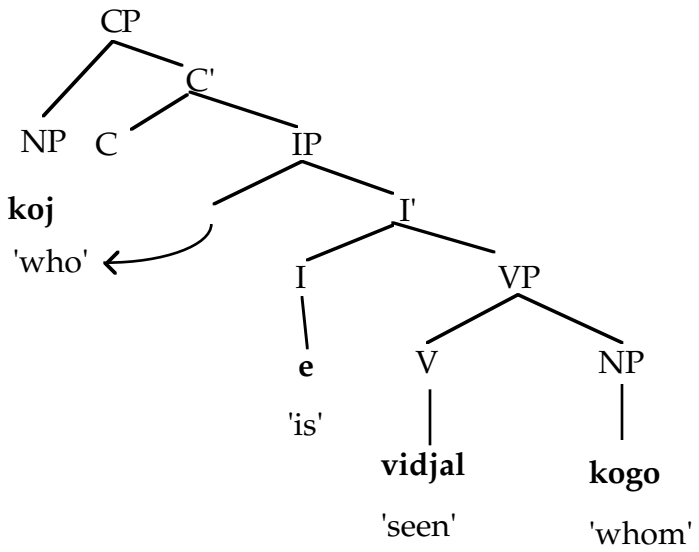
46. a. Koj kogo e vidjal
who whom is seen
'Who saw whom?'

b. *Kogo koj e vidjal
whom who is seen

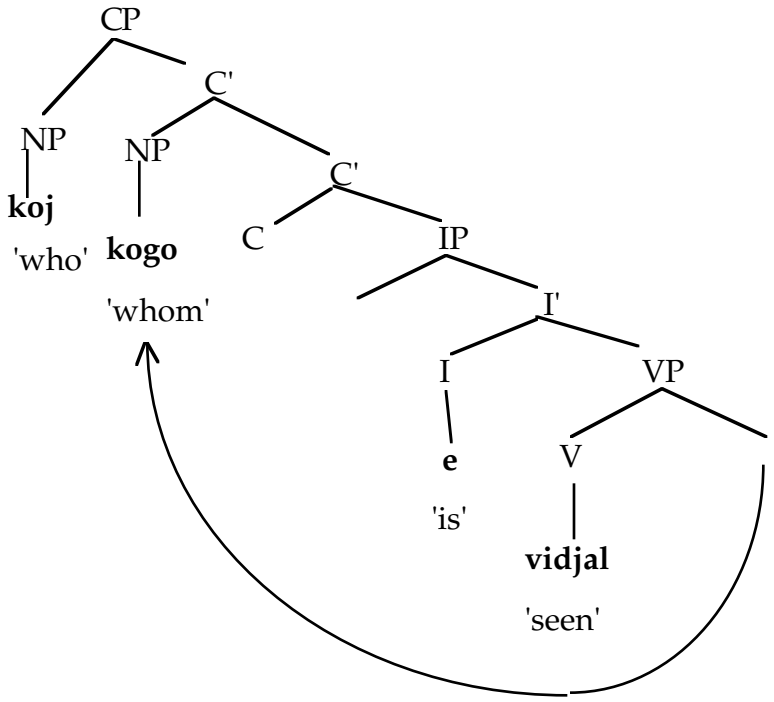
The observation was made above that on standard assumptions about how multiple adjunction works, the order of the wh-words in 46 suggests that the first wh-word to raise is the direct object, followed by the subject; that is, the order is the opposite of that found in English. We have seen that this seems to be the case generally when a given attractor triggers multiple movements on a single level.

One possible account of the facts of 46 would be to deny the standard claim that adjunction always expands the tree. We might claim that in 46.a, *koj* 'who' moves to Spec CP first, followed by movement of *kogo* to a lower Spec CP:

47. a.



b.



In fact, there is some evidence suggesting that this is in fact the way in which the derivation proceeds. Recall the Bulgarian examples in 5, repeated as 48:

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

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

I suggested that these examples provided evidence for the Subjacency Tax theory over the Levels theory. In 5.b, a well-formed wh-movement redeems the ill-formedness of a wh-movement out of a complex NP, despite the fact that both movements are in the overt syntax. This suggests that Subjacency, if satisfied by the first wh-word to move to a given Spec CP, need not be satisfied by other movements to that position. This generalization also covers the well-known English contrast in 49:

49. a. *What do you wonder who bought t?

b. Who t wonders who bought what?

49.b has a wide-scope reading for *what*, as Baker (1970) points out, which involves a wh-island violation of the type seen in 49.a. On the account sketched here (and further developed in Brody (1995b) and Richards (to appear)), this can be seen as the same fact as the Bulgarian facts shown above in 48; the well-formed move of *who* to the matrix Spec CP in 49.b licenses the Subjacency-violating move of *what* to that position at LF. The fact that the second move is at LF is irrelevant, as the Bulgarian examples show, and we have seen further evidence suggesting that Subjacency constraints LF movement in the course of this paper.

Note that this phenomenon is sensitive to the order in which operations occur; the Subjacency-obeying move must precede the Subjacency-violating move[7], as shown in 50:

50. 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?

50.a is a well-formed case in which *who* licenses later movement of *which car* to the matrix Spec CP. 50.b represents a derivation in which the ill-formed movement of *which car* out of the complex NP precedes the well-formed move of *which hubcaps*, and the result is much worse. We can therefore use this phenomenon as a diagnostic for the order of wh-movements.

Looking back at the Bulgarian examples again, what we see is that multiple wh-movement does indeed seem to follow a kind of "tucking in" strategy of the kind outlined above:

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

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

Note that the well-formed wh-movement in 51.b leaves the wh-phrase *koj senator* 'which senator' to the left of the other wh-phrase. If the principle that the well-formed movement must precede the ill-formed movement is operative here, then we have evidence that multiple wh-movement involves tucking in.

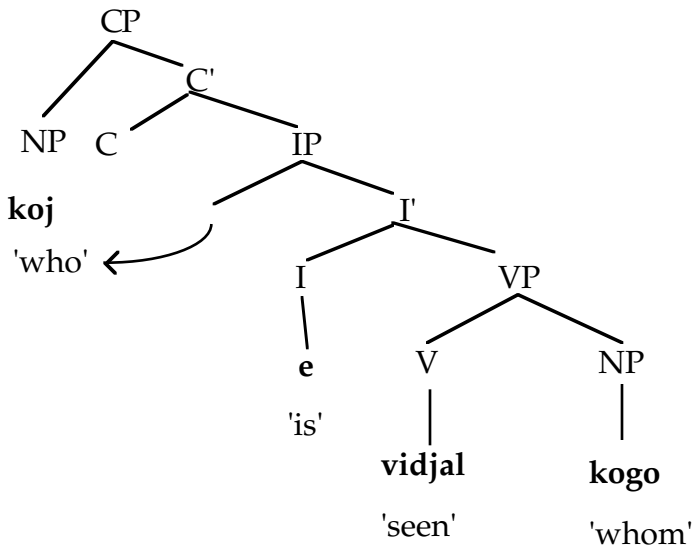
---

[7] assuming, as is standard but by no means universal, that covert movements follow overt movements in the derivation.

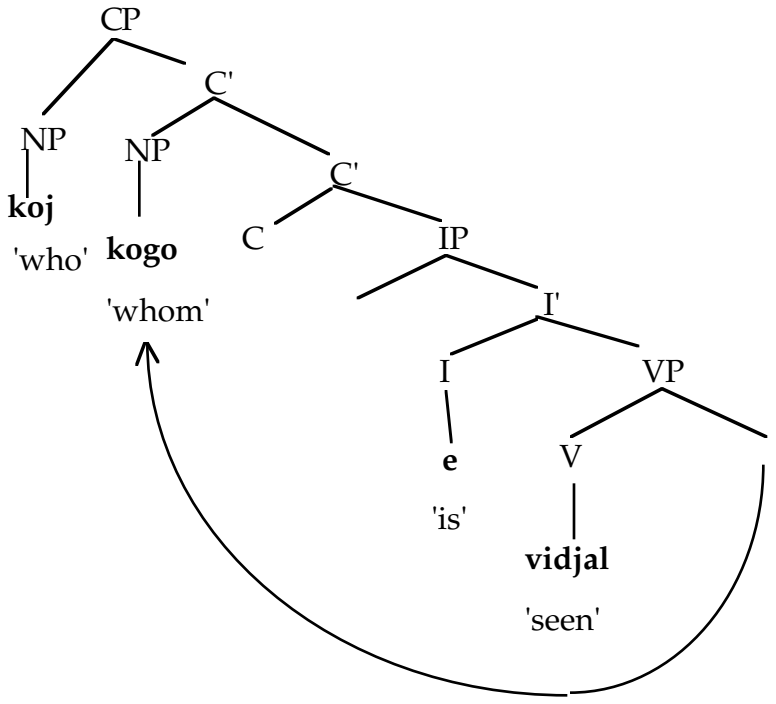
If this is true, several questions arise; first, why is tucking in possible, given that it appears to be countercyclic? Second, why is it necessary? The answer to the first question might be found in the version of cyclicity developed in Chomsky (1996, class lectures). He argues that cyclicity consists not of a requirement that every operation expand the tree but of a requirement that that every operation involve a feature on the head which projects the maximal projection at the root of the tree. If this is correct, then tucking in movement is not countercyclic; all the wh-movements involve feature-checking with the same C^0, so they all occur on the same "cycle". The fact that they do not expand the tree is no longer relevant.

The answer to the second question is presumably a straightforward application of Shortest Move. Consider again the derivation in 47, repeated as 52:

52. a.



b.



In 52.a, C^{o} attracts the nearest wh-word, as required by Shortest Move, and moves it to its specifier. In 52.b, the C^{o} attracts the next wh-word, again obeying Shortest Move, and must again move it to a specifier. We have seen that Cyclicity allows, in principle, movement to either side of the existing

specifier. However, we might reasonably expect Shortest Move to force movement to an "internal" specifier, since this move is shorter than one to an "external" specifier. And in fact this expectation is well-founded; tucking-in movement is forced, as we have seen. Shortest Move can therefore derive both Superiority and Anti-Superiority[8], as desired.

References

Aoun, Joseph, and Yen-hui Audrey Li. 1993. On some differences between Chinese and Japanese Wh-elements. *Linguistic Inquiry*24.365-372.

Bos kovic, Zeljko. 1995. Superiority effects with multiple wh-fronting in Serbo-Croatian. ms., University of Connecticut.

Brody, Michael. 1995a. Focus in Hungarian and bare checking theory. *Tübingen Working Papers in Linguistics*.

Brody, Michael. 1995b. *Lexico-logical Form: a radically minimalist theory*. Cambridge: MIT Press.

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

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

Horvath, Julia. 1986. *FOCUS in the theory of grammar and the syntax of Hungarian*. Foris: Dordrecht.

Huang, C.-T. James. 1982. Logical relations in Chinese and the theory of grammar. Doctoral dissertation, MIT.

Kiss, Katalin E. 1987. *Configurationality in Hungarian*. Reidel: Dordrecht.

---

[8] Actually, Saito and Watanabe's account of Japanese anti-superiority no longer follows, since the claim is that the highest wh-word *always* moves first; we can no longer have an ECP-based account of the kind they sketch. But the ECP is pretty disreputable these days anyway. Hopefully I can come up with something else; one can imagine a representational view in which *why* has to end up as the structurally lowest wh-word, for some reason (say, so it's the nearest potential binder for its trace).

Kiss, Katalin E. 1994. Sentence structure and word order. Ferenc Kiefer and Katalin E. Kiss, ed.s, *Syntax and semantics 27: The syntactic structure of Hungarian*, 1-90. Academic Press: San Diego.

Mahajan, Anoop. 1992. The A/A-bar distinction and movement theory. Doctoral dissertation, MIT.

Marácz, László. 1989. Asymmetries in Hungarian. Doctoral dissertation, Rijksuniversiteit Groningen.

Richards, Norvin. To appear. The Principle of Minimal Compliance. To appear in *Proceedings of CONSOLE 7*.

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.

Saito, Mamoru. 1992. Long-distance scrambling in Japanese. *Journal of East Asian Linguistics* 1.69-118.

Sauerland, Uli. 1996. The interpretability of scrambling. ms., MIT.

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

Watanabe, Akira. 1992. Subjacency and S-structure movement of wh-in-situ. *Journal of East Asian Linguistics* 1:255-291.