Sometimes in SpecCP, sometimes in-situ[1]

Željko Bošković

University of Connecticut

1. Introduction

Languages differ systematically with respect to how they treat wh-phrases in multiple questions. In English, one and only one wh-phrase undergoes fronting in multiple questions.

(1) Who did John give what?

In Slavic languages, on the other hand, all wh-phrases undergo fronting, though they do not all land in SpecCP in all Slavic languages. (See Rudin 1988. In Bulgarian, from which the example in (2) is drawn, all fronted wh-phrases are located in SpecCP).

(2) Koj kakvo e kupil?
    who what is bought

Languages such as Japanese are the exact opposite of Slavic: they allow all wh-phrases to remain in-situ in multiple questions:

(3) John-wa dare-ni   nani-o   ageta ka?
    John-top who-dat what-acc gave   Q

Finally, there are 'mixed' languages such as French: French displays both the English and the Japanese pattern. Thus, both (4a), where one wh-phrase moves to SpecCP, and (4b), where all wh-phrases remain in-situ, are good as true non-echo multiple questions. The Bulgarian pattern, on the other hand, is not attested in French, as illustrated by (4c).

(4) a. Qu' a-t-il donné à qui?

   what has-he given to whom

    'What did he give to whom?'

   b. Il a   donné quoi à qui?

   he has given what to whom

   c. *Qu'à qui a-t-il donné?

In this paper I will examine wh-constructions in French. It turns out that the English and the Japanese pattern are not always both available in French. In fact, whereas the English pattern is always available, the distribution of the Japanese pattern is rather limited in French, which suggests that the two languages employ different strategies in in-situ questions. The goal of this paper is to examine when exactly the Japanese pattern is allowed in French and to account for the limited distribution of the in-situ strategy in this language. The analysis proposed below will be shown to have interesting theoretical consequences. It will provide evidence that lexical insertion can take place in LF. In fact, even elements with strong features will be shown to be capable of entering structure in LF under certain well-defined conditions. The analysis will also provide evidence that feature-movement is subject to locality conditions on movement. Consequences of the proposed analysis of French for the Principles of Economy, as well as pure syntactic wh-movement languages such as English and pure wh-in-situ languages such as Japanese, will also be examined.

2. Embedded questions in French

As noted above, in simple matrix questions French allows both the English and the Japanese pattern. Thus, both (5a) and (5b) are grammatical. It is well-known, however, that the Japanese pattern is disallowed in embedded questions. Whereas (6a) is acceptable, (6b) is not and sharply contrasts with (5b) and Japanese (7).

(5) a. Qui   as-tu   vu?

     whom have-you seen

   b. Tu as vu qui?

(6) a. Pierre a   demandé qui   tu   as   vu

         Pierre has asked   whom you have seen

   b. *Pierre a demandé tu as vu qui

(7) Peter-wa [anata-ga dare-o   mita ka] tazuneta

     Peter-top you-nom who-acc saw   Q   asked

The contrast between (5b) and (6b), which has always resisted a satisfactory account (see here n. 5), follows given certain minimalist assumptions in Chomsky (1995). In particular, the contrast follows from the way lexical insertion is performed in the minimalist system and Chomsky's (1995) definition of strong features. In the minimalist system, which dispenses with D-structure, lexical insertion takes place through the operation Merge, which is defined in such a way that it always must expand structure, i.e., make the existing tree bigger. As a result, Merge can take place only at the root of a tree; it cannot take place in embedded positions. In non-minimalist terms, it is subject to the cycle. Merger generally takes place in overt syntax. As noted in Chomsky (1995), this follows without stipulation. Thus, if an NP such as *John* is inserted in LF the derivation crashes because LF cannot interpret phonological features of *John*. If, on the other hand, *John* is inserted in PF, PF will not know how to interpret the semantic features of *John.* The only way to derive a legitimate PF and a legitimate LF is for *John* to be inserted before the level of SS is reached. PF will then strip off the phonological features of *John* and the semantic features of *John* will proceed into LF. This line of reasoning allows lexical insertion to take place in PF and LF under certain conditions. To be more precise, it allows PF insertion of semantically null lexical elements and it allows LF insertion of phonologically null elements. We are interested in this second possibility here. (Chomsky actually suggests that PF insertion should be altogether banned without giving empirical justification for this move. The *do* of *do-*support, however, appears to be a plausible candidate for PF insertion.)

        Returning now to (5b), notice that the complementizer in (5b) is phonologically null.

Since wh-movement must take place overtly at least in some constructions in French I assume that the +wh feature is strong in French. Suppose now that we insert the complementizer with strong +wh-feature in the LF of (5b). In Chomsky's (1994) system this would not be possible. Chomsky (1994) considers strong features illegitimate LF objects. As a result, a derivation that enters LF with a strong feature necessarily crashes. This is, however, not the case under Chomsky's (1995) approach to strong features, which I will adopt here.[2] Chomsky (1995) defines strong features derivationally. In particular, he defines strong features as features that are not tolerated by the derivation and therefore must be eliminated via checking as soon as they are inserted into the tree. This conception of strong features allows insertion of strong features in LF as long as they are checked off immediately upon insertion. Nothing, then, seems to prevent the complementizer with a strong +wh-feature from being inserted in the LF of (5b).[3] Once the complementizer is inserted, the wh-phrase moves into the checking domain of the complementizer, checking off its strong +wh-feature. Note also that since the lexical insertion of the complementizer takes place at the top of the tree, it is a legitimate instance of Merge. Under this analysis, (5b) then has the S-structure in (8a) and LF in (8b). The reason why no overt wh-movement takes place in (5b) is trivial. The CP projection, containing the +wh-feature, is not present in the overt syntax.

(8) a. SS [$_{IP}$ Tu as vu qui]

   b. LF [$_{CP}$ Qui C [$_{IP}$ tu as vu]]

Notice that in-situ questions in French are not restricted to arguments. Thus, (9) is also a well-formed non-echo question. The construction in (9) can be derived in the same way as (5b), with the LF insertion of the complementizer with a strong +wh-feature.[4]

(9) Jean a   réparé la voiture comment?
    Jean has fixed the car   how

Consider now embedded questions such as (6b). In order to account for the ungrammaticality of

4

(6b) we need to rule out the derivation in which the interrogative complementizer is inserted into the embedded clause in LF. The derivation is in fact straightforwardly ruled out because it involves merger that does not expand the tree, i.e., applies in an embedded position. To satisfy the requirement that merger expand the structure, the complementizer with strong +wh-feature must be introduced into the structure in (6b) in overt syntax before the higher structure is built. The wh-phrase then must move overtly to SpecCP to check the strong feature of the interrogative C as soon as the C is merged into the structure. Overt wh-movement is clearly the only option with embedded questions. The contrast between (5b) and (6b) is thus straightforwardly accounted for.[5]

This account of the contrast between (5b) and (6b) crucially depends on the interrogative C being phonologically null in the constructions under consideration. This is what enabled us to delay the insertion of the complementizer until LF in (5b), which eventually made it possible to account for the contrast between (5b) and (6b). That the fact that the complementizer is phonologically null in the constructions in question plays a crucial role in the phenomena under consideration and is not just a coincidence is confirmed by certain facts concerning wh-constructions in Serbo-Croatian (SC).

SC is a multiple wh-fronting language which, as shown in Rudin 1988, cannot place more than one fronted wh-phrase in SpecCP. In Bošković 1997b I show that SC exhibits an interesting pattern with respect to the Superiority Condition. In particular, SC exhibits Superiority effects exactly in those contexts in which French must have wh-movement. (This holds for all the contexts in which French has obligatory wh-movement discussed below.) Where French does not have to have syntactic wh-movement, SC does not exhibit Superiority effects. This is illustrated in (10-11), which show that fronted wh-phrases in SC are freely ordered in matrix questions but not in embedded questions.[6]

(10) a. Ko šta   kupuje?
         who what buys
         'Who buys what?'
     b. Šta ko kupuje?

(11) a. Zavisi   od toga [ko šta   kupuje]

        depends on it   who what buys

        'It depends on who buys what.'

  b. ?*Zavisi od toga [šta ko kupuje]

  c. [Ko   koga voli],   taj   o   njemu i   govori.

       who whom loves,   that-one about him   even talks

       'Everyone talks about the person they love.'

  d. ?*[Koga ko voli], taj o njemu/o njemu taj i govori.

  e. Ima [ko   šta da   kupi]

      has   who what particle buys

      'Someone has something to buy.'

  f. ?*Ima [šta ko da kupi]

To account for the parallelism between the contexts in which SC exhibits Superiority effects and the contexts in which French must have wh-movement I proposed in Bošković 1997b that SC is a French-type language with respect to when it must have wh-movement. Given that, as in French, no wh-movement must take place in SC matrix questions, the Superiority Condition is trivially satisfied in (10). Since, as in French, wh-movement must take place in SC embedded questions, the Superiority Condition is operative in (11). (11b,d,f) are then ruled out by the Superiority Condition. Given the analysis, the only difference between French and SC is that even wh-phrases that do not move overtly to SpecCP in SC still must undergo fronting. Thus, no wh-phrase is allowed to stay in-situ in (10-11). For example, *ko kupuje šta 'who buys what' is ungrammatical. This fronting of wh-phrases is clearly independent of wh-movement, as indicated by the fact that the construction is ungrammatical even on the echo question reading of šta. The same holds for *Jovan kupuje šta 'John buys what'. For the purposes of this paper we may assume that fronting of SC wh-phrases that does not land in SpecCP takes place in PF, as suggested in Bošković 1997b, and thus follows wh-movement. Stjepanović (1995) shows that the fronting takes place for focusing reasons. In what follows I will ignore this focus fronting of wh-

phrases and concentrate on wh-movement, i.e., movement to SpecCP, which, as suggested above, takes place in (11) but not in (10), SC being a French-type language with respect to when it must have wh-movement. As noted above, this enables us to account for the contrast between (10) and (11) with respect to Superiority.

Note that the interrogative complementizer in SC (10-11) is phonologically null. As a result, the analysis of the contrast between French (5b) and (6b) with respect to the obligatoriness of wh-movement can be readily extended to the SC constructions in question. SC, however, also has a phonologically realized interrogative complementizer *li*.[7] If the LF C-insertion analysis is on the right track we would expect that, in contrast to (10), matrix *li*-questions will exhibit Superiority effects, which is an indication of when wh-movement must take place in SC. Since *li* is not phonologically null it must be introduced into the structure overtly, in contrast to phonologically null complementizers. Overt movement to SpecCP then must take place in *li*-questions to eliminate the strong +wh-feature. (Since SC does not differ from French with respect to when it must have overt wh-movement, I assume that the +wh-feature in SC is strong.) The prediction is borne out. This is illustrated in (12). Example (12b) clearly contrasts with (10b) on the true question reading of the second wh-phrase.

(12) a. Ko li šta   kupuje?
          who C what bought
          'Who on earth buys what?'
     b. ?*Šta li ko kupuje?

The contrast between (10b) and (12b) indicates that in order for the wh-movement option not to be forced in matrix questions in SC, the interrogative complementizer must be phonologically null, as expected under the current analysis.

It is well-known that even in some dialects of French C may be phonologically realized in questions. More precisely, some French dialects allow the complementizer *que* to appear in questions, as illustrated in (13):

(13) Qui   que tu as   vu?

    whom   C   you have seen

    'Who did you see?'

Under the LF C-insertion analysis we predict that in-situ questions will not be possible with the overt complementizer *que*. Since, due to its phonological content, the complementizer must be introduced in the overt syntax, we would expect it to always trigger syntactic wh-movement. As (14) shows, the prediction is borne out. (14) is unacceptable for all speakers, including those who accept (13).

(14) * Que tu as vu qui?

The contrast between (13) and (14), as well as the fact that matrix overt C questions differ from matrix null C questions with respect to the availability of the in-situ strategy, provide a strong confirmation of the LF C-insertion analysis.

Two questions arise at this point: what is the status of (5b) in English and why are both the in-situ and the wh-movement option allowed with matrix questions in French? Consider first the counterpart of (5b) in English.

(15) (*) You saw who

The status of the construction is not quite clear. (15) does not seem to be as good as French (5b) on the true question (i.e. non-echo) interpretation. However, there still seems to be a contrast between (15) and embedded questions such as (16).

(16) *John wonders you saw who

If (15) is good on the true question reading, its grammaticality, as well as the contrast between (15) and (16), can be accounted for in the same way as (5b) and the contrast between (5b) and

(6b). If, on the other hand, (16) is bad on the true question reading we can account for it by assuming that the interrogative complementizer in English constructions such as (15) is lexically specified as a phonological affix, a rather natural move given *do*-support and S-Aux inversion in English matrix questions, which must be affixed to a verbal head at PF.[8] *Do*-support and S-Aux inversion can be assumed to take place in order to provide an appropriate host for the affixal complementizer. The derivation on which the interrogative complementizer is introduced in the LF of (15) is then ruled out straightforwardly because the requirement that the complementizer be affixed to a verbal head in PF cannot be satisfied in the derivation in question. (The derivation crashes because LF cannot interpret the phonological requirement on the interrogative C.)

Let us now turn to embedded questions. Notice first that, for some reason that remains unclear, the interrogative C in embedded questions in English does not seem to be a PF affix, given that neither *do*-support nor S-Aux inversion take place in English embedded questions. (Compare *What did John buy* and *What can John buy* with *I wonder what John bought* and *I wonder what John can buy*.)[9] The derivation in which the interrogative complementizer in (16) is introduced into the structure in LF, which would delay movement of *what* to the embedded CP projection until LF, then cannot be ruled out by appealing to PF requirements on the complementizer. However, as discussed above with respect to French (6b), the derivation in question is still ruled out because it involves merger, i.e. lexical insertion, that does not expand the tree. Since Merge can take place only at the root of a tree, C with strong +wh-feature (I assume that the +wh-feature is strong in English) must be introduced in the overt syntax before the higher structure is built. Wh-movement then must take place overtly in the construction in question to check the strong +wh-feature of C upon insertion. We thus derive *John wonders what Peter bought* instead of (16). Notice that we can still account for the fact that (16) is worse than (15) on this analysis because the constructions are ruled out in different ways: Merger in an embedded position and the failure to check off a strong feature immediately upon insertion should then be taken to result in stronger unacceptability than the failure of a PF affix to be affixed to an appropriate host at PF due to its LF insertion.

Consider now the question of why French allows (5a) in addition to (5b) in spite of

Procrastinate, which requires that all operations take place as late as possible. We can account for the existence of both (5a) and (5b) in French if French has both a PF affix and a non-affix interrogative C, the former being used in (5a) and the latter being used in (5b). The fact that, in contrast to (5a), inversion does not and cannot take place in (5b) (see (17)), may be relevant here. (Note that the ungrammaticality of (17) is fully consistent with the IP analysis of French in-situ questions.) However, the grammaticality of (18), where nothing seems to 'support' the interrogative C, casts doubt on this analysis.

(17) *As-tu vu qui?
(18) Qui   tu   as   vu?
       whom you have seen
       'Who did you see?'

The most promising way of accounting for the grammaticality of both the in-situ and the wh-movement strategy in French matrix questions seems to be to exempt Merge from Procrastinate. The phonologically null C could then be merged either overtly or covertly in French, giving us either (5a) or (5b). Having Merge freely apply either overtly or covertly in spite of Procrastinate is actually a rather natural move in Chomsky's (1995) system. Procrastinate was originally posited in order to delay movement until LF whenever this is possible (see Chomsky and Lasnik 1993: 535). In Chomsky's (1995) system, however, this can be achieved without Procrastinate. Chomsky observes that a natural consequence of the standard minimalist assumption that movement is driven by feature-checking is that, all else being equal, the operation Move should apply to features and not to syntactic categories. Overt movement, which feeds PF, still has to apply to whole categories, given the natural assumption that lexical items with scattered features cannot be interpreted/pronounced at PF. Since the considerations of PF interpretability are not relevant to LF, in LF the operation Move applies only to features. As observed in Chomsky (1995) and attributed to Hisa Kitahara and Howard Lasnik, in this system LF movement is in a sense always cheaper than overt movement since it carries less material: whereas overt movement affects the whole categories, covert movement affects only features. Since covert

movement is more economical than overt movement, it should be preferred to overt movement. Note now that, given feature-movement, there is no longer any need for an independent principle such as Procrastinate to ensure that movement is delayed until LF whenever possible. Since this was the main motivation for positing Procrastinate in the first place, it seems that Procrastinate is now eliminable. And if Procrastinate is eliminated, Merge should be free to apply either overtly or covertly. Introduction of feature movement into the theory does not affect Merge, i.e., it does not appear to make LF Merge in any sense cheaper than overt Merge. There is only one scenario under which it actually does: Suppose that a merged element triggers movement. The movement itself could be cheaper if the element in question is merged in LF rather than in the overt syntax, since in the former case we would be dealing with feature movement. Notice, however, that we need global economy of derivation to make LF Merge cheaper in the case in question. This cannot be achieved with local economy of derivation, which is clearly conceptually more appealing than global economy of derivation (see Chomsky 1995 and Collins 1995 for much relevant discussion), because local economy of derivation does not allow look-ahead. If the insertion itself is not cheaper in point A than in point B, by local economy of derivation it should be allowed to apply either in point A or point B. Merger of the interrogative C in the constructions under consideration then should be able to apply either covertly or overtly. As a result, we account for the fact that both (5a) and (5b) are good. In (5a), the interrogative C is merged into the structure overtly, which triggers overt wh-movement, and in (5b) it is merged covertly. As discussed above, covert applications of Merge are still restricted, since only phonologically null elements can be introduced into the tree in LF. And both covert and overt Merge are constrained by the requirement that Merge must expand the structure, i.e., apply only at the top of the tree.[10]

To summarize, we have seen that for the no overt wh-movement option to be available in French the interrogative complementizer must be phonologically null and located at the top of the tree. The LF insertion of the complementizer analysis provides a natural way of capturing these two conditions: As discussed above, only phonologically null elements can be inserted covertly and lexical insertion can take place only at the top of the tree. The behavior of French in wh-in-situ questions thus follows straightforwardly given the way lexical insertion takes place in

the minimalist system and given Chomsky's (1995) conception of strong features. The fact that under the current analysis, the strength of the +wh-feature in French can be kept constant, i.e., there is no need to assume that the +wh-feature of French C is sometimes weak and sometimes strong to account for the availability of more than one option with respect to wh-movement in French, should be considered a plus for the analysis. Assuming that the +wh-feature can be either strong or weak in French would amount to sneaking in optional movement, which is by its very nature incompatible with the minimalist system. The optionally strong/weak +wh-feature analysis is also empirically inferior to the current analysis since it fails to account for the fact that the in-situ strategy is not always allowed in French. Chomsky (1995) assumes that the +wh-feature is interpretable and therefore does not have to be checked when it is weak (see the discussion in section 4). It is not clear, then, how the ungrammaticality of (6b) can be accounted for if the +wh-feature can be weak in French. The same holds for other contexts in which wh-movement is obligatory in French discussed below (see examples (19a), (20a), and (21a)). As shown below, the data in question receive a straightforward account under the strong +wh-feature+LF C-insertion analysis. To the extent that it is successful, the analysis provides evidence that lexical insertion, in fact even the insertion of elements with strong features, can take place in LF in certain well-defined configurations.[11] This in turn provides evidence for Chomsky's (1995) approach to strong features, which leaves room for covert lexical insertion of elements with strong features. We have also seen that if we adopt the local rather than the global view of economy of derivation we can account for the fact both the in-situ and the wh-movement strategy are available in French root questions, which provides evidence that the local view of economy of derivation is superior to the global view on empirical grounds. Turning to English, we have seen that, if real, the different behavior of French and English with respect to the availability of the in-situ strategy in matrix questions can be accounted for given the plausible assumption that the interrogative C is a phonological affix in English.

It should be pointed out here that Chomsky (1995) explicitly bans covert insertion of elements with strong features, which plays a crucial role in the current analysis. However, Chomsky's empirical reasons for doing this are very weak, and the way of doing it is conceptually very problematic. Chomsky appears to have two empirical reasons for banning LF

insertion of elements with strong features: accounting for the ungrammaticality of English (15) and what he calls the general unacceptability of in-situ questions with adjuncts even in French-type languages. However, we have seen above that the ungrammaticality of (15), as well as the contrast between (15) and (5b), can be accounted for even if we allow covert insertion of elements with strong features. As for in-situ adjunct questions, they are clearly allowed in French, as illustrated in (9). Note also that Chomsky does not discuss the fact that the in-situ strategy is not always available in French even with arguments, as illustrated by the contrast between (5b) and (6b), and the contrast between (13) and (14) (see also section 3.). It is difficult to see how the contrasts in question can be accounted for in a principled way without invoking LF insertion of phonologically null elements with strong features, which is certainly a possibility in the theory (i.e., all the mechanisms needed in the LF C-insertion analysis are expected to be available given the basic minimalist assumptions in Chomsky 1995.)[12]

3. Long-distance and negated questions in French

In the previous section I have discussed the well-known difference between matrix and embedded null C questions in French with respect to the availability of wh-in-situ. The wh-in-situ strategy, however, is not always available in French matrix null C questions. Thus, my informants find long-distance questions in which a finite clause boundary intervenes between a +wh SpecCP and a wh-in-situ fully acceptable only on the echo-question reading. Wh-movement is obligatory in the contexts in question.[13]

(19) a. ?*Jean et   Pierre croient que Marie a   vu   qui?

         Jean and Pierre believe that Marie has seen whom

       'Whom do Jean and Pierre believe that Marie saw?'

    b. Qui Jean et Pierre croient-ils que Marie a vu?

(20) a. *Jean et   Marie se demandent si Pierre aime qui?

       Jean and Marie wonder   if Pierre loves whom

       'Whom do Jean and Marie wonder if Peter loves?'

b. ??Qui Jean et Marie se demandent-ils si Pierre aime?

A question arises now why, in contrast to short-distance questions, the in-situ strategy is not available in French long-distance questions.[14] The question again receives a principled answer in the minimalist system. Furthermore, the answer turns out to have interesting theoretical consequences.

Consider (19a). (Example (20a) can be accounted for in the same way.) What we need to do is rule out the derivation in which the complementizer with a strong +wh-feature is inserted at the top of the tree in the LF of (19a), with the wh-phrase undergoing movement to check off the strong +wh-feature of the complementizer immediately upon the insertion. If this derivation were to yield a legitimate output, we would not be able to account for the ungrammaticality of (19a) on the true non-echo question reading of the wh-phrase. If the movement of *qui,* motivated by checking the strong +wh-feature of the complementizer, were overt it would involve substitution into SpecCP. Note, however, that since the complementizer is inserted in the covert syntax in the derivation under consideration, the movement in question will also have to be covert. Recall now that all LF movement necessarily involves pure feature movement. Chomsky (1995) argues that in LF formal features move or, to be more precise, adjoin to the heads bearing matching features. LF movement thus necessarily involves head movement, i.e., adjunction to X^o-elements. Instead of moving to SpecCP, *qui* will then have to undergo head movement by adjoining to the matrix C in the LF of (19a) in the derivation under consideration. It is well-known, however, that movement to X^o positions is subject to very strict locality restrictions. As far as I know, no grammatical instance of head movement of the 'necessary' length has ever been reported. I, therefore, suggest that the derivation in question is ruled out because it involves movement to an X^o position that violates locality restrictions on head-movement.

Notice that the movement of *qui* to the matrix C in (19a) violates locality restrictions on head-movement even if we adopt Roberts' (1992) (see also Rivero 1991) relaxed version of the Head Movement Constraint, which subjects head-movement to relativized minimality and by doing this allows movement to an X^o-position of type α to cross X^o-positions of type β but not type α. To reach the matrix C, an A'-head position, in the derivation in question, *qui* must skip

another A'-head, namely, the embedded clause C, which violates even Roberts' relaxed relativized minimality version of the Head Movement Constraint. On the other hand, in short-distance questions such as (5b), LF movement of the wh-phrase to the interrogative complementizer does not cross any A'-heads, so that relativized minimality is respected. In other words, it represents a grammatical instance of movement to X^o-positions under Roberts' approach to locality restrictions on such movement. This approach thus enables us to account for the contrast between (5b) and (19a).

It also enables us to account for the degraded status of (21a) on the true non-echo question reading. My informants find negative wh-in-situ constructions such as (21a) degraded on the true question reading of the wh-phrase.

(21) a. ?*Jean ne   mange pas quoi?

        Jean neg eats   neg   what

        'What doesn't John eat?'

   b. cf. Que ne mange-t-il pas?

In order to derive a legitimate output for (21a) on the true question, non-echo reading of the wh-phrase in-situ, the complementizer must be inserted in the LF of (21a) and the wh-phrase must move to the complementizer to check off its strong +wh-feature immediately upon insertion of the complementizer. The movement of the wh-phrase, however, violates Roberts' version of the Head Movement Constraint. Like the movement of the wh-phrase to the matrix C in (19a), movement of *quoi* to C, an A'-head, crosses another A'-head, this time negation, and is expected to yield an ungrammatical output under this approach to movement to X^o-positions. Adopting this approach to locality conditions on movement to X^o-positions thus enables us to account not only for the ungrammaticality of (19a), but also the ungrammaticality of (21a), as well as the contrast between (5b) and (19a, 21a). Since this analysis is crucially based on Chomsky's feature-movement hypothesis, which forces all LF movement to be adjunction to X^0-positions, to the extent that it is successful, the analysis provides evidence for the feature movement hypothesis.[15] The analysis also provides evidence that feature-movement is subject to locality restrictions on

movement.[16] In fact, feature movement seems to be subject to essentially the same locality restrictions as overt movement to X^0-positions, which confirms Chomsky's intuition that feature-movement in some sense involves head-movement. It remains to be seen, however, how the relativized minimality version of the Head Movement Constraint can be integrated into Chomsky's (1995) system and what its consequences for the system will be. The needed version of the Head Movement Constraint does not readily fit into a system based on Attract instead of Move. However, the same holds for some other locality constraints on movement, for example the traditional Condition on the Extraction Domain (CED) phenomena of Huang (1982), in fact almost everything except the Wh-Island Constraint. There thus still seems to be a need for a conception of the Make the Shortest Move Principle which considers movement from the point of view of the moved element. In contrast to Attract, such a conception of the Make the Shortest Move Principle can readily accommodate the relativized minimality version of the Head Movement Constraint. As shown in Takahashi 1994, it can also accommodate the traditional CED phenomena and the Coordinate Structure Constraint, which remain mysterious under Attract.[17]

4. Wh-phrases that do not move in LF

Consider now multiple questions in which a clausal boundary or negation intervenes between a wh-in-situ and a +wh C whose Spec is overtly filled by another wh-phrase. (22a) and (22b) are better than (19a) and (21a) on the true question, non-echo reading of the wh-phrases in-situ.

(22) a. Qui croit   que Marie a   vu   qui?
         who believes that Marie has seen whom
     b. Qui ne mange pas quoi?
         who neg eats   neg what

The grammaticality of the constructions on the true question reading of the wh-phrases in-situ indicates that the wh-phrases in-situ are not undergoing LF movement to C. If they were, we

would expect (22a,b) to have the same status as (19a) and (21a), which is not the case. This is expected in the minimalist system. In this system, movement to a +wh SpecCP, or adjunction to a +wh C, take place to check off the strong +wh-feature of C. In (22a-b) this is done by the wh-phrases that are located in SpecCP at SS. There is therefore no need for the wh-phrases that are located in-situ in the SS of (22a-b) to undergo LF movement to C, in contrast to the wh-phrases in (19a) and (21a). (Note that in Chomsky's 1995 system, the +wh-feature of the wh-phrases is considered to be interpretable and weak and, therefore, does not have to be checked.) The Last Resort Condition, which bans superfluous steps, then prevents them from undergoing LF movement to C.[18]

It seems plausible, however, that a wh-phrase in-situ must be somehow associated with a +wh C, possibly as a way of scope taking, in order to ensure proper interpretation (see Baker 1970, Chomsky 1973, Reinhart 1995, Williams 1986, among many others). In (22a-b), this can be done through the mechanism of unselective binding, with the interrogative C binding the wh-phrases in-situ. Notice that this cannot be done in (19a) and (21a), since the strong +wh-feature of the complementizer remains unchecked if the wh-phrases in-situ are associated with the interrogative complementizer through unselective binding. To check the strong +wh-feature, the wh-phrases must enter into a checking relation with the complementizer, which is done through LF adjunction to the complementizer, after the complementizer is inserted into the structure.

A question that arises now is why (23a-b), with an in-situ adjunct, are ungrammatical.

(23) a. *Qui a réparé la voiture comment?

    b. *Who fixed the car how?

Under the standard pre-minimalist analysis, (23a-b) are ruled out by the Empty Category Principle (ECP). Since all wh-phrases are assumed to be located in SpecCP in LF, the wh-phrases in-situ in (23a-b) must undergo LF movement to SpecCP. The constructions are then assumed to be ruled out because the trace left by the LF wh-movement of the wh-adjuncts is not properly governed. This account is clearly not available in the minimalist framework. Furthermore, it is empirically inadequate.

Rudin (1988) shows convincingly that in Bulgarian constructions such as (24), both wh-phrases are located in SpecCP with the adjunct wh-phrase being adjoined to SpecCP, since the subject wh-phrase moves to SpecCP before the adjunct to satisfy the Superiority Condition. (The condition requires that given two wh-phrases in-situ, the higher wh-phrase move to SpecCP. In the Bulgarian case, the highest wh-phrase moves first.)[19]

(24) [$_{CP}$ [[Koj] kak] [$_{C'}$ e   kupil   kniga]]
       who how   is bought book
       'Who bought the book how?'

The SS of (24) is thus the same as the LF of (23) in the pre-minimalist framework. The grammaticality of (24), then, provides evidence that an element adjoined to SpecCP can license its trace, contrary to the standard assumption that antecedent government from the SpecCP adjoined position is not possible.[20]

Furthermore, Haider (1986) and Müller and Sternefeld (1996) show that in German, a wh-adverb can remain in-situ overtly even in constructions in which the Spec of the interrogative C in which the adverb is interpreted is filled by another wh-phrase. ((25a) is from Haider 1986, and (25b) from Müller and Sternefeld 1996.)

(25) a. Wer ist weshalb weggegangen?
         who is why   away-walked
      'Who left why?'
      b. Wer is warum gekommen
         who is why   come
         'Who came why?'

As Haider points out, the standard ECP account of (23) incorrectly rules out (25).

Finally, notice that if the adjuncts in (23a-b) were to undergo LF wh-movement, under the feature movement theory the movement would be feature-adjunction to the interrogative C and

not movement to SpecCP. It seems that this would be a perfectly legitimate instance of feature movement. In fact, the movement would not differ in any respect from the movement the adjunct must undergo in the grammatical (9) to check the strong +wh-feature of the complementizer, which is inserted in LF. Clearly, we need a new account of the ungrammaticality of (23a-b).

Recall that I have assumed above that a wh-phrase must be associated with an interrogative complementizer in order to ensure proper interpretation. The association can be established in two ways: either by moving the wh-phrase to a position within the projection of the complementizer, or by having the complementizer unselectively bind the wh-phrase in-situ. Now, Tsai (1994) and Reinhart (1995) argue that wh-adjuncts or, to be more precise, wh-adverbs, cannot be unselectively bound due to the lack of a variable (i.e. an open position).[21] As a result, Tsai and Reinhart argue, wh-adverbs cannot be interpreted in-situ. If this is true, the only way for the wh-adverbs in (23) to be associated with the interrogative C is for them to undergo LF movement to the interrogative C.[22] As noted above, this appears to be a perfectly well-formed instance of feature movement which does not violate any locality conditions on movement or licensing of traces. Notice, however, that the movement has no morphological motivation. Neither the complementizer nor the wh-adverb contain any features that need to be checked prior to the movement. The structures that (23a-b) have prior to the movement of the adjuncts to the interrogative C are fully legitimate LFs, with all features that need to be checked checked. As a result, the movement of the adverbs to the interrogative complementizer violates the Last Resort Condition. The problem with the movement in question is that it is driven by what Chomsky (1993) calls 'the search for intelligibility' rather than the satisfaction of formal requirements (feature-checking), which is the only legitimate driving force for movement according to Chomsky (1993, 1995). I conclude, therefore, that (23a-b) converge as gibberish without LF movement of the wh-adverbs, which would ensure proper interpretation but is banned by the Last Resort Condition.

Let us now see how we can account for German (25a-b), which contrast with French and English (23a-b). The grammaticality of the German constructions indicates that the wh-adverbs are succeeding in establishing a relation with the interrogative C. We could account for this by assuming that German wh-adverbs differ from their French and English counterparts in that they

can be unselectively bound by C. This seems unlikely, given that, as argued in Tsai (1994) and Reinhart (1995), the possibility of unselective binding depends on the availability of an open position, i.e., a variable, which exists in NPs but not in adverbs (in this respect, see also Higginbotham 1983, 1985). It then seems more appropriate to account for the grammaticality of (25a-b) by having the wh-adverbs undergo LF feature movement to the matrix C. The driving force for the movement has to be independent of the +wh-feature, i.e., the feature driving it cannot be the +wh-feature and should not be available in English and French. It is tempting to relate this movement to the availability of partial wh-movement in German, which indicates that German wh-phrases can move to within a CP projection independently of the +wh-feature, the relevant C not being specified as +wh. That relating the relevant feature to the one involved in partial wh-movement may be on the right track is indicated by the fact that partial wh-movement is not available in English and French.[23] It is worth noting in this context that Bošković (1997b) and Stjepanović (1995) show that Serbo-Croatian wh-phrases can also be fronted independently of the +wh-feature. (They also show that the movement is not an instance of scrambling). In fact, Bošković (in press a) argues with respect to Bulgarian, a language that allows multiply filled interrogative SpecCPs in overt syntax, that even movement of wh-phrases to an interrogative SpecCP does not have to be driven by +wh-feature checking, i.e., it can have feature-checking motivation independent of the +wh-feature. Interestingly, Müller and Sternefeld (1996) observe that, in contrast to short-distance questions, German wh-adverbs cannot be left in situ in long-distance questions. This is exactly what is expected if they are undergoing LF feature movement to C. The ungrammaticality of (26) then can be accounted for in the same way as the ungrammaticality of (19a).

(26) a. *Wer hat gesagt daß Fritz warum ein Buch gelesen hat
            who has said   that Fritz why   a   book read   has
            'Who has said that Fritz has read a book why?'
     b. *Wen   hast du empfohlen   daß man wie bestrafen soll
            whom have you recommended that one how punish   should

Whatever the precise identity of the feature that is responsible for the LF movement of wh-adverbs in German is, it cannot be obligatorily present in German wh-phrases or the interrogative C.[24] This is indicated by the fact, observed by Müller and Sternefeld (1996), that argument wh-phrases can be left in-situ in long-distance questions in German, which in turn indicates that not all German wh-phrases in-situ are forced to move to C in LF.

(27) Wer hat gesagt daß Fritz was   lesen soll
      who has said   that Fritz what read should

In contrast to wh-adverbs, the wh-phrase in (27) can be unselectively bound and therefore does not depend on undergoing movement to the interrogative C to be associated with it.

A question that arises now is what the proper treatment of pure wh-in-situ languages such as Japanese is? Notice first that Japanese (28a) should not receive the same analysis as French (5b), i.e., it should not be analyzed as involving LF insertion of a complementizer with a strong +wh-feature since, in contrast to French, embedded questions with wh-phrases in-situ are allowed in Japanese. If Japanese were to be given the same analysis as French, we would incorrectly predict (28b) to be ungrammatical. Notice also that the interrogative complementizer in Japanese (28) is lexically realized. As a result, in contrast to the interrogative complementizer in French (5b), the interrogative complementizer in Japanese (28) clearly cannot be introduced into the structure in LF.

(28) a. Anata-ga dare-o   mita ka
        you-nom who-acc saw   Q
     b. Peter-wa [anata-ga dare-o   mita ka] tazuneta
         Peter-top you-nom who-acc saw   Q asked

It is well-known that long-distance in-situ questions are also good in Japanese. Thus, both (29a) and (29b) are good on the matrix true question non-echo reading of the wh-phrase and contrast with French (19a) and (20a). Negated questions are also good on the non-echo reading, again in

contrast with French (cf. (21a)).

(29) a. John to   Mary-wa [Peter-ga   dare-o   mita to] sinziteiru ka
        John and Mary-top Peter-nom who-acc saw that believe   Q
    b. (?)John to   Mary-wa [Peter-ga   dare-o   mita kadooka] siritagatteiru   ka
          John and Mary-top Peter-nom who-acc saw   whether   want-to-know Q
    c. John-wa nani-o   tabenakatta ka
       John-top what-acc ate-neg   Q

The data in (28-29) indicate that, in contrast to French, Japanese wh-phrases do not undergo LF feature movement. If they did, it would be difficult to account for the contrasts between Japanese and French noted above. However, it is well-known that Japanese wh-phrases are not completely insensitive to constraints on movement. Thus, adjunct wh-phrases cannot be separated from the interrogative complementizer by an island, as illustrated in (30). According to Watanabe (1992), some degradation is found even in argument questions such as (29b), though this claim is somewhat controversial.

(30) *John to   Mary-wa [Peter-ga   naze kuruma-o naosita kadooka] siritagatteiru   ka
       John and Mary-top Peter-nom why car-acc   fixed   whether   want-to-know Q

Given (30), some kind of movement must be taking place in Japanese questions. As noted above, the movement cannot be covert, or we would not able to account for the contrast between (29) and (19a), (20a), and (21a). The movement, then, must be overt. In other words, the movement in question seems to behave like XP movement, rather than X^0-movement. Wh-phrases themselves are clearly not undergoing wh-movement in the constructions under consideration, since they are located in-situ. Therefore, the only possibility seems to be that a null wh-operator is undergoing movement. This is exactly what is proposed in Watanabe (1992) and Aoun and Li (1993), who argue that the difference between languages such as English and languages such as Japanese with respect to what moves in wh-questions is a result of the shape of

wh-phrases and their quantificational force in the languages in questions. Under both Watanabe's and Aoun and Li's analyses, interrogative SpecCPs are always overtly filled in languages such as Japanese. Under Watanabe's analysis, the wh-operator always moves to SpecCP, whereas under Aoun and Li's analysis it is sometimes base-generated there and sometimes it moves, which is intended to account for island effects. Both analyses imply that the +wh-feature in Japanese is strong and, therefore, requires presence of an element with a +wh-feature in SpecCP overtly. In the current system, we seem to be led to a Watanabe/Aoun and Li style analysis. There are still many wrinkles left to iron out, doing which would go well beyond the scope of this paper. For some relevant discussion, see Watanabe (1992) and Aoun and Li (1993).[25]

5. Conclusion and some theoretical consequences of the proposed analysis

In this paper I have examined when wh-movement must take place in French and offered a minimalist account of the limited distribution of the in-situ strategy in French. I have argued that in French constructions involving wh-in-situ such as (5b), the interrogative CP projection is inserted only in LF. This is a necessary condition for the in-situ strategy to be available in French. Wh-movement then does not take place overtly in (5b) for a trivial reason: no CP projection is present in the SS of (5b). Where the in-situ option is not allowed, either LF insertion of the interrogative CP projection is blocked or the insertion leads to a violation of locality constraints on movement. To be more precise, with embedded questions, LF insertion of the interrogative complementizer is blocked because it violates the requirement that merger expand the structure, and with long-distance and negative questions the insertion leads to a violation of locality constraints on movement. The LF C-insertion analysis has interesting theoretical consequences. It provides evidence that lexical insertion, in fact even lexical insertion of elements with strong features, can take place in LF under certain well-defined conditions, namely, when the insertion takes place at the root of the tree and when the element undergoing insertion is phonologically null. This in turn provides evidence that Chomsky's (1995) conception of strong features is superior to earlier conceptions, for example Chomsky (1994),

since it allows LF insertion of elements with strong features.

The current analysis also provides evidence that feature movement is subject to at least some locality constraints on movement. In fact, it seems to be subject to the same constraints as overt movement to X^0-positions, which confirms Chomsky's claim that feature movement involves X^0-adjunction. Note that the analysis presented here is crucially based on the assumption that all LF movement involves adjunction to X^0-elements. (LF wh-movement then involves adjunction to C^0, rather than movement to SpecCP.) As a result, to the extent that it is successful, the analysis presented here also provides evidence for Chomsky's feature movement hypothesis.

Under the current analysis we would expect to regularly find wh-elements in-situ in matrix clauses in languages with a strong +wh-feature as long as the interrogative complementizer is phonologically null and is not lexically specified as a phonological affix, a reflex of which could be S-Aux inversion or a *do*-support type process. (A phonological affix C must be PF-adjacent to its host.) However, S-Aux inversion is cross-linguistically a very common phenomenon in questions. This is not surprising, given Pesetsky's (1992) conjecture that null morphemes are by and large specified as phonological affixes. Another interfering factor is that seeing a fronted wh-phrase does not necessarily indicate that wh-movement is taking place. Thus, as discussed in Bošković 1997b and illustrated briefly above, Serbo-Croatian wh-phrases, including wh-phrases with echo interpretation, are forced to undergo fronting quite independently of the wh-feature. (Stjepanović (1995) argues that the fronting is motivated by a focusing requirement on Serbo-Croatian wh-phrases.) Fronting of a wh-phrase in Serbo-Croatian thus does not necessarily provide evidence that wh-movement is taking place. In fact, as discussed above, in spite of obligatory fronting of wh-phrases, Serbo-Croatian turns out to behave like French with respect to when it must have wh-movement. Apparently, care should be taken in diagnosing wh-movement. Fronting of wh-phrases cannot be taken as a sole diagnostic of wh-movement. The same point is made in Cheng 1991 with respect to several languages.

There is one potential empirical consequence of the analysis developed above that still remains to be discussed. We have seen that nothing in the minimalist system prevents a phonologically null complementizer with a strong +wh-feature from being inserted in LF at the

top of the tree in short-distance root questions. A question that arises now is whether other phonologically null elements can also be inserted in LF if their insertion takes place at the top of the tree. This could leave us with the possibility of matrix clauses being bare VPs in certain circumstances in the overt syntax, with the rest of the structure being built in the covert syntax. Consider first the question with respect to English constructions such as (31):

(31) John wanted a house

A question that arises is whether INFL, and with it the IP projection, can be inserted in the LF of (31), as a result of which the construction would be a bare VP in the overt syntax.[26] It appears that this would be possible if the INFL of (31) is phonologically null and has no PF requirements, for example, it is not specified as a PF affix. Halle and Marantz (1993), Bobaljik (1994), and Lasnik (1995) have, however, recently revived Chomsky's (1957) analysis of finite clauses by arguing that English finite INFL is a PF affix. In fact, according to these authors, INFL in (31) is not simply specified as a phonological affix, it even has phonological content. They argue that *ed* is actually located under INFL and undergoes affix hopping to the main verb, or in more recent terms, undergoes morphological merger (which should not be confused with Merge) with the main verb in PF under adjacency. If this is indeed the case, INFL, and with it the IP projection, would have to be present in the overt syntax in (31).

   Consider, however, the French construction in (32):

(32) Jean voulait une maison.
  Jean wanted a house

Lasnik (1995) argues that, in contrast to English finite INFL, French finite INFL is neither phonologically realized nor specified as a phonological affix. It is simply a bunch of features checked against the features of the verb after the verb raises to it. (In other words, while in English *want* but not *wanted* is present in the lexicon, *ed* being a phonological realization of INFL, in French *voulait* as whole is present in the lexical entry of the verb in question.) If INFL

in French (32) is indeed phonologically null, a question arises as to whether it could be inserted into the structure in LF, which in turn opens up the possibility that the construction could be a bare VP at SS.[27] It is very difficult to tell empirically whether this is a possibility since, as discussed above with respect to the phonologically null complementizer, the derivation in which INFL is inserted in the overt syntax is also available in (32). Allowing LF insertion of INFL could not then rule out any constructions, though it might rule in some constructions that were ungrammatical on the IP-in-syntax derivation. (Recall that this was the case with allowing the possibility of the interrogative C being inserted in LF, which did rule in some constructions that were ungrammatical on the CP-in-syntax derivation.) At this point I have no empirical evidence that could conclusively settle the issue under consideration. What makes our job particularly difficult here is that it is not enough that INFL is phonologically null and has no PF requirements of its own to ensure that INFL can be inserted in LF. We need to ensure that no element present in the structure is required to move to INFL, or an element within the IP projection, overtly. This probably makes the ungrammaticality of negative sentences such as (33) irrelevant here, since it seems plausible that French negation is required to move to INFL overtly (see Acquaviva 1995, Belletti 1990, Haegeman 1992, Moritz 1989, Pollock 1989, among others.)

(33) a. *(ne) pas Jean aime Marie.

        neg neg Jean loves Marie

   b. *(ne) aime pas Jean Marie

If the negative marker *ne* indeed must end up in INFL overtly, the movement should probably be driven by some inadequacy of negation, since non-negative sentences where negation is not present in INFL are fine. (I assume that in negative constructions where *ne* is missing, *ne* is deleted in PF.) This means that the presence of negation requires overt insertion of INFL, which in turn triggers subject and V-movement to SpecIP and INFL respectively.

     There is a similar interfering factor in (34), another potentially relevant construction.

(34) *Complètement, Jean oubliera   cette histoire.

     Completely   Jean will-forget this story

(34) contains what is traditionally referred to as a VP adverb. (See, however, Belletti 1990, who suggests that *complètement* can be either a VP or a TP adverb). If *complètement* is indeed a VP adverb, a question arises as to why (34) is bad on the derivation in which INFL is inserted covertly. Kayne (1994) argues that the distinction between specifiers and adjuncts should be abolished. As noted by Ian Roberts (personal communication), in Kayne's 1994 system the adverb in (34) has to occupy the Spec position of some functional head, which is phonologically null. Suppose now that the phonologically null head, which is clearly phonologically weak, is a clitic and that French clitics in general must move to INFL, as suggested in Belletti 1990 and Kayne 1991, among many others, neither of which is an unreasonable assumption. Given these assumptions, (34) would be ruled out in the same way as (33). The presence of the functional head that the adverb introduces would require overt insertion of INFL, which in turn would require overt subject and V-movement to SpecIP and V respectively.   (Note that, as shown in Belletti 1990, *complètement* cannot undergo topicalization, which rules out the option of analyzing (34) with the subject in SpecIP, V in INFL, and *complètement* undergoing topicalization.)[28] The ungrammaticality of (34) thus cannot help us draw a definite conclusion concerning the possibility of covert INFL insertion. Covert INFL insertion may actually be quite generally banned for independent reasons. Consider again the derivation in which INFL is inserted overtly in (32), with the verb moving to adjoin to INFL and the subject moving to SpecIP. If these movements are driven by strong features of INFL, it appears that INFL could still be inserted covertly in (32). However, if these movements are driven by some inadequacy of the moved elements, which would require them to be located in the INFL-adjoined position and SpecIP respectively, then the option of inserting INFL and the IP projection in LF would be independently ruled out. The option would also be ruled out if the EPP effect is a result of some general requirement that there must be a filled SpecIP in every sentence. It is difficult to find empirical evidence concerning what drives the movements in question. Chomsky (1995) assumes without real empirical evidence that the movements are driven by strong features of

INFL, in which case it seems that there are no obstacles to LF insertion of INFL in (32). However, Chomsky (MIT Lectures 1995) suggests that the EPP is a result of some universal thematization requirement, which cross-linguistically requires overtly filled SpecIPs. The requirement could be interpreted in a way that would completely rule out the option of inserting INFL in LF and with it the option of inserting covertly any projection lower than INFL, given that Merge must expand the structure.[29]

So, where does this leave us? We seem to have good empirical and conceptual evidence (the theory allows it) that under certain circumstances C and the CP projection can be inserted into the structure covertly, though the covert insertion option did not rule out the overt insertion option. As for INFL and the IP projection, the situation is unclear. At this moment I have no conclusive empirical evidence that would go one way or the other, and it is unclear whether the option would be allowed theoretically. In other words, firm conceptual evidence of the relevant sort (this is allowed by the theory) is still not here, mainly because it is unclear what is responsible for movements to positions within the IP projection. What is clear is that the covert lexical insertion option is theoretically very limited: it can apply only at the root of the tree to phonologically null elements that furthermore have no PF requirements and do not have to be present overtly to provide a landing site for another element that must move overtly. The fact that we have been able to find one instance of LF insertion that exactly fits the conditions stated above provides a confirmation of the theoretical framework adopted in this paper.

References

Acquaviva, Paolo. 1994. The representation of operator-variable dependencies in sentential negation. Studia Linguistica 48:91-132.

Aoun, Joseph. 1986. Generalized Binding. Dordrecht: Foris Publications.

Aoun, Joseph and Yen-hui Audrey Li. 1993. Wh-elements in situ: Syntax or LF? Linguistic Inquiry 24:199-238.

Baker, C. L. 1970. Note on the description of English questions: The role of an abstract question morpheme. Foundations of Language 6:197-219.

Belletti, Adriana. 1990. Generalized verb movement. Turin: Rosenberg and Sellier.

Bobaljik, Jonathan. 1994. What Does Adjacency Do? In MIT Working Papers in Linguistics 22, 1-32. Department of Linguistics and Philosophy, MIT, Cambridge, Mass.

Bošković, Željko. 1992. Clausal selection, subjacency, and minimality. Ms., University of Connecticut, Storrs.

Bošković, Željko. 1996. Selection and the categorial status of infinitival complements. Natural Language and Linguistic Theory 14:269-304.

Bošković, Željko. 1997a. On certain violations of the Superiority Condition, Agro, and economy of derivation. Journal of Linguistics 33:227-254.

Bošković, Željko. 1997b. Superiority effects with multiple wh-fronting in Serbo-Croatian. Lingua 102:1-20.

Bošković, Željko. 1997c. The syntax of nonfinite complementation: An economy approach. Cambridge, Mass.: MIT Press.

Bošković, Željko. in press a. Multiple wh-fronting and economy of derivation. Proceedings of the West Coast Conference on Formal Linguistics 16. Stanford Linguistics Association, Stanford University, Stanford, Calif.

Bošković, Željko. in press b. LF movement and the Minimalist Program. In Proceedings of the North Eastern Linguistic Society 28. GLSA, University of Massachusetts, Amherst.

Cheng, Lisa. 1991. On the typology of wh-questions. Doctoral dissertation, MIT, Cambridge, Mass.

Chomsky, Noam. 1957. Syntactic structures. The Hague: Mouton.

Chomsky, Noam. 1973. Conditions on transformations. In A festschrift for Morris Halle, ed. Stephen R. Anderson and Paul Kiparsky, 232-286. New York: Holt, Reinhart, and Winston.

Chomsky Noam. 1991. Some notes on economy of derivation and representation. In Principles and parameters in comparative grammar, ed. Robert Freidin, 417-454. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 1993. A Minimalist Program for Linguistic Theory. In The view from building 20: Essays in linguistics in honor of Sylvain Bromberger, ed. Ken Hale and S.

Jay Keyser, 1-52. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 1994. Bare phrase structure. In <u>MIT Occasional Papers in Linguistics</u> 5. Department of Linguistics and Philosophy, MIT, Cambridge, Mass.

Chomsky, Noam. 1995.Categories and transformations. In <u>The minimalist program</u>, 219-394. Cambridge, Mass.: MIT Press.

Chomsky, Noam, and Howard Lasnik. 1977. Filters and control. <u>Linguistic Inquiry</u> 8:425-504.

Chomsky, Noam, and Howard Lasnik. 1993. The theory of principles and parameters. In <u>Theory of syntax: An international handbook of contemporary research</u>, ed. Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld, and Theo Vennemann, 506-569. Berlin: De Gruyter.

Collins, Chris. 1995. Toward a theory of optimal derivation. <u>MIT Working Papers in Linguistics</u> 27, 65-103. Department of Linguistics and Philosophy, MIT, Cambridge, Mass.

Haegeman, Liliane. 1992. Sentential negation in Italian and the Neg Criterion. <u>Geneva Generative Papers</u> 0, 1-38. Université de Genève.

Haider, Hubert. 1986. Affect α: A reply to Lasnik and Saito, On the nature of proper government. <u>Linguistic Inquiry</u> 17:113-126.

Halle, Morris, and Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In <u>The view from building 20: Essays in linguistics in honor of Sylvain Bromberger</u>, ed. Ken Hale and Samuel J. Keyser, 53-110. Cambridge, Mass.: MIT Press.

Higginbotham, James. 1983. Logical form, binding, and nominals. <u>Linguistic Inquiry</u> 14:395-420.

Higginbotham, James. 1985. On semantics. <u>Linguistic Inquiry</u> 16:547-593.

Hoffman, Joel. 1995. Syntactic and paratactic word order effects. Doctoral dissertation, University of Maryland, College Park.

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

Izvorski, Roumyana. 1996. The syntax and semantics of correlative proforms. <u>Proceedings of NELS</u> 26, 133-147.

Izvorski, Roumyana. 1997. On the nature of wh-infinitival complements of possessive and

existential predicates. Paper presented at the 6th Annual Workshop on Formal Approaches to Slavic Linguistics, University of Connecticut, Storrs.

Kayne, Richard. 1991. Romance clitics, verb movement, and PRO. Linguistic Inquiry 22:647-686.

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

Lasnik, Howard. 1995. Verbal morphology: Syntactic Structures meets the Minimalist Program. In Evolution and revolution in linguistic theory: Essays in honor of Carlos Otero. ed. Paula Kempchinsky and Héctor Campos, 251-275. Washington D.C.: Georgetown University Press.

Moritz, Luc. 1989. Aperçu de la syntaxe de la négation en français et en anglais. Mémoire de licence, Université de Genève.

Müller, Gereon, and Wolfgang Sternefeld. 1996. A'-chain formation and economy of derivation. Linguistic Inquiry 27:480-511.

Pesetsky, David. 1987. Wh-in-situ: Movement and unselective binding. In The representation of (in)definiteness. ed. Eric J. Reuland and Alice G.B. ter Meulen, 98-129. Cambridge, Mass.: MIT Press.

Pesetsky, David. 1992. Zero syntax, vol. 2. Ms., MIT, Cambridge, Mass.

Pollock, Jean-Yves. 1989. Verb movement, universal grammar, and the structure of IP. Linguistic Inquiry 20:365-424.

Reinhart, Tanya. 1995. Interface strategies. OTS Working Papers, Research Institute for Language and Speech, Utrecht University.

Rivero, Maria-Luisa. 1991. Long head movement and negation: Serbo-Croatian vs. Slovak and Czech. The Linguistic Review 8:319-351.

Rizzi, Luigi. 1990. Relativized minimality. Cambridge: MIT Press.

Roberts, Ian. 1992. Verbs and diachronic syntax. Dordrecht: Kluwer.

Rudin, Catherine. 1988. On multiple questions and multiple wh-fronting. Natural Language and Linguistic Theory 6:455-501.

Stjepanović, Sandra. 1995. Short-distance movement of wh-phrases in Serbo-Croatian matrix clauses. Ms., University of Connecticut, Storrs.

Takahashi, Daiko. 1994. Minimality of movement. Doctoral dissertation, University of
    Connecticut, Storrs.

Takahashi, Daiko. 1997. Move F and null operator movement. <u>The Linguistic Review</u> 14:181-
    196.

Tsai, Wei-tien Dylan. 1994. On economizing the theory of A-bar dependencies. Doctoral
    dissertation, MIT, Cambridge, Mass.

Watanabe, Akira. 1992. Wh-in-situ, subjacency, and chain formation. In <u>MIT Occasional Papers
    in Linguistics</u> 2. Department of Linguistics and Philosophy, MIT, Cambridge, Mass.

Williams, Edwin. 1986. A reassignment of the functions of LF. <u>Linguistic Inquiry</u> 17:265-299.

Notes

1. For helpful comments and suggestions, I thank Howard Lasnik, Roger Martin, and audiences at various seminars at the University of Connecticut, Formal Approaches to Slavic Linguistics 5 hosted by Indiana University and Wabash College, Kanda University, Nanzan University, and the Linguistic Symposium on Romance Linguistics 27 hosted by University of California at Irvine, where portions of this material were presented. Particular thanks are due to Jun Abe, Masao Ochi, Mamoru Saito, Yuji Takano, Juan Uriagereka, and Akira Watanabe. Thanks are also due to a number of people, particularly Michèle Bacholle, Cédric Boeckx, Viviane Déprez, and Géraldine Legendre, for their help with judgments. Some material from Bošković (1997b) has served as a starting point for sections 2 and 3, where this material is considerably expanded and revised.

2. The same holds for Chomsky (1993). Chomsky (1993) considers strong features illegitimate PF objects, which does not in principle rule out LF insertion of elements with strong features.

3. Chomsky (1995) also considers the possibility that the null complementizer in constructions such as (5b) is inserted in LF, but ultimately rejects it for reasons to be discussed below.

4. *Pourquoi* 'why' is actually an exception in that it cannot be left in-situ. Rizzi (1990), however, shows convincingly that the reason for this is trivial: *Pourquoi* must be base-generated within the CP projection.

5. Aoun (1986) offers an account of the ungrammaticality of (6b) based on his proposal that selectional restrictions must be satisfied at SS in French. The construction is then ruled out because *demander* does not take a +wh-complement at SS. This account is not available in the current theoretical framework, which does not attach any syntactic significance to SS. The fact that the current analysis does not invoke SS to account for the ungrammaticality of (6b) can in fact be considered a plus for the analysis. Under this analysis, there is no need to check whether selectional restrictions are satisfied in French before LF. Note also that in order to account for the

contrast between French (6b) and constructions such as Japanese (7), Aoun proposes that, in contrast to French, in Japanese selectional restrictions must be satisfied only at LF. The difference between French and Japanese concerning when selectional restrictions must be satisfied is inconsistent with the current view of cross-linguistic variation, where languages are considered to differ only in their morphological properties. (I return to pure wh-in-situ languages below.) As will become obvious below, Aoun's (1986) analysis is also empirically inadequate. It leaves several facts concerning the availability of the in-situ strategy in French discussed below unaccounted for.

6. Note that SC indirect questions have the same form as matrix questions. As a result, special care must be taken to rule out the possibility of analyzing a SC indirect question as a superficial direct question, with the superficial matrix clause being analyzed as an adsentential. To rule out this possibility, no pause should be made before the embedded question. (Note that the matrix clause needs to be given more content, which means avoiding a matrix clause that contains only a verb, to make it less likely to be analyzed as an adsentential). No danger of analyzing the embedded question as a superficial matrix question arises in SC correlative constructions (see (11c-d)) and existential constructions such as (11e-f), whose wh-clause has all the formal properties of questions (see Izvorski 1996, 1997). Note also that I ignore the irrelevant echo question reading of wh-phrases in (11).

7. *Li* is used in regular yes-no questions. With wh-phrases, it is used in emphatic questions.

8. Pesetsky (1987) also assumes that the interrogative complementizer in English is a phonological affix. For relevant discussion see also Pesetsky (1992), who conjectures that null morphemes are by and large specified as phonological affixes.

9. An alternative that may be worth pursuing is to assume that the interrogative C in embedded questions adjoins to the higher verb overtly, thus satisfying its affixal requirement. This option is not available in matrix questions such as (15) because it would involve overt lowering.

10. This seems to allow for the possibility of having a number of separate trees in the overt syntax, which are then assembled together in LF. This possibility is taken advantage of in Hoffman (1995), who argues that this is exactly what happens in free-word order languages. However, it seems that the possibility should be ruled out in PF, since it is difficult to see how some PF operations, for example, intonational phrasing, can apply properly unless the whole structure is assembled into one tree before it enters PF.

11. It is true that the notion of strength was originally posited to drive overt movement. However, the way it is implemented in Chomsky (1995) leaves room for having strength drive covert movement in certain well-defined configurations. I see no reason why this possibility should not be taken advantage of, especially when this move is empirically well-motivated, which I attempt to show in this paper. (See also the discussion in the following section.)

12. To rule out LF insertion of phonologically null elements with strong features Chomsky has to adopt an additional assumption, given in (i):

(i) α enters the numeration only if it has an effect on output.

Chomsky argues that strength in an interrogative complementizer has neither LF nor PF effect if the complementizer is introduced into the structure covertly. As a result, given (i), strength cannot be present in the numeration, which is defined as an array of lexical items that is mapped by the computational system into a linguistic expression (a PF-LF pair), to start with. The condition in (i) is conceptually very problematic. In order to determine the effects of (i) we need to know PF and LF outputs. However, the numeration, which is determined by (i), must be present in the initial stage of the derivation. This globality raises a very serious conceptual problem. (i) also has a number of what appear to be undesirable empirical consequences that are not explored in Chomsky (1995). Thus, (i) appears to ban constructions containing INFL with a strong D feature and null subjects (pro, PRO, null operator), since the movement of the null subject to SpecIP, motivated by the strength of INFL, has no effect on the outputs. That is, it seems that languages that have null subjects of any kind cannot have a strong D INFL in the contexts in which these null subjects appear. This may be too strong.

The basic idea behind (i) is that superfluous projections are not allowed in representations. A number of ways of implementing this idea have been proposed in the literature that do not face the problems that (i) faces and do not have the effect of altogether ruling out covert insertion of elements with strong features (see Bošković 1997c and references therein). Bošković (1997c) in fact shows that given a small modification of the notion of numeration and the way lexical insertion works, the ban on superfluous projections in representations can be made to follow from the independently needed Last Resort Condition, which bans superfluous steps in derivations. That is, Bošković (1997c) shows that it is possible to use the Last Resort Condition to prevent creation of representations with superfluous projections, which makes any condition that is intended to specifically rule out such projections redundant and therefore eliminable.

13. The judgments of my informants for (19a) range from fully unacceptable as a true non-echo question to clearly worse than (5b). Note that (19a) sounds best with stress on the wh-phrase in-situ, an indication of the echo-question reading. (The similar situation holds for (21a), discussed below, which has a similar status as (19a)).

It is worth noting here that (19a) is not as bad as (6b) on the true question reading. This can be accounted for under the analysis presented below, since (6b) and (19a) violate different conditions.

Note that the direct object wh-phrase in (20a) has the matrix scope on the relevant reading. The construction is clearly worse than Subjacency violations on the matrix true question reading of the direct object wh-phrase.

14. In this paper I concentrate on finite wh-in-situ complements. Infinitival wh-in-situ complements are discussed in Bošković in press b.

15. It is worth noting here that Chomsky (MIT Lectures 1995) suggests that when X undergoes overt XP movement to SpecYP there are actually two movements involved: Move F first adjoins formal features of X toY for feature checking and then the rest of X undergoes XP movement to SpecYP ('pied-piping') followed by a repair strategy that makes X pronounceable. Under this analysis, it is not possible to make LF movement more constrained than overt movement by appealing to Move F, which I attempt to do here based on the data under consideration (for relevant data, see also section 4). Therefore, if the discussion here is on the right track the two separate movements analysis cannot be correct: the decision to 'pied-pipe' must be made immediately so that only one *actual* movement takes place (XP moves to SpecYP), as originally suggested by Chomsky (MIT Lectures 1994) and Chomsky (1995). (Chomsky 1995 is somewhat ambivalent on this issue. However, he crucially assumes throughout chapter 4 that the checking configuration is Spec-head for overt syntax and FF(adjoined to head)-head for covert syntax, which goes against the spirit of the two movements analysis.)

16. Takahashi (1997) also reaches the conclusion that feature-movement is subject to locality restrictions on movement. He claims that elements from which feature movement takes place can block movement of their own features via Huang's (1982) Condition on Extraction Domain (CED), i.e., if they are subjects or adjuncts. The data considered here, however, provide evidence against this claim. Recall, for example, that French (9) must involve feature movement of the wh-phrase to C, after C is inserted in LF. Since the wh-phrase is an adjunct, (9) must involve feature movement out of an adjunct. Its grammaticality, then, provides evidence that phrases do not block movement of their own features. The same conclusion is reached in Bošković (1997c). (It is worth pointing out here that, as noted by Takahashi (1997), Takahashi's (1994) minimalist account of CED effects does not rule out feature movement of or, to be more precise, out of adjuncts and subjects.)

17. Actually, the Attract system fails to account even for the full range of Wh-Island effects, which are supposed to be its show-case. In this system, the Wh-Island effect is captured by appealing to feature-checking instead of the A/A' distinction. Thus, (i) is ruled out because the matrix C, which needs to check its +wh-feature, fails to attract the closest +wh-feature bearing element (*where*).

(i) ??Which book$_i$ do you wonder where$_j$ John put t_i t_j

This seems to leave (ii) unaccounted for.

(ii) ??(Peter thinks that) That book$_i$ you wonder where$_j$ John put t_i t_j

It is not at all clear why the +wh-feature should be relevant in attraction of topics. A similar problem arises with respect to a number of other constructions, for example, relativization out of wh-islands (cf. ?*The book that you wonder where John put*) and *tough*-movement out of wh-islands (cf. ??*This car is tough to ask Peter when to repair*). Chomsky's (1995) system, where feature-checking is intended to do the job of the A/A' distinction with respect to relativized minimality, thus fails to account for the full range of wh-island effects. Several other types of

relativized minimality effects with A'-movement also appear to remain unaccounted for in this system (for example, Rizzi's 1990 Pseudo-Opacity effects and Inner Island effects.) It is my belief that a fully successful way of capturing the effects of the A/A' distinction on relativized minimality type phenomena would also cover the cases discussed in the text.

18.Note that since the relevant wh-phrases do not undergo LF wh-movement, we can now dispense with the stipulation that Subjacency does not apply in LF, a position which I adopt here.

19.The construction is bad if the order of the wh-phrases is reversed, which, as argued in Rudin (1988) and Bošković (1997a), is a Superiority effect. Note that, as argued extensively in Rudin (1988) and Bošković (in press b), adjunction to SpecCP in Bulgarian proceeds to the right, so that the wh-phrase that comes first in the linear order is the one that moves first to SpecCP. For an account of the full range of Superiority effects in Bulgarian based on the Economy approach to Superiority, see Bošković (1997a, in press a).

(i) *[$_{CP}$ [[Kak] koj] [$_{C'}$ e   kupil   kniga]]
           how who   is bought   book

20.The assumption was rather strange, given that c-command seems to hold between the adjoined element and its trace and given that at the same time it was apparently assumed that antecedent government is possible from the SpecIP adjoined position. This was necessary to account for expletive *there* constructions, where the associate of the expletive was assumed to undergo adjunction to the expletive in LF (see Chomsky 1991).

21.They show that the standard argument-adjunct asymmetry is actually a noun-adverb asymmetry (for relevant discussion, see also Huang 1982), resulting from the presence of an open position in NPs, but not adverbs. The contrast between *Who left how* and *Who left what way* is instructive here.

22.Tsai and Reinhart also argue that wh-adverbs must be located within the interrogative CP projection in order to be properly interpreted.

23.Admittedly, the account of the different behavior of German and English/French wh-adverbs with respect to the possibility of remaining in situ in questions adopted here is not very deep. However, neither are the alternative accounts of this fact I am aware of. Thus, Müller and Sternefeld (1996) (see also the references therein) account for the fact under consideration by stipulating that German differs from English in that in German, but not in English, antecedent-government is possible from the SpecCP adjoined position. Haider (1986), on the other hand, suggests that the trace left by LF wh-movement of adverbs is lexically governed in German, but not in English.

24.This is in line with the assumption that the same kind of feature, with a potential difference in strength (the situation is not clear here given that, as argued in Bošković 1997c, strength is not the only thing that can force overt wh-movement), is involved in partial wh-movement, which is not an obligatory operation.

25.A potential problem here is that the movement of the wh-operator could be feature movement in spite of taking place overtly. Since the wh-operator is phonologically null, there seems to be no need for category pied-piping to ensure proper PF interpretation. This is un undesirable result, since we clearly want the movement in question to at least have the option of being XP movement, in order to avoid strict locality restrictions on X^0 movement. There are two ways out of the potential problem. One possibility is that the wh-operator is actually not phonologically null. The operator moves to SpecCP, after which it undergoes PF deletion, perhaps as a reflex of the traditional Doubly Filled Comp Filter (see Chomsky and Lasnik 1977). To be able to undergo PF deletion, the operator would have to enter PF as a whole, rather than 'broken' into features. This is actually exactly what Takahashi (1997) proposes happens in English constructions such as *the reason that John said that Mary left,* which, according to Takahashi, who essentially follows Chomsky and Lasnik (1977), involves movement of a phonologically overt element to SpecCP, followed by PF deletion of the element. (Note that the movement in question clearly has properties of XP movement.)

   An alternative analysis is available that seems more principled to me. Chomsky argues that when one formal feature undergoes movement, all formal features are affected by the movement. Movement minimally affects all formal features (FF). Returning now to the wh-operator, it is not clear that the operator contains anything but formal features. In fact, +wh may be the only feature the operator has. FF (wh-operator) would then be the whole wh-operator, (i.e., FF would not be a part of the lexical item in question, but the whole lexical item.) Given the contextual definition of X^{max} and X^{min} that Chomsky adopts (for evidence for this approach to the X^{max}/X^{min} distinction, see Bošković 1997c), FF(wh-operator) would also be a maximal projection (it is a lexical item that does not project any more) as well as a minimal projection (it is a lexical item that is not a projection at all.) Chomsky suggests that such elements are able to undergo both XP and X^0-movement. This would enable FF(wh-operator) to undergo XP-movement, which is a less constrained option. (It is not clear whether the final landing site of the movement could still be the C-adjoined position, and not only SpecCP.) In fact, as Masao Ochi (personal communication) points out, the null operator could undergo XP movement even if it were to move in LF (by definition, X^{max} would be undergoing movement) so that, for our purposes, it is actually not necessary for the null operator movement to be overt.

26.As for C and the CP projection, it is not quite clear whether they would ever be present in (31). For arguments that there is no C or CP projection in *that*-less declarative clauses, see Bošković (1992, 1996, 1997c).

27.Whether (32) would be a bare a VP if the IP projection could be inserted in LF is not clear, since it depends on whether there are any projections between VP and IP. A number of proposals concerning the existence of such 'intermediate' projections have been made in the literature. If there are such projections, and if any of them is specified as a phonological affix, it would have to be present overtly. Note that what I called the IP projection could be split into a number of functional heads, some of which could be PF affixes and, if so, would have to be present overtly. Recall also that Pesetsky (1992) argues that phonologically null heads are by and large cross-linguistically PF affixes, which would make any phonologically null head above the VP in (32) likely to be an affix.

28.There is another potentially interfering factor in (34). As noted by Belletti (1990), even sentential adverbs in French cannot occur in front of a subject, though this should be a base-generated option for them even when the subject is located in SpecIP (cf. English (ib)).

(i) a. ?*Probablement   Jean vendra   ces   livres.
          probably   Jean will-sell   these books
    b. Probably, John will sell these books.

It is possible that Adv-subj-V sequences are for some reason ruled out in French, i.e., it is possible that whatever rules out (ia) rules out (34). This is plausible if the ungrammaticality of (i) is not due to structural reasons. (See, however, Belletti 1990 for an attempt at a structural explanation of the ungrammaticality of (ia).) Notice, however, that some adverbs can appear before the subject in French, as illustrated by *Souvent Pierre se trompe* 'Often Pierre makes mistakes'. (The adverb must have strong stress though.)

29.Chomsky (MIT Lectures 1995) leaves open what drives movement of V to INFL, which seems to completely fall outside his system.