Optimal Questions[*]

Peter Ackema & Ad Neeleman
Utrecht University/OTS

1. *Constraint Interaction*

The traditional view of the organization of grammar is that, although facts are derived by a conspiracy of principles, the principles themselves do not compete. That is to say, no principle is violated in order to satisfy another principle. In a grammatical sentence, all principles are satisfied.

In the minimalist program (cf. Chomsky 1995) this picture is slightly different, due to the important role assigned to economy principles. Economy principles are not meant to evaluate a single derivation, but to choose from several possible derivations the one that best satisfies them. This means that with respect to economy conditions, though not with respect to other conditions, there is a difference with the traditional view in that economy conditions do not impose absolute requirements on a derivation or representation. This shifted perspective also implies that it is no problem if an economy condition conflicts with another constraint. In fact, all economy conditions disfavoring movement conflict with the conditions that trigger movement. For example, economy principles like Procrastinate and Shortest Steps conflict with the requirement to check strong features. This conflict is resolved by the assumptions that non-checking of a strong feature leads to a crashing derivation (i.e. 'check strong feature in overt syntax' is an absolute constraint on converging derivations) and that global economy takes into consideration converging derivations only. For example, in English it is necessary to place a WH in spec-CP to get a well-formed WH-question, despite the economy requirement to minimize overt movement. The fact that some constraints (the need to check strong features) must overrule others (Procrastinate, Shortest Steps) implies that these constraints are of different importance, which virtually means that they are ranked.

Constraint ranking is the key notion of optimality theory (Prince & Smolensky 1993). In the minimalist program 'ranking' is restricted to two types of constraints, absolute constraints and economy constraints, where in every language the absolute constraints outrank the economy constraints. Optimality theory proposes a more radical

---

breach with the traditional view. There are no absolute constraints any longer (though see note 7). All constraints are ranked and evaluated collectively. Evaluation proceeds as follows. The structures to be compared are first evaluated with respect to the highest-ranked constraint. In the event of a tie, the remaining candidates are judged by the next highest constraint, and so on. Constraints therefore need not be surface true. A lower ranked constraint can be violated in an optimal form when this form scores better on higher ranked constraints than its competitors. (Note that even the highest-ranked constraint can be violated, namely in case there is no potential output that does not violate it).

Moreover, whereas in the minimalist program the ranking of economy constraints and other constraints is constant across languages, optimality theory seeks to explain language variation as a consequence of different rankings of constraints. Each possible ranking of the constraints must determine the grammar of a natural language (and each grammar of a natural language must be determined by some ranking of the constraints). This means that there is an interesting way of testing the validity of optimality theory. Given some well-motivated principles handling data in a specific language, reranking these constraints should result in a (partial) grammar of another language. If this indeed turns out to be the case, it is a strong argument for constraint ranking and reranking as proposed in optimality theory.

In this paper we will argue that language variation in the syntax of questions, in particular multiple questions, can indeed be accounted for by different rankings of a few well-motivated constraints. The paper is structured as follows. First, we will try to establish the constraints operative in the syntax of questions by considering English (section 2). We will then show how different rankings of these constraints account for the syntax of WH-questions in different languages. We will discuss Bulgarian, Czech, Chinese and French (section 3). The paper concludes with a brief summary of the proposed analysis (section 4).

2. *The Syntax of Questions*

Before we can test whether constraint reranking can account for the syntax of questions cross-linguistically, we must find out which constraints are operative in question formation. This is the topic of this section. In section 2.1, we will outline the theory of clause structure within which the analysis will be couched. In section 2.2, we will develop a minimal theory of question formation. From this theory two constraints can be derived, which in combination with the well-known Shortest Steps condition explain the syntax of questions in English (section 2.3).

2.1 Functional Projections and Head Movement

In the minimalist program, or more specifically in checking theory, WH-movement and the verb movement that accompanies it in some languages must be driven by the need to check WH-features. The difference between overt and covert WH-movement should be regulated by the strength of these features. In this paper we will take issue with this analysis, for the following reasons.

The theory of feature checking is only descriptively adequate, and not explanatorily, if no satisfactory definition of the strength of features can be given. Especially in the case of WH-features, this seems problematic. Compare for instance (1a) and (1b).

(1)   a.   Who(m)$_i$ did John talk to t$_i$

      b.   What$_i$ did John give t$_i$ to who(m)

      c.   *John talked to who(m)

In a theory based on feature checking, one is forced to conclude on the basis of (1a) that *who(m)* has a strong WH-feature. If that were so, however, (1b) should crash. The hypothesis that *who(m)* can have either strong (in 1a) or weak (in 1b) WH-features is untenable. First, it is ad hoc since no morphological difference is detectable between strong and weak *who(m)*. Second, the ungrammaticality of (1c) (under a nonecho reading) is left unexplained, if a *who(m)* with weak features exists. Third, checking theory leads to a conceptually unattractive duplication of features in the syntactic representations. The WH-feature is not only present on the WH-expression (a necessity), but also in the functional position in which checking must take place.[1] Similar problems arise in connection with the accompanying verb movement. Given the principle of *Greed*, as assumed in checking theory, this movement must be caused by the need for the verb to check a strong WH-feature. The presence of inversion in (2a) versus the absence of inversion in (2b) then indicates that a verb like *will* has a variant with and without (strong) WH-features, even though no morphological evidence for this difference exists.

---

[1] This double representation of features might be used to solve the problem posed by (1) for checking theory. The solution consists of the assumption that the WH-feature of WH-words can be either strong or weak in English (as noted in the text this is an ad hoc assumption anyway, since there never is any morphological difference between the strong and the weak variant of the WH-word), while the WH-feature of C is always strong. This solution seems unattractive to us, since it assumes that the same feature can be strong or weak according to the position it happens to be generated in. This option is not otherwise attested, as far as we know. (In standard minimalism, features are assumed to be strong or weak, period). Note that it cannot be assumed that the WH-feature of C is a different feature from the WH-feature of the WH-expression, since in that case checking would not be possible.

(2) a.   What$_i$ will$_j$ John [t$_j$ see t$_i$]
     b.   John [will see a cow]

It seems, then, that in the domain of WH-movement, checking theory faces some serious difficulties, because ad hoc assumptions must be made about the featural content of lexical items.[2]

An alternative theory of question formation (that is, a theory of WH-movement and verb movement to functional positions) seems to be in order. We will now try to develop such a theory, based on the following two assumptions. First, functional structure is only generated when necessary. Second, generation of functional structure can be necessitated by the requirement that certain semantic functions of VP are syntactically marked. One of these functions would be that of 'question'. The remainder of this section deals with the first assumption. The second one is discussed in section 2.2.

Let us assume that, given the central role of economy, structure is only generated when required for wellformedness. Sentences minimally contain a proposition, which is syntactically encoded as a juncture of a subject and a verbal predicate (VP or V', depending on one's assumptions about the position of the subject). The question is under what circumstances further (functional) structure on top of this VP will be present.

Given a theory of phrase structure as proposed in Chomsky (1994), the insertion of a functional marker after the VP is generated leads to a functional projection headed by that marker. There is no other option, since further projection of the VP does not provide the functional marker with an appropriate position.

(3) a.   F'   b.   *   VP

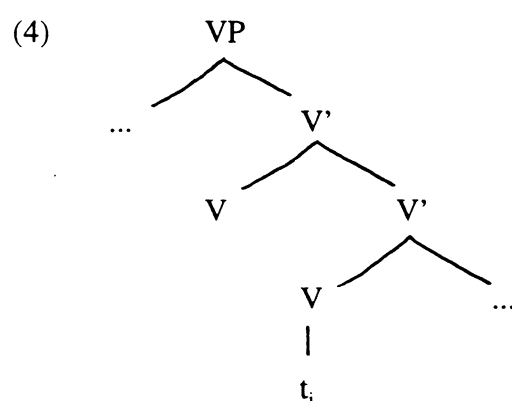   F   VP   F   VP/V'
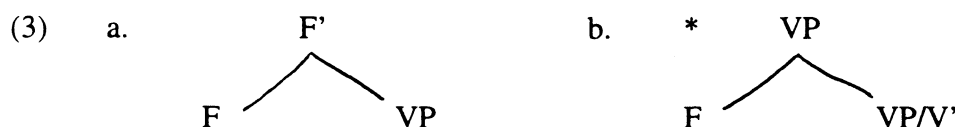
If we now take seriously the idea that projection depends on the presence of a lexical head, we are forced to the conclusion that the structure in (3a) is not base-generated if no functional marker is present. If only an inflected verb is taken from the lexicon, then only a VP will be base-generated.[3]

---

[2] In fact, there is an additional problem, pointed out by Reinhart (1993). According to her and others, WHs in situ do not move at all, not even at LF. If this is so (as we will assume for the largest part, see section 3), it is hard to see how their features could be checked.

[3] In accordance with the lexicalist perspective on inflection, we assume that verbs are inserted in fully inflected form. See Lapointe (1980), Lieber (1980), Jensen & Stong-Jensen (1984), Di Sciullo & Williams (1987), Chomsky (1993) and Ackema (1995) for various proposals. Moreover, the view expressed here presupposes that the VP-internal subject hypothesis is correct.
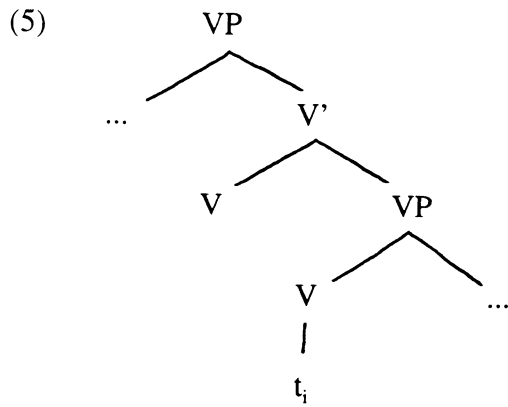
However, when necessary, functional structure may be derived in the absence of a functional marker in the numeration. In order to accomplish this we adopt a procedure proposed by Ackema et al. (1993). This procedure generates functional structure as a result of head movement. As opposed to Chomsky (1994), we would like to allow for the possibility of *self-attachment*. By this we mean that a head can move out of its base position and adjoin to a node of its own projection.

Suppose that, after VP has been generated, the verb moves out of its base position and adjoins to V'. The resulting structure would seem to be ungrammatical. It is standardly assumed that every head must be dominated by a maximal category and every (possibly segmented) maximal category must dominate a head. But in (4) one maximal projection dominates two heads.

(4)

```
           VP
          /  \
        ...    V'
              /  \
            V      V'
                  /  \
                V      ...
                |
                t_i
```
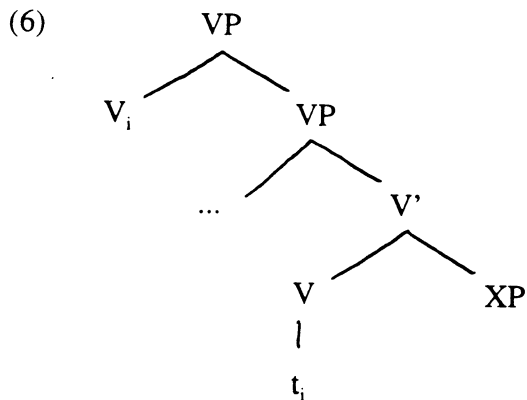
This problem is solved by Ackema et al. by assuming that X-bar theory has the following two properties. First, it applies at every stage of the derivation (cf. Van Riemsdijk 1989, Chomsky 1993). Second, the familiar X-bar levels are decomposed into the features [±M(aximal)] and [±P(rojected)] (cf. Muysken 1982), where at least the value of [±M] is defined contextually, namely as the highest projection of a head X. As a result, it follows that nodes which are nonmaximal at one point in the derivation may be maximal at another (and vice versa).
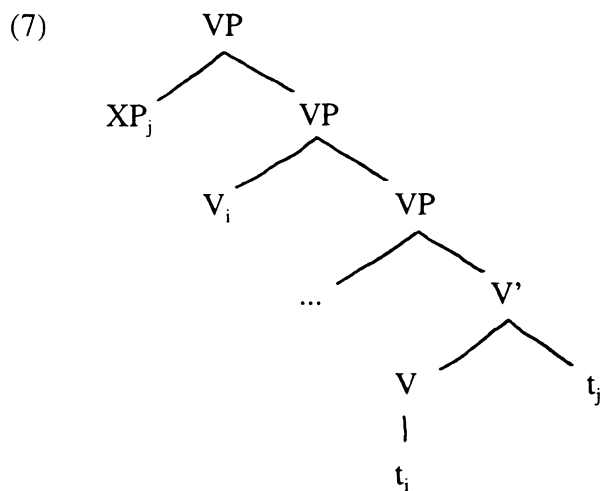
The structure in (4) is now rescued if the lower segment of V' is reanalyzed as VP, a reanalysis which takes place within the limits of the X-bar theory sketched above. The V'-node, which initially has the features [+P,-M], turns into a node with the features [+P,+M] at surface structure. In other words, this node is now conceived of as the highest projection of the verbal trace:

(5)

```
              VP
             /  \
          ...     V'
                 /  \
                V    VP
                    /  \
                   V    ...
                   |
                   tᵢ
```
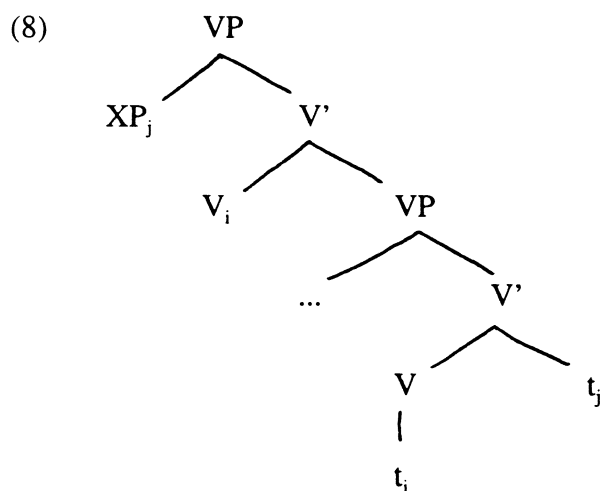
This is one way of establishing the structure in (3a) by movement. A second way, which will be the crucial one in this paper, also involves self-attachment. In this case, the verb adjoins not to V', but to VP:

(6)

```
              VP
             /  \
          Vᵢ     VP
                /  \
             ...    V'
                   /  \
                  V    XP
                  |
                  tᵢ
```

There is an important difference between self-attachment of the verb to V' and self-attachment to VP. In the former case, the specifier of the functional VP will contain the same element that the specifier of the lexical VP contained in the initial structure (without this element having actually been moved, compare (4) and (5)). Self-attachment to VP, however, creates the possibility of deriving a different specifier in the functional projection. This is accomplished as follows. First, a second element moves out of its base position and adjoins to VP. This could be the complement of V, as in (7), but in fact any element will do.

6

(7)

```
                VP
              /    \
          XP_j      VP
                  /    \
               V_i      VP
                      /    \
                   ...      V'
                          /    \
                        V       t_j
                        |
                        t_i
```

The element adjoined to VP can be turned into the specifier of the functional VP by reanalysis of the intermediate segment of VP. If this node's positive marking for maximality is changed into a negative marking, a structure results in which the moved V heads a functional VP with the moved XP in its specifier:[4]

(8)

```
                VP
              /    \
          XP_j      V'
                  /    \
               V_i      VP
                      /    \
                   ...      V'
                          /    \
                        V       t_j
                        |
                        t_i
```

An instantiation of (8) is question formation in (English) root sentences, as we will show below.

---

[4] It might be that the intermediate structures in (6) and (7) face some X-bar theoretical problems. The higher VP-nodes in these structures can either be analyzed as a projection of the moved verb, or as a projection of the lower VP (with the verb a [-P,+M] adjunct to this VP). If Chomsky (1994) is correct in claiming that syntactic representations may not have ambiguous structural properties, structures like (6) and (7) would be ruled out. This problem is solved by reanalysis as in (8), resulting in a structure that is not ambiguous. This means that, if Chomsky is correct, self-attachment of the verb to VP must be accompanied by the fronting of some constituent. See for discussion Koeneman (1995).

In sum, functional structure can be generated in two ways, namely by the insertion of a functional marker or by self-attachment. The latter option in turn consists of two subcases. If the head adjoins to an intermediate projection, reanalysis will result in a structure with a functional projection on top of VP that contains the same specifier as the VP contained in the base. If the head adjoins to its maximal projection, fronting of some element and reanalysis will result in a functional projection on top of VP with the fronted element in its specifier position.

2.2 Q-Marking

We have now seen how functional structure can be derived, but we have not yet considered the question why and when it is derived. We assume that one important reason to create functional structure is to derive a configuration in which certain properties of VP (i.e. the proposition) can be syntactically marked. VP can be marked for such notions as tense, modality and aspect, and, as we will argue in this section, for questionhood.

It is a pervasive property of natural language that heads may mark properties of their complements. A verb may, for instance, assign a thematic role and possibly case to its complement. It would therefore be a minimal extension of existing theories to assume that sentence-type marking in the sense just outlined involves a head and its complement. The general scheme for the marking of a property [+P] is the one in (9).

(9)   F'
           ⟋‾‾‾⟍
       F [+P] ⟶ YP
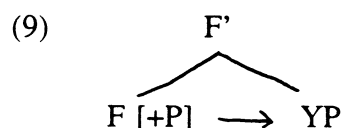
The marking of VP as a question is an instance of the general scheme in (9). In English, for instance, questions are clearly marked as such. A straightforward instance of Q(uestion)-marking is found in embedded yes/no questions. Here a base-generated functional head bearing the Q-feature takes the VP to be marked as its complement:

(10)   John wonders [if [$_{VP}$ you have seen the soccer match]]
   +Q ⤴
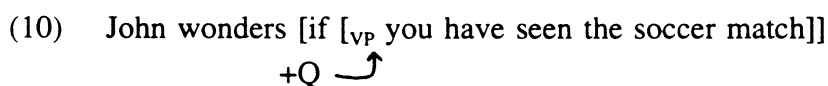
It is not obvious, however, that question marking always involves a head and its complement. In matrix WH-questions, for instance, the relevant feature seems to be present on the WH-expression, and not on the verb. Moreover, one may wonder whether Q-marking by a complementizer, as in (10), is an instance of the same process as Q-marking by a WH-expression, as in (11).

(11)   [What$_j$ [have$_i$ [you t$_i$ seen t$_j$]]]

To start with the first question, it is hardly a problem that the relevant feature is present in the specifier of the functional projection. By the general mechanism of specifier-head agreement, we may expect that a WH-specifier will transfer this feature to the related head, after which this head will be capable of marking its complement. This means, of course, that such a head must be present. In the absence of a complementizer in this root environment, self-attachment of the verb is necessary. If the verb did not move, there would be no head to mark the proposition as a question.

Assuming for the moment that the relevant feature indeed is the Q-feature we encountered earlier, (11) can be depicted as in (12), where, in line with assumptions made earlier, the projection hosting the WH-expression is labelled VP since it is headed by (the moved) V:

(12)   VP
          /   \
    WH [+Q]   V'
              /   \
          V [+Q]   VP
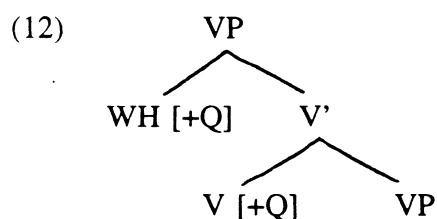
So, the derivation of a WH-question contains the following steps. First, the verb adjoins to VP (cf. 13b). Then, the WH-expression adjoins to VP (cf. 13c), and then VP-to-V' reanalysis takes places (cf. 13d).[5]

(13)   a.   [$_{VP}$ You have seen what]
        b.   [$_{VP}$ Have$_i$ [$_{VP}$ you t$_i$ seen what]]
        c.   [$_{VP}$ What$_j$ [$_{VP}$ have$_i$ [$_{VP}$ you t$_i$ seen t$_j$]]]
        d.   [$_{VP}$ What$_j$ [$_{V'}$ have$_i$ [$_{VP}$ you t$_i$ seen t$_j$]]]

Let us now turn to the second question posed above: can the Q-feature of the interrogative complementizer and the WH-feature of a WH-expression be equated? The null hypothesis would be that both are identical. Although this might seem problematic, given that the Q-feature marks yes/no questions (cf. 10), while the WH-feature marks the questioning of a particular constituent (cf. 11), there are two observations that support unification. First, there are verbs like *wonder*, which select both *if* clauses and WH-questions (cf. 14), but there are no verbs that select one of these categories but not the other:

---

[5] This derivation does not comply with Greed, at least not in its strictest form. We therefore assume that this principle does not hold.

(14)   John wonders [what [Mary will say]]

Second, in embedded WH-questions in Dutch, both a Q-complementizer and a WH-expression may be present (cf. 15). If the features of these elements are incompatible, one would expect the derivation to crash.

(15)   Jan vraagt zich af [wat [of [Marie zal zeggen]]]
       John wonders REFL PART *what if Mary will say*

If there is a uniform category 'question', the different interpretations of WH-questions and *if* questions must be derived in some other way. In fact, this can easily be achieved. WH-questions typically contain a variable accessible to the Q-marking head.[6] We assume that as a result of this the feature Q is interpreted as concerning this element. In yes/no questions there is no gap, and therefore the feature Q can only be interpreted as concerning the whole expression (i.e. its truth value).

The hypothesis that there is a general category 'question' and that English requires syntactic marking of this category faces two problems. First, in embedded WH-questions, there is a Q-bearing element, but there does not seem to be a head (cf. 16a). Second, in matrix yes/no questions there is a head, but there does not seem to be a Q-bearing element (cf. 16b). By earlier assumptions we are forced to say that (16a) contains a null head and that (16b) contains a null Q-operator. We believe that this is indeed what is going on.

---

[6] By 'accessible', we mean a variable that is not bound within the c-command domain of the Q-marking head. So, in (ia) a WH-question interpretation is possible, but in (ib) and (ic) it is not.

(i)   a. WH_i C[+Q] [$_{VP}$ ... t_i ...]
       b.   C[+Q] ... WH_i ... [$_{VP}$ ... t_i ...]
       c.   C[+Q] [$_{VP}$ ... WH ...]

In (ia) the variable is accessible to the Q-marking head, since it is unbound within the c-command domain of this head. In (ib) there is a variable in VP, but it is not accessible to the Q-marking head, as it is already bound by its WH-antecedent. In (ic) there is no variable at all (at surface structure). As pointed out by Eddy Ruys (p.c.), this explains the illformedness of the following example:

(ii)   ??Wie_i vraag je [t_i of [Jan slaat t_i]]
       who ask you if/whether John hits

This example cannot be ruled out on a par with WH-islands, since *of* 'if' is the head, not the specifier of CP. The example is not so much ungrammatical, however, but uninterpretable. The trace in the embedded clause is accessible to C, hence the embedded clause should be interpreted as a WH-question. At the same time, the matrix clause must be interpreted as a WH-question as well. Since there is only one WH-operator present, this is impossible.

(16) a. John wonders [what ___ Mary will say]
      b. ___ have you seen the soccer match

Concerning (16a), we have already seen that in certain languages the Q-marking head can be lexicalized. A Dutch example was given in (15). Apparently, embedded WH-questions do contain a Q-marking head. It is unclear, however, what the status of this head is in the theory of phrase structure we have adopted. Recall that functional structure could only be generated by insertion of lexical markers or self-attachment. Neither of these possibilities seems to apply in (16a).

This problem can be solved in several ways. The solution that seems most promising at the moment would be to say that a [+Q] complementizer is in fact inserted in embedded WH-clauses in English as well, but that, as a consequence of phonological constraints on the lexicalization of 'COMP', it is deleted (i.e. not pronounced) when the syntactic representation is mapped into a phonological one (cf. Pesetsky 1994). The relevant phonological constraint is usually referred to as the doubly filled COMP filter.

One might object to such an analysis by saying that *if* (as opposed to *that*) may usually not be omitted. This, however, can be seen as a result of the condition of recoverability. As is well-known, deletion is only possible if the features of the deleted material are recoverable from the overtly expressed parts of the sentence. This has the consequence that the complementizer in an embedded yes/no question like (10) may not be deleted: the Q-feature of C is not recoverable. In (16a), however, the Q-feature is encoded on the WH-expression as well as on the complementizer, and hence this property of the complementizer is recoverable after deletion:

(17) a.   John wonders if Mary loves him --->
      a'. *John wonders __ Mary loves him
      b.   John wonders what if Mary will say --->
      b'. John wonders what __ Mary will say

Concerning (16b), it can be observed that, like the null head in (16a), the null Q-operator is phonetically realized in certain languages. Some examples are given below.

(18) a. **An** bpósfaidh tú mé   (Irish)
          Q *will-marry you me*
      b. **Is** idda hmad s tmazirt   (Berber)
          Q *went Ahmed to country*
      c. **Tsi** hot er geleient dos bux   (Yiddish)
          Q *has he read the book*

In addition to this, it might be that X-bar theory necessitates the presence of such an operator. Consider how a matrix yes/no question is derived. The sentence starts out as a VP (cf. 19a). Then the verb undergoes self-attachment, as in (19b). However, (19b) as it stands may be ungrammatical, because it is structurally ambiguous (see note 4). It can either be analyzed as an adjunction of a (minimal maximal) element to VP or a case of the shifted verb taking a VP-complement. Ambiguity of this type is resolved by VP-to-V' reanalysis if a phrasal element is adjoined to the top of the structure, as outlined in section 2.1. In (16b), this would be the (base-generated) null Q-operator:

(19) a.   [$_{VP}$ You have seen the soccer match]
      b.   [$_{VP}$ Have$_i$ [$_{VP}$ you t$_i$ seen the soccer match]]
      c.   [$_{VP}$ Q [$_{VP}$ have$_i$ [$_{VP}$ you t$_i$ seen the soccer match]]
      d.   [$_{VP}$ Q [$_{V'}$ have$_i$ [$_{VP}$ you t$_i$ seen the soccer match]]

Just as in the case of *if* deletion, one could argue that English does have a Q-operator in its lexicon, but that this element is not spelled out when the syntax is mapped onto a PF representation. This might follow from the constraints proposed by Pesetsky (1994). In particular, Left Edge (F), the constraint that requires left alignment of a functional projection with its head, would favor deletion of material preceding the finite verb in (19), up to recoverability. The Q-operator is in fact recoverable after its Q-feature has been copied onto the verb.

We will not pursue these issues here, assuming that the solutions suggested above, or something like them, are correct. In conclusion, Q-marking can be taken to be an instantiation of the general process of sentence-type marking, as outlined in section 2.1. We will now consider what consequences this has for the topic central in this paper: the syntax of multiple questions.

2.3 *Multiple Questions*

In this section we will begin our argument to the effect that the syntax of multiple questions is determined by constraint interaction in an optimality-theoretic way. First, we will formulate two constraints that follow from the theory of Q-marking introduced above. We will then show that interaction of these constraints with the well-known Shortest Steps constraint can account for the syntax of multiple questions in English.

The first constraint that follows from section 2.2 is obvious: a question must be marked as such, where marking requires VP to be the complement of the Q-marker. We will refer to this constraint as *Q-Marking*.

12

(20)   *Q-Marking*
        A question must be overtly Q-marked

We should point out here that Q-Marking (and the other two constraints we will introduce in this section) is a constraint holding of surface structures. That is to say, it requires that questions are visibly marked, thus forcing, as we will see below, overt movement of both one WH-expression and the verb (in main clauses). One can think of constraints of this type in two ways: either as S-structure constraints (re-introducing this level of representation) or as PF constraints (stretching the subject matter of this level of representation). We will adopt the former option here.

        Q-Marking is a constraint about the element to be marked. From the theory, a plausible constraint about the marking element also follows. Due to the general notion of economy, it is desirable not to insert features into a tree without these features having a suitable function. If this is so, insertion of Q-bearing expressions requires usage of their Q-feature. We have already seen that 'Q' is a property added to propositions. Hence, the Q-feature must have a proposition in its scope. Since propositions are expressed by VPs (given the VP-internal subject hypothesis), the Q-feature must take scope over VP. The constraint we would like to propose requires that this scope be marked overtly.

(21)   *Q-Scope*
        [+Q] elements must c-command VP at surface structure

It will be clear that there is some overlap in the empirical effects of Q-Marking and Q-Scope. Both principles may trigger WH-movement. However, as we will show at length below, there are constructions in which one is crucially satisfied, while the other is not. For example, movement of a WH-expression without accompanying head movement will result in a structure that violates Q-Marking, but satisfies Q-Scope (for this WH-expression).

        The third and last constraint that will be relevant needs no further introduction:

(22)   *Shortest Steps*
        Minimize the distance between chain links

Since zero is the minimal distance imaginable, (22) subsumes the 'Move-only-when-necessary' constraint that turns up in various forms in various works. It is an open issue how distance in chains should be measured. The most straightforward interpretation seems to be that each node in the path between two chain links results in a violation of Shortest Steps (cf. Collins 1994 for a similar approach). Length of chains is defined in (23).

(23) *Length*

      The length of a chain is the (total) cardinality of the path(s) connecting the head of the chain and the foot of the chain, such that there are no paths connecting the head and the foot with a lower (total) cardinality.

We assume that, in calculating the cardinality of a path, more segments of a single category count only once. So, a path $<X, Y, Y, Z>$ counts as equally long as a path $<X, Y, Z>$. The notion of 'path' and the relation of 'connectedness' (which we assume to be symmetrical and transitive) are defined as in (24) and (25).

(24) *Path*

      A path is an ordered set of nodes $<N_1\ N_2\ N_3\ ...\ N_M>$ such that, if N_N and N_{N-1} are contained in this set, N_N immediately dominates N_{N-1}.

(25) *Connectedness*

    a.   A node N is connected to a path P if and only if there is a node in P that immediately dominates N.

    b.   A path P is connected to a path Q if and only if there is a node N that is contained in both P and Q.

Consider, for instance, the case of simple upward movement depicted in (26) (which might be a case of NP raising in a VP containing an adjunct).

(26)

        B
       ／＼
    A_i C
        ／＼
    D   C
        ／＼
     E t_i
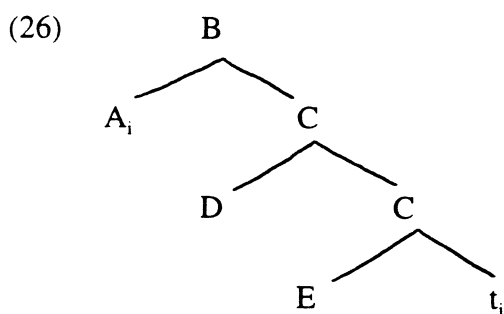
There is a single path connecting A and its trace, namely $<B, C, C>$. This set of nodes is a wellformed path because B immediately dominates the higher segment of C, while this segment immediately dominates the lower segment of C. The antecedent A is connected to this path since it is immediately dominated by B; the trace t is connected to this path since it is immediately dominated by (the lower segment of) C. The length of the chain between A and t is two, because the two segments of the single category C count only once. Hence, this movement results in two Shortest Steps violations.

      Next, consider the slightly more complicated case in (27) (which might depict head-to-head adjunction).

14

(27)
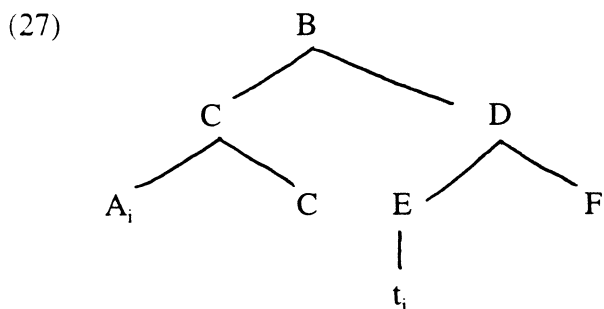
In this structure, there is no single path connecting A and t, because there is no set of nodes between A and t such that each of these nodes immediately dominates the next. Instead, A and t are connected by two paths: P_1 <B, C> and P_2 <B, D, E>. A is connected to P_1, t is connected to P_2, and P_1 and P_2 are connected because they both contain B. Hence, the length of the chain between A and t is five.

After having introduced the relevant constraints, we need to define the candidate set. Following work by Grimshaw (1993) and Reinhart (1993), we assume that sentences belong to the same candidate set if they are projected from the same set of lexical items (the same numeration) and target the same semantic representation. The intuition is that the syntax is a device mapping the lexicon to a semantic representation (and of course to a phonological representation). This mapping can take place in various ways, yielding various derivations and thus various candidates.[7] The actual derivation is the one that is optimal with respect to a number of ranked (S-structure and LF) constraints. This is schematized in (28), where the optimal route is bold-faced.

(28)   SS   LF
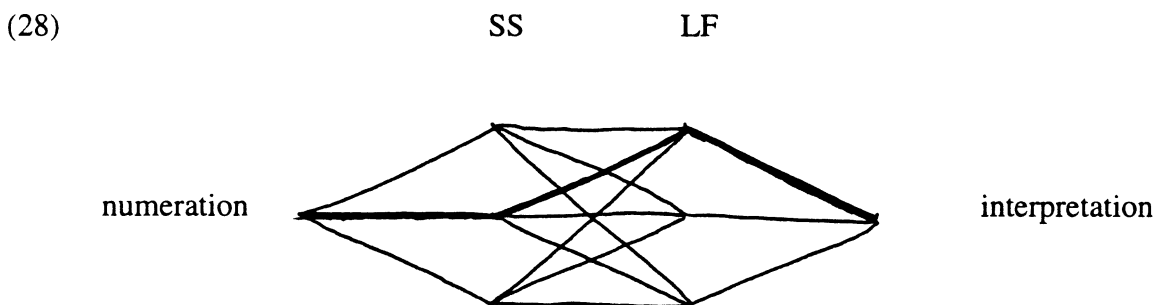
This view has several consequences. Trivially, two sentences that are projected from different lexical items are not part of the same candidate set. Two sentences that target different interpretations are not in competition either. Finally, sentences that are uninterpretable cannot block derivations that do lead to an interpretation, because such sentences belong to a different candidate set (if they belong to a candidate set at all).

---

[7] The device producing the various possible derivations (what is called GEN in optimality theory) might have certain intrinsic properties which (since they are part of the generator) cannot be violated, like the wellformedness conditions imposed by X-bar theory.

Consider now how the constraints interact in a nonmultiple question like *what have you seen*. In English, any question is Q-marked, so Q-Marking must be relatively high in the hierarchy. In particular, it must be ranked above Shortest Steps, or we would not expect WH-movement. This is shown in the tableau in (29), where Q-Scope is ordered lowest for reasons to be discussed below.[8]

(29)   *English simple questions*

	Q-Marking	Shortest Steps	Q-Scope
☞ What have [you t seen t]		*******	
[You have seen what]	*!		*
What [you have seen t]	*!	**	
Have [you t seen what]	*!	***	*

The sentence *what have you seen* violates Shortest Steps seven times: three times because the number of nodes separating *have* and its trace is three (the functional V', VP and V'), and four times because the path connecting *what* to its trace consists of <VP, V', VP, V'> (cf. 13d). However, Q-Marking is satisfied. Due to spec-head agreement in the derived functional projection, *have* acquires a Q-feature with which it marks its VP-complement. Crucially, all other candidates violate Q-marking. For the second one, this is immediately

---

[8] A complication seems to arise in embedded clauses. Consider the following examples:

(i)   a.   *I wonder if John loves who
        b.   *I wonder whether John loves who
        c.   I wonder who if John loves

The problem is that, because Q-marking is possible without WH-movement in the above examples, one would expect the WH to remain in situ, given that Shortest Steps outranks Q-Scope in English. This problem is only apparent, however. As noted earlier, Q-marking a VP without there being an accessible variable contained in this VP leads to a yes/no question interpretation. We assume that the interpretation of the Q-feature in CP is determined when Q-marking takes place, i.e. at surface structure in English. A consequence of this is that (ia) and (ib) are not in the same candidate set as (ic), and hence they cannot block (ic). The same reasoning applies to a main clause structure like (ii):

(ii)   *Q have$_i$ [you t$_i$ seen what]

Consequently, examples like (ia,b) and (ii) can only be interpreted as yes/no questions, but in that case the WH in situ must be interpreted as an echo-WH. (In this reading, the sentences are indeed possible).

16

clear, since nothing has been moved. In the third one, *what* has been adjoined to VP (thus complying with Q-Scope). An adjunction structure, however, is not a proper structure for marking, only a head-complement structure is, and no head has been fronted. In the last candidate, a head has been fronted, but this does not inherit a Q-feature, due to lack of movement of *what*. Note that it follows that there will be no language in which verb movement is triggered in WH-sentences (as it is in English), but in which the WH-element itself remains in situ (as in Chinese). Since moving the head without moving the WH-element does not lead to any improvement with respect to either Q-Marking or Q-Scope, Shortest Steps will rule out this possibility regardless of the ranking of these constraints. As far as we know, this prediction is correct.

Let us now turn to multiple questions. The high ranking of Q-Marking again ensures that the head and at least one WH-phrase must move in order to create the proper Q-marking environment. The question now is what will happen to the other WH-phrase(s). Here, the ranking between Shortest Steps and Q-Scope becomes relevant. If Shortest Steps is ranked higher than Q-Scope, as in the tableau in (30) below, it follows that the other WH-phrases remain in situ. Under the reverse order, one would expect all WH-elements to move out of VP.

(30) *English multiple questions*

	Q-Marking	Shortest Steps	Q-Scope
☞ Who has [t t seen what]		******	*
What has [who t seen t]		*******!	*
Who what has [t t seen t]		*******!***[9]	
Has [who t seen what]	*!	***	**

Although Shortest Steps is violable in English (the actual outputs in both (29) and (30) violate it), it has its effects. It does not only account for the fact that all WH-elements but one remain in situ, but also for the fact that the one WH-element that *is* moved (to comply with Q-Marking) is the one which makes the shortest possible move, i.e. the subject in (30) and the direct object in *what did Mary give to whom* (cf. Barrs & Lasnik 1986). In general, superiority effects follow from the global evaluation of Shortest Steps, as argued

---

[9] The way in which this number of violations is derived will be clearer in section 3.1, when Bulgarian multiple questions are discussed. It should be clear, however, that moving *what* is worse than leaving it in situ as far as Shortest Steps is concerned.

by Golan (1993) and Reinhart (1993) (a very similar analysis was already proposed by De Haan (1979:157 ff.), based on his Minimal String Principle).

Concerning the WH in situ, we assume, following Reinhart (1993) and others, that this element does not move at LF either. It is interpreted via a semantic procedure we need not go into here. One of the advantages of this assumption is that the null hypothesis can be adopted that all movement is subject to Subjacency, not just overt movement. The well-known fact that WH-in-situ can apparently violate Subjacency follows, since it is not interpreted by movement. (We will come back to this issue in section 3.3 on Chinese/Japanese).

So, constraint ranking is crucial for English. As noted in section 1, however, the most important type of evidence for constraint ranking and global evaluation lies in cross-linguistic variation. This variation should be explicable just by reranking the constraints. We will now try to establish whether the typology of multiple question formation can be accounted for in this way.

3. *Reranking*

There are six possible rankings of the three constraints we have proposed. It will turn out that in two instances two orderings have the same effects. This means that in practice four different types of question formation are to be expected. One of these was discussed in section 2.3, the English type. The remaining three are also attested, as we will show in sections 3.1 - 3.3. In section 3.4, we will discuss a possibility not yet mentioned, which is that two constraints are not ranked with respect to each other. This will account for optional WH-movement in French.

3.1 *Bulgarian*

Suppose the constraints proposed in section 2 are ranked as in (31).

(31)   Q-Marking >> Q-Scope >> Shortest Steps

The high ranking of Q-Marking ensures that, just like in English, the proper structure for Q-marking must be derived. That is, it ensures head movement plus movement of at least one WH-phrase to the specifier of the newly created VP (CP in traditional theories; for ease of reference we will henceforth refer to functional VPs as VP*). Compared to English, however, the ranking of Q-Scope and Shortest Steps has been reversed. This means that not just one, but all WH-phrases must move out of VP. The shortest possible

18

move is adjunction to VP, but this is ruled out for the following reason. The structure that would result if the other WHs would adjoin to VP is the one in (32).

(32)

```
            VP*
          /     \
        WH_i     V'*
               /     \
             V_j*      VP
                     /    \
                   WH_k     VP
                          /    \
                        t_i     V'
                              /    \
                             V      t_k
                             |
                             t_j
```
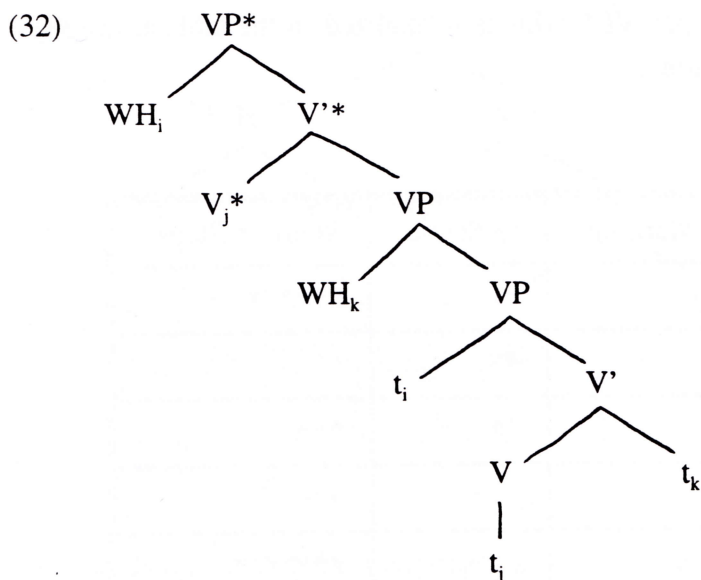
In this structure, Q-Marking is violated after all, because of a Relativized Minimality effect. Relations between an element of type A and another element are generally blocked by intervening elements of type A. This phenomenon is well-known from the literature on extraction, but it can also be observed in other domains. For example, a verb may select for the case of its complement, but not if another case-assigning head intervenes, such as a preposition. If Relativized Minimality holds of syntactic relations in general, Q-marking of the VP by the derived functional head in (32) is blocked by the Q-bearing WH_k, which is a potential Q-marker. Since WH_k is not in the proper configuration for marking, there is no Q-marking possible at all in (32).[10]

---

[10] Probably, even if the intervening element would not be a potential Q-marker, marking by V* would be blocked anyway in a structure like (32). This is because presumably adjunction to direct complements (that are to be marked) is generally ruled out. Adjunction to complement clauses, for instance, is ungrammatical (cf. Chomsky 1986, Grimshaw 1993, and others). In general, marking is possible in a configuration like (ia), but not in a configuration like (ib)

(i)   a.

```
        X'
      /    \
     X      YP
```

b.

```
        X'
      /    \
     X      YP
          /    \
        ZP_i    YP
              /    \
             Y      t_i
```

19
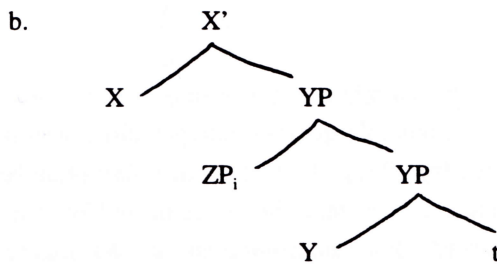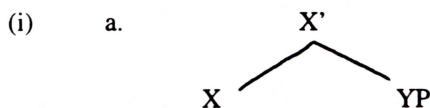
As a consequence, WHs may not adjoin to VP if Q-marking is necessary. If they must be moved out of VP, they must move to a position above V*. Hence, not just one, but all WHs must move to spec-VP*.[11] In short, the effect of the ordering in (31) is that all WH-expressions cluster together in spec-VP*. This is formalized in the tableau in (33) (where the brackets indicate VP-boundaries).

(33) *Bulgarian multiple questions*

	Q-Marking	Q-Scope	Shortest Steps
WH V [t t WH]		*!	******
[WH V WH]	*!	**	
V [WH t WH]	*!	**	***
[WH [WH [t V t]]]	*!		***
WH V [WH [t t t]]	*!		********
☞ WH WH V [t t t]			*********

---

[11] In principle, there still seems to be another possibility, namely moving one WH to spec-VP* to ensure that there will be Q-marking, and adjoining all other WHs to VP*, as in (i).

(i) * VP*

WH$_i$

VP*

WH$_j$

V'*

V$_k$*

VP

t$_i$

V'

V

t$_j$

|

t$_k$
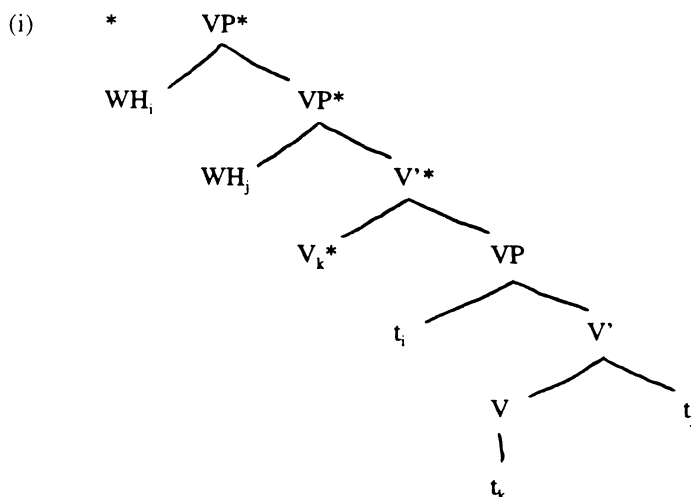
In order to obtain a multiple question interpretation, absorption must take place (cf. May 1977, 1985, Higginbotham & May 1981). We assume that absorption between moved WHs is possible under the standard condition of mutual m-command. Since the moved WHs in (i) do not mutually m-command each other (WH$_j$ is dominated by VP*, which does not dominate WH$_i$), absorption is impossible. This structure is therefore uninterpretable, so it falls outside the relevant candidate set (see section 2.3).

It is predicted, therefore, that there are languages in which multiple questions are structured as in (34).

(34)

```
                          VP*
                 _____/  _____
              WH_i                    V'*
            __/  \__            ____/    \____
         WH_i      WH_j     V_k*              VP
                            t_i        ____/    \____
                                   V'
                               __/    \__
                              V            t_j
                              |
                              t_k
```
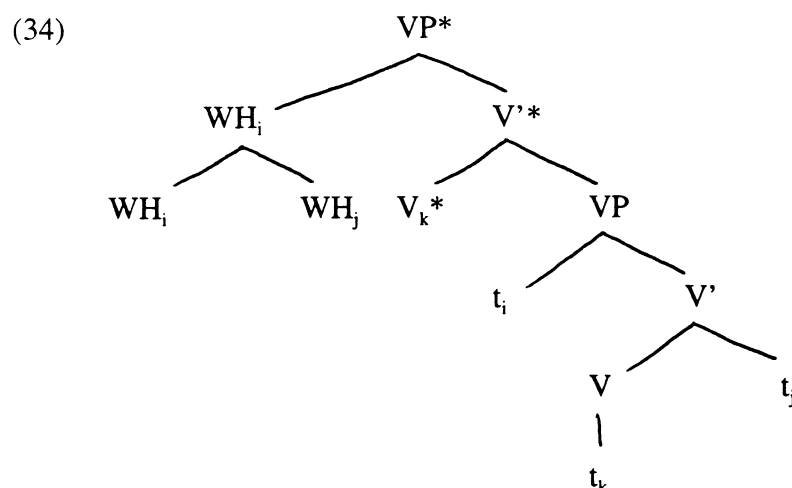
This prediction is correct. Languages in question are, for example, Bulgarian and Romanian, as shown by Rudin (1988). Rudin cites several pieces of evidence showing that all WH-expressions form an indivisible constituent in spec-CP (here: spec-VP*). Focusing on Bulgarian, it can be observed that no WH-expression remains in situ (all examples from Bulgarian are taken over from Rudin 1988):[12]

(35)   Koj kogo vižda
       who whom sees

Moreover, the WH-cluster cannot be interrupted. For instance, second place clitics in Bulgarian must come at the end of the entire WH-sequence, indicating that this sequence occupies a single syntactic position. This is shown for the pronominal clitic *ti* and the auxiliary clitic *e* in (36).

(36)   a.   Koj kakvo ti e kazal
            who what you has told
            'who told you what'
       b.   *Koj ti e kakvo kazal
            who you has what told

Similarly, adverbials may not split up the fronted WH-sequence, as shown by the examples in (37).

---

[12] For reasons of exposition, we will restrict our attention to cases with two WH-phrases (subject and object) here, but the results can immediately be extended to cases with more than two WH-phrases.

(37) a.   Zavisi ot tova, koj kogo prův e udaril
          depends on this who whom first has hit
          'It depends on who hit whom first'

      b.   *Zavisi ot tova, koj prův kogo e udaril
          depends on this who first whom has hit

Another piece of evidence cited by Rudin comes from WH-words in free relatives (assuming that the syntax of relatives is parasitic on the syntax of questions, at least in these languages). Relative WH-words must be marked with a definitizing *-to*. It is possible that all WH-expressions are affixed with *-to* separately, but it suffices to add one *-to* at the end of the complete WH-sequence to yield the whole sequence relative. Adding *-to* to the first WH-constituent only is ungrammatical, however. As Rudin notes, this suggests that *-to* is suffixed to one WH-constituent, which consists of all fronted WH-phrases:

(38) a.   Kojto kakvoto iska ...
          who-DEF what-DEF wants
          'whoever wants whatever'

      b.   Koj kakvoto iska ...
          who what-DEF wants

      c.   *Kojto kakvo iska ...
          who-DEF what wants

Another important ingredient of the structure in (34) is that there is obligatory verb movement in order to derive the VP* needed for Q-marking. This means that, if the subject is not a WH-element itself, we predict there to be obligatory inversion in WH-questions. This is correct. As noted by Kraskow (1992), languages of the Bulgarian-type do indeed display obligatory inversion in questions:

(39) a.   Kakvo kupuva Ivan
          what buys John
          'What does John buy'

      b.   *Kakvo Ivan kupuva
          what John buys

It seems, then, that the syntax of Bulgarian WH-questions indeed follows from the reranking in (31) of the constraints proposed in section 2.3.

    A further striking characteristic of the WH-sequence in Bulgarian resembles the superiority effects that can be observed in English. While in multiple WH-questions in English a WH-subject will be the element that is overtly moved, in Bulgarian a WH-subject must precede all other fronted WH-elements (cf. 40).
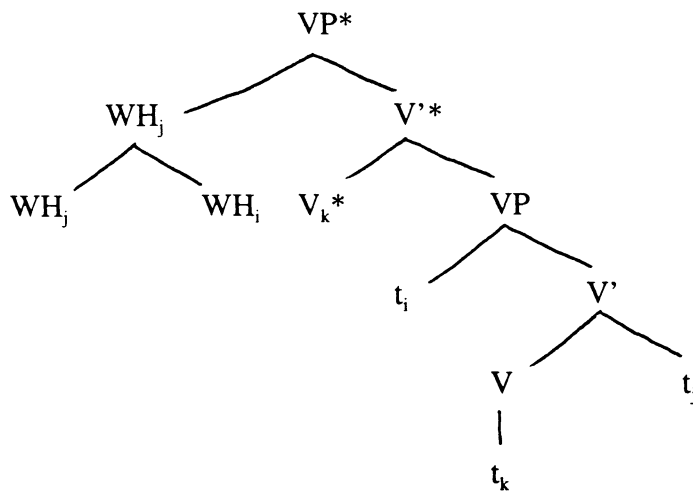
(40) a. Koj kogo vižda
          who whom sees
      b. *Kogo koj vižda
          whom who sees

We will now argue that this is an indirect effect of the lowest ranked constraint in Bulgarian, Shortest Steps. This is an illustration, therefore, of the fact that low-ranked constraints can still have their effects when the different candidates are equal with respect to all the higher constraints. Note that (40) indicates that the relation between subject and object in the WH-cluster is asymmetric. Either the subject must be adjoined to the object, in which case (40) shows that WH-phrases must be left-adjoined, or the object must be adjoined to the subject, in which case (40) shows that WH-phrases must be right-adjoined. It cannot be the case, however, that it does not matter which WH-phrase adjoins to which, since free word order would be predicted then, even when assuming a uniform direction of adjunction. According to Rudin, the cluster in spec-VP* is headed by the subject, with the other WH-phrases right-adjoined to it. We will now show that this indeed follows from Shortest Steps in our analysis.

Suppose that both WH-phrases move to spec-VP* independently. No matter whether the subject is moved first, and the object adjoined to it (cf. 41a), or the other way around (41b), Shortest Steps will be violated equally, namely twelve times.
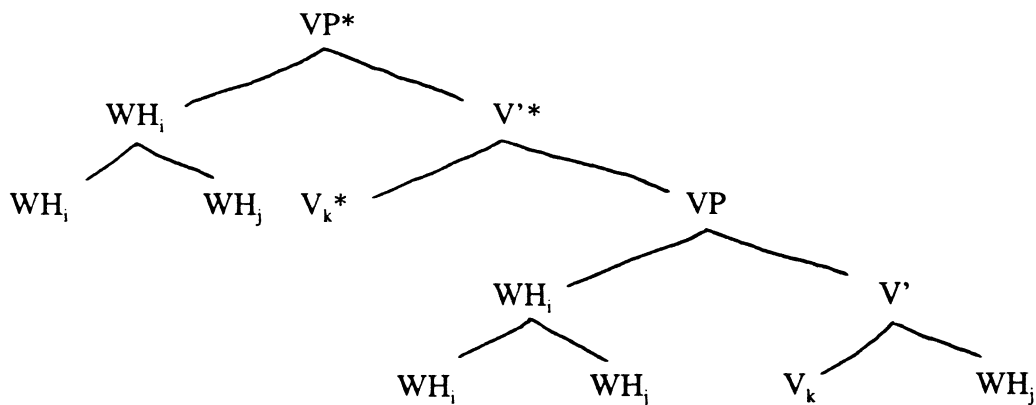
(41) a.

b.

```
                              VP*
                 _____/_____
              WH_j                              V'*
           ___/\___                       _____/_____
        WH_j       WH_i      V_k*        VP
                                    ____/\____
                                  t_i          V'
                                          _____/\_____
                                        V              t_j
                                        |
                                        t_k
```
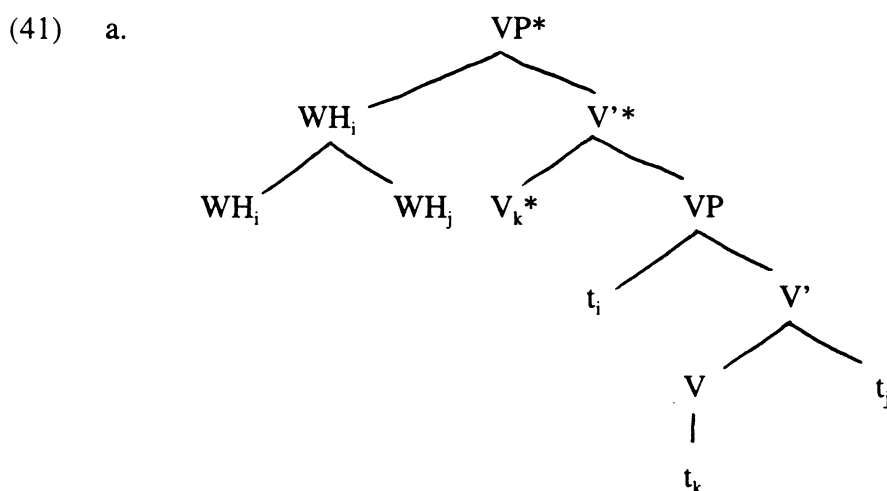
The paths that connect antecedents and traces in (41a) are <V'*, VP, V'> for the moved verb, <VP*, V'*, VP> for the moved subject, and the two paths <VP*, WH_i> and <VP*, V'*, VP, V'> for the moved object. This amounts to twelve violations of Shortest Steps. In (41b) the paths are <V'*, VP, V'> for the verb, <VP*, V'*, VP, V'> for the object, and <VP*, WH_j> plus <VP*, V'*, VP> for the subject, which also results in twelve Shortest Steps violations. If these were the only options, we would therefore predict optionality as to which WH-phrase adjoins to which, while we just noted that this cannot be correct.

Fortunately, there is another option to derive the correct structure for a multiple WH-question in Bulgarian, which has the object adjoined to the subject and which turns out to be scoring better on Shortest Steps. This option consists of first adjoining the WH-object to the WH-subject and then moving this cluster as a whole to spec-VP*. The point is that moving the two WH-phrases together, as one constituent, is cheaper than moving them apart over the same distance. If the object can be taken for a ride by the subject for part of the distance it must move, this will therefore be cheaper. See the tree in (42), where for clarity we have spelled out the internal structure of the traces (we assume the copy theory of movement, cf. Chomsky 1993).

(42)

```
                              VP*
                 _____/_____
              WH_i                              V'*
           ___/\___                     _____/_____
        WH_i       WH_j     V_k*       VP
                                  _____/\_____
                                WH_i            V'
                              ___/\___       ___/\___
                            WH_i    WH_j   V_k      WH_j
```

24

The paths we now obtain are the following: <V'*, VP, V'> for the moved verb again, <VP, WH$_i$> and <VP, V'> for the moved object, and <VP*, V'*, VP> for the moved WH-cluster. This results in only ten violations of Shortest Steps, and therefore is better than either of the trees in (41). Note that this derivation crucially depends on adjunction of the object to the subject, thus explaining the asymmetry in the WH-cluster.[13] The correct word order then follows from uniform right-adjunction of WH-phrases, as proposed by Rudin.[14]

To conclude the section, it should be pointed out that the constraint ranking in (31) is not the only one which results in the Bulgarian-type of multiple WH-questions. Reversing Q-Scope and Q-Marking has no effects, as long as Shortest Steps is ranked lowest. Consider the tableau in (43). Compared to (33), the column for Q-Scope and Q-Marking are reversed.

(43)   *Bulgarian multiple questions (alternative ranking)*

	Q-Scope	Q-Marking	Shortest Steps
WH V [t t WH]	*!		******
[WH V WH]	*!*	*	
V [WH t WH]	*!*	*	***
[WH [WH [t V t]]]		*!	***
WH V [WH [t t t]		*!	*******
☞ WH WH V [t t t]			*********

Q-Scope dictates that all WHs must move out of VP. The derivation that complies best with Shortest Steps would involve adjunction of the WHs to VP, without head movement. But since Q-Marking outranks Shortest Steps, the syntactic configuration for Q-marking must be derived. This can only be achieved by head movement plus WH-movement of at least one WH to the spec of the derived VP*. For reasons outlined above, the other WHs must then move to spec-VP* as well (if they have to move out of VP in the first place, as

---

[13] Note that first adjoining the subject to the object and then moving the whole cluster to spec-VP* is ruled out by the ban on downward movement, even though it also results in ten Shortest Steps violations at S-structure.

[14] There are in fact some cases where left-adjunction is possible as well, i.e. cases where the order object-WH before subject-WH is possible, dependent on pragmatic factors. See Billings & Rudin (1995) for discussion of this issue within an OT framework.

is the case because of the high ranking of Q-Scope), or Q-marking would be blocked after all. Hence, the same pattern reappears.

From the six possible rankings of our three constraints, three have now been discussed, leading to two different patterns of WH-question formation. The ranking Q-Marking >> Shortest Steps >> Q-Scope leads to movement of the verb and one WH, as in English. The rankings Q-Marking >> Q-Scope >> Shortest Steps and Q-Scope >> Q-Marking >> Shortest Steps trigger clustering of WHs in spec-VP*/CP and supporting verb movement in main clauses, as in Bulgarian. In the next section we will discuss a third pattern, derived from yet another ranking.

3.2 *Czech*

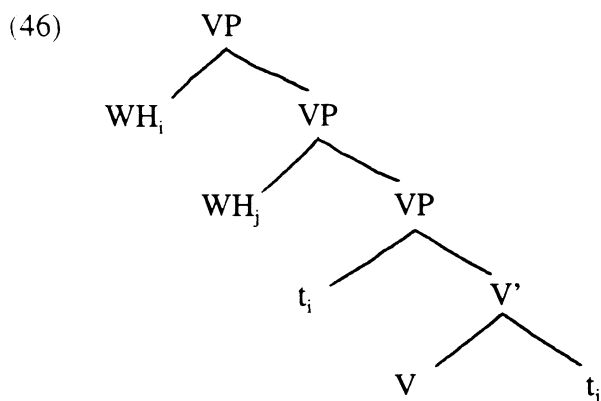Suppose the constraints proposed in section 2.3 are ranked as in (44).

(44)   Q-Scope >> Shortest Steps >> Q-Marking

The high ranking of Q-Scope dictates that all WH-phrases be moved out of VP, like in Bulgarian. Unlike in Bulgarian, Shortest Steps outranks Q-Marking, so that it is more important that movement be minimized than it is to derive the proper Q-marking structure. As noted in section 3.1, the derivation that complies with the demand that all WHs take scope over VP which is best with respect to Shortest Steps is one in which all WHs are adjoined to VP and in which there is no head movement (and hence no derived VP* that would be necessary for Q-marking). This is illustrated by the tableau in (45).

(45)   *Czech multiple questions*

	Q-Scope	Shortest Steps	Q-Marking
WH V [t t WH]	*!	******	
[WH V WH]	*!*		*
V [WH t WH]	*!*	***	*
☞ [WH [WH [t V t]]]		***	*
WH V [WH [t t t]]		****!****	*
WH WH V [t t t]		****!******	

The optimal candidate is structured as in (46).

26

(46)

```
          VP
         /  \
      WH_i   VP
            /  \
         WH_j   VP
               /  \
              t_i   V'
                   /  \
                  V    t_j
```

This representation violates Shortest Steps only three times. The path between the moved subject-WH and its trace consists of VP-segments only. As proposed in section 2.3, crossing segments of the same category does not lead to more than one violation. The path between the object-WH and its trace consists of two VP-segments and V', and hence induces two violations.

In section 3.1 it turned out that it was cheaper with respect to Shortest Steps to first adjoin the object-WH to the subject-WH, and then move this cluster out of VP. One might expect that a derivation like this is cheaper as well in the grammar under discussion, where WH-movement is adjunction to VP. It turns out, however, that in this case clustering the WHs and then moving them together is more costly. Such a derivation would result in the following tree (where the traces are spelled out).

(47)

```
                        VP
                      /    \
                   WH_i      VP
                  /   \     /    \
               WH_i   WH_j  WH_i   V'
                           /   \   /   \
                        WH_i  WH_j V   WH_j
```
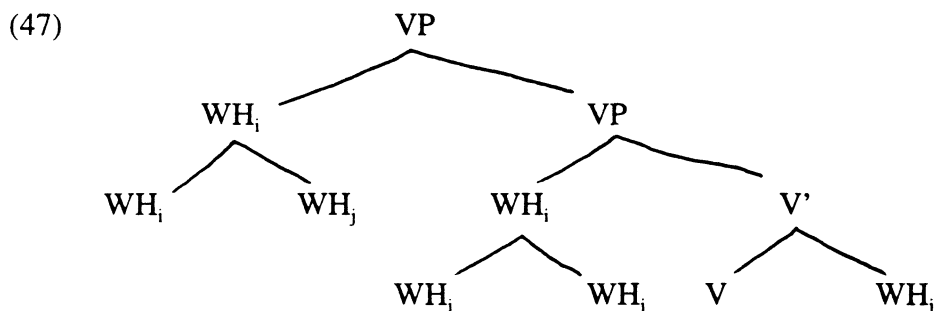
The path for the moved WH-cluster is <VP, VP>, while the paths relevant to the moved object are <VP, WH_i> and <VP, V'>. This amounts to five Shortest Steps violations, two more than in the optimal candidate in (46).[15]

---

[15] Clustering is also suboptimal where the indirect object and the direct object of a double object construction are concerned. We assume the structure in (i) for these constructions. Adjoining DO- and IO-WHs separately to VP gives five Shortest Steps violations. First adjoining the DO-WH to the IO-WH, and then adjoining this cluster to VP gives six Shortest Steps violations.

It is predicted therefore that the ranking in (44) defines a grammar in which all WHs are fronted, but in which they do not form a cluster. There are indeed languages of this type. Examples are Czech and Polish. Focusing on Czech, all WH-expressions must be fronted in this language, just like in Bulgarian.

(48)   Kdo koho viděl

      who whom saw

      'who saw whom'

Unlike in Bulgarian, however, elements can appear between the fronted WH-expressions. Consider for example second place clitics, which in Czech must always follow the first major constituent of the clause, as shown by Toman (1986) and Rudin (1988). These authors further show that these clitics always follow the first WH-word in multiple WH-constructions. This indicates that the WH-expressions form separate constituents, not one:

(49)   a.   Kdo ho kde viděl je nejasné

           who him$_{CL}$ where saw is unclear

           'it is unclear who saw him where'

      b.   *Kdo kde ho viděl je nejasné

           who where him$_{CL}$ saw is unclear

Moreover, parentheticals can appear anywhere in the WH-sequence in Czech. This is shown below for an example containing three WH-expressions (Rudin's 49):

(50)   a.   Kdo, podle tebe, co komu dal

           who according to you what to whom gave

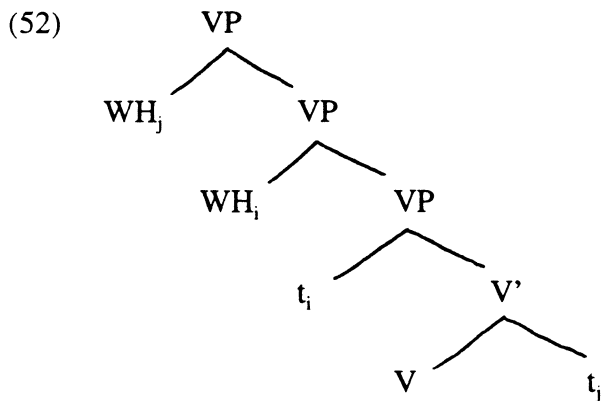   'who, according to you, gave what to whom'

(i)
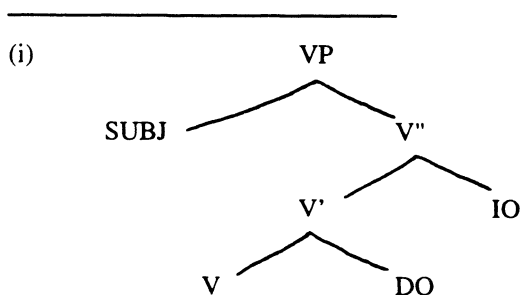
This analysis of double object constructions comes close to the traditional analyses in Blom & Daalder (1977) and De Haan (1979) (where our V' is VP, our V" is PredP, and our VP is S). We assume (i) to be valid for English as well as a basic structure. The surface structure is derived by rightward movement of the DO, which takes place for reasons outlined in Neeleman (1994).

28

b.   Kdo co, podle tebe, komu dal
who what according to you to whom gave

The same phenomenon occurs with adverbials. (All Czech data below were provided by Ludmila Menert, p.c.).

(51)   a.   Kdo rychle co komu dal
who quickly what to whom gave
'who quickly gave what to whom'

b.   Kdo co rychle komu dal
who what quickly to whom gave

From the proposed ranking for Czech it does not only follow that the WH-words do not form a cluster, but also that no superiority effects should exist (unlike in Bulgarian again). Whereas in the Bulgarian WH-cluster the subject-WH precedes the other WHs (cf section 3.1), it is predicted that in a grammar defined by (44) the order between the fronted WHs is free. This is because the number of Shortest Steps violations will always be the same when all WH-words adjoin separately to VP, regardless of the order in which they are moved. Compare, for instance the tree in (52) with the one in (46). In (46) the object was adjoined first, after which the subject was adjoined. In (52) the order of adjunction is reversed.

(52)

$$
\begin{array}{c}
\text{VP} \\
\diagup\!\!\!\diagdown \\
\text{WH}_j \qquad \text{VP} \\
\qquad \diagup\!\!\!\diagdown \\
\qquad \text{WH}_i \qquad \text{VP} \\
\qquad\qquad \diagup\!\!\!\diagdown \\
\qquad\qquad t_i \qquad \text{V'} \\
\qquad\qquad\qquad \diagup\!\!\!\diagdown \\
\qquad\qquad\qquad \text{V} \qquad t_j
\end{array}
$$

Because segments count only once, there is no difference in the number of Shortest Steps violations of the paths connecting the subject and the object to their traces. The paths are <VP, VP> for the subject and <VP, VP, VP, V'> for the object, again giving three violations in total. We thus predict that in Czech-type languages there is free word order between the WHs (modulo possible stylistic preferences). In particular, it should be the case that the subject-WH does not have to precede the other WHs. This prediction is borne out. Alongside (48), (53) exists.

(53)   Koho kdo viděl

whom who saw

'who saw whom'

Given the trees in (46/52) a further prediction follows. As opposed to the Bulgarian-type, there should be no verb movement in the Czech-type languages (since the proper structure for Q-marking need not be derived). It is predicted therefore that there should be no obligatory inversion in questions with a non-WH-subject. As noted by Kraskow (1992), this correlation indeed exists. Compare (54) with (39).[16,17]

(54)   a.   Co kupuje Ivan

what buys John

'What does John buy'

b.   Co Ivan kupuje

what John buys

Although our analysis of Czech multiple WH-questions correctly predicts that the fronted WHs do not form a cluster, that they are ordered freely, and that there is no verb movement, there is a final point to consider. This concerns the position of the first WH-element. Until now, we have assumed that, like the other WHs, it is adjoined to VP. According to Rudin (1988), however, the first WH-expression is moved to spec-CP. This can, in fact, not be verified for main clauses, because C (V* in our terms) is not present.

---

[16] Note that inversion is possible, but this need not be due to verb movement, since Czech has an Italian-type of free subject inversion.

(i)   a.   Včera Jan koupil knihu

Yesterday John has-bought book

'John has bought a book yesterday'

b.   Včera koupil knihu Jan

Yesterday has-bought book John

The same appears to be the case for the other languages of this type, like Serbo-Croatian (Maaike Schoorlemmer, p.c.).

[17] It also follows that there will be no languages in which one WH-expression is fronted without supporting verb movement while the other WHs remain in situ:

(i)   *WH$_i$ [$_{VP}$ ... t$_i$ ... WH ...]

The reason for this is that, in the absence of verb movement, WH-movement must be triggered by Q-Scope, and not by Q-Marking (since this constraint is violated in (i)). This means that Q-Scope would also trigger movement of the other WHs. As far as we know, this prediction is correct.

30

Obviously, there is no complementizer in main clauses, and as we have seen above there is also no verb movement. Therefore, given the Shortest Steps requirement, we take it that the initial WH, like the others, adjoins to VP in main clauses.

In embedded clauses the situation is different. Rudin (1988) claims that the doubly-filled-COMP filter is present in languages like Czech and Polish. However, we are informed that in colloquial registers it is possible to combine an overt complementizer with a fronted WH-expression in Czech (though not in Polish):

(55)   Chtěl bych vědět ...
       I would like to know
       a.   *že co Marie četla
             that what Mary read
       a'.   co že Marie četla
             what that Mary read
       b.   *že komu co Marie dala
             that who what Mary gave
       b'.   komu že co Marie dala
             who that what Mary gave
       b".   *komu co že Marie dala
   *who what that Mary gave*

These facts show that in embedded clauses the situation is as Rudin assumes: one WH-phrase has moved to spec-CP, while the others are again adjoined to VP (IP in Rudin's system). Apparently, in embedded clauses there is an extra factor forcing a longer movement of one of the WHs, in spite of Shortest Steps. This factor is tied to the presence of the complementizer *že*:

(56)   Myslím/vím/říkám *(že) Janek přijde
       I think/know/say that John comes

This complementizer marks its complement as a declarative. This implies that WH-expressions, at least at the level of interpretation, should take scope over this complementizer, or the clause would be uninterpretable.[18] However, as argued by Van

---

[18] In Czech embedded WH-questions, the complementizer that can appear is the declarative one, never the interrogative one. However, movement of at least one WH to spec-CP would be triggered just as well in the presence of an interrogative complementizer. This is because, as noted in section 2.2, a clause introduced by an interrogative complementizer and containing WH-phrases below this complementizer cannot be interpreted as a WH-question, only as an embedded yes-no question with an echo reading for the WH(s) (see also notes 6 and 8).
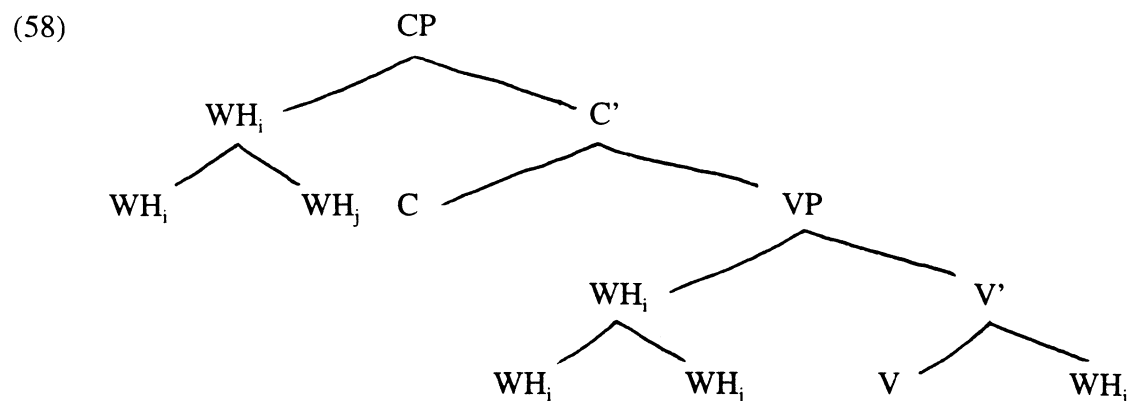
Riemsdijk (1978), a WH-expression that is A'-moved at surface structure cannot be moved at LF: its syntactic position is fixed. This implies that if all WHs would adjoin to VP in embedded questions (as they do in main clauses in Czech), a structure would result that is uninterpretable. So, in derivations leading to the desired interpretation the WHs must take scope over the complementizer already in overt syntax. In order to achieve this, it suffices to move one WH-expression to a position c-commanding the complementizer at surface structure, i.e. to spec-CP. This becomes apparent if we consider the resulting structure in (57). (The construction contains a CP as opposed to a VP*, since we are dealing with an embedded question. Hence, the functional projection is derived by insertion of a complementizer, rather than by movement of the verb).

(57)
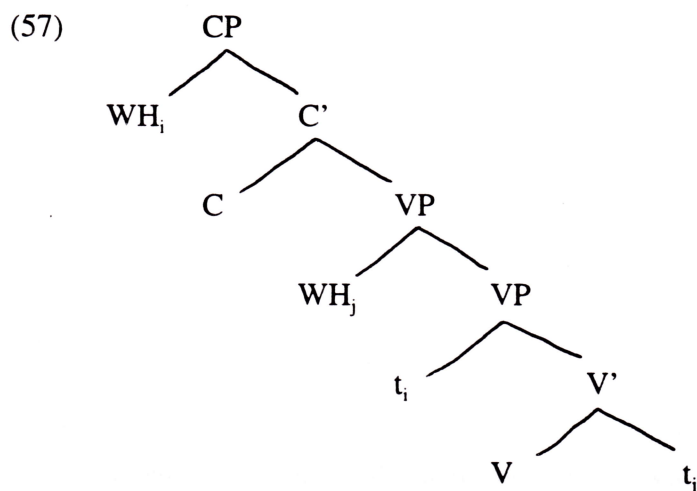
Absorption between the WHs must take place in order to obtain a multiple question interpretation. As noted in footnote 11, the condition on absorption between moved WHs is mutual m-command. The WHs do mutually m-command each other in (57), so absorption is possible. As a result, both WHs will have the same scope, namely that of WH_i (since, as noted, they must take scope over the complementizer). So, in order for interpretation to be successful one WH must move to spec-CP in order to provide an absorption site for the other WHs.

The other WHs will not move to spec-CP but adjoin to VP, because of Shortest Steps. In (57), Shortest Steps is violated five times (the path <CP, C' VP, VP> for the subject giving three violations, the path <VP, VP, V'> for the object giving two violations). The alternative derivation would involve formation of a WH-cluster within VP plus movement of this cluster to spec-CP (à la Bulgarian), as in (58), with traces spelled out. Here, Shortest Steps is violated seven times (the paths <VP, WH_i> plus <VP, V'> for

the object resulting in four violations, the path <CP, C', VP> for the WH-cluster resulting in three violations).[19]

(58)

```
                        CP
            WH_i                  C'
        WH_i    WH_j   C                  VP
                              WH_i              V'
                          WH_i   WH_j      V        WH_j
```

Note that in the optimal structure in (57), as opposed to the suboptimal structure in (58), Q-marking is blocked by adjunction of a WH to VP. However, Q-Marking is ordered below Shortest Steps in Czech, and is therefore overruled. Hence the contrast with Bulgarian, in which (58) is the optimal structure. In this language, Q-Marking is ranked higher than Shortest Steps (see section 3.1 for discussion).

Finally, it is predicted that, like in main sentences, no superiority effects will obtain in Czech embedded clauses, not even with respect to the movement of one WH-expression to spec-CP. This is because all WHs must be moved out of VP anyway. Then, for Shortest Steps it does not matter which WH-phrase is selected to move further to spec-CP. This prediction is correct: both (59a) and (59b) are possible.

---

[19] It would seem to be predicted that when the distance to be covered by the WH-expressions in a multiple question increases, clustering will be favored. This is not the case, however. The structure of multiple WH-questions in Czech remains the same when the WHs are moved out of an embedded clause to a position c-commanding the matrix-clause (insofar as this is possible in the first place, cf. Rudin 1988:455). In (i), for instance, no clustering seems to take place:

(i)   Co, podle tebe, komu Petr řekl [že Jan dal]
           what, according to you, who Peter said that John gave
           'What, according to you, did Peter say that John gave to whom'

This problem can be solved if we follow the suggestion of Fox (1995) that evaluation proceeds cyclically. That is to say, the ranked constraints first apply to the embedded CP, then to the combination of the embedded clause and the matrix clause. In the embedded CP, Shortest Steps favors separate movement of the two WH-expressions, as explained in the main text. This means that clustering can only take place when the larger cycle is taken into account, i.e. when the two WHs have already been adjoined to the embedded VP. However, then it is no longer possible, because it would have to take place within the embedded clause (the initial landing site of the WHs), which would be a violation of strict cyclicity.

33

(59)   Chtěl bych vědět ...
       I would like to know
       a.   kdo (že) co četl
            who that what read
       b.   co (že) kdo četl
            what that who read

In conclusion, like the first three, the fourth possible ranking of the constraints we have proposed results in an actually occurring type of question formation as well. In the next section we will show that the remaining two possible rankings lead to one more type, which is attested as well.

3.3 *Chinese and Japanese*

Suppose the constraints proposed in section 2.3 are ranked as in either (60a) or (60b), the two remaining possibilities.

(60)   a.   Shortest Steps >> Q-Marking >> Q-Scope
       b.   Shortest Steps >> Q-Scope >> Q-Marking

The high ranking of Shortest Steps ensures that it is more important not to move than it is to satisfy Q-Marking or Q-Scope. The result is that all WH-expressions will have to remain in situ in grammars defined by (60) (see the tableau in (61), in which the ranking between Q-Marking and Q-Scope is irrelevant).

(61)   *Chinese multiple questions*

	Shortest Steps	Q-Scope	Q-Marking
WH V [t t WH]	*!*****	*	
☞ [WH V WH]		**	*
V [WH t WH]	*!**	**	*
[WH [WH [t V t]]]	*!**		*
WH V [WH [t t t]]	*!*******		*
WH WH V [t t t]	*!********		

34

As is well-known, languages that adhere to the tableau in (61) exist. According to Huang (1982), Chinese is an example. In root questions as well as embedded questions, WH-expressions remain in their base position:

(62)   a.   Ni xihuan shei
               you like who
               'who do you like'
       b.   Wo xiang-zhidao Lisi mai-le sheme
               I wonder Lisi bought what
               'I wonder what Lisi bought'

At this point we need to be more precise about the interpretation of WH-expressions in languages like Chinese and Japanese. We have assumed earlier, following Reinhart (1993), that in English WHs in situ are not assigned scope by LF movement, but by an interpretational operation. One of Reinhart's arguments for this was that the interpretation of WH-in-situ is not subject to Subjacency. Although we cannot go into the details of the interpretational operation here, it is important to notice that in essence it makes use of the fact that a Q-operator obligatorily occupies spec-VP*/spec-CP in English WH-questions. Unmoved WH-phrases are parasitic on this Q-operator in a fashion reminiscent of unselective binding, so that for our present purposes the LF of a question like (63a) can be represented as (63b):[20]

(63)   a.   [Which persons]$_i$ did you ask [whether John read what]
       b.   Which$_{<i,j>}$ did you ask [t$_i$ persons] [whether John read what$_j$]

We would like to assume that the Q-operator on which WH-phrases in situ depend for their interpretation must be introduced by lexical material. That is to say, it is not an abstract operator that can be freely inserted (as in English yes/no questions), but it is introduced by a moved Q-bearing element (i.e. *which* persons in the example above). Returning now to languages like Chinese and Japanese, this assumption has the effect that one WH-expression must move by covert movement to a position c-commanding the proposition. After the movement of this one WH, possible other WH-expressions can be interpreted in situ through the semantic operation proposed by Reinhart. So, an S-structure like (64a) must be mapped into a representation like (64b) at LF:[21]

---

[20] Unlike absorption of moved WHs, this interpretational mechanism obviously does not require mutual m-command of the WH in situ and the WH in spec-VP*/CP, which determines the scope of the question.

[21] Note that LF movement is possible in Chinese and Japanese, but not in Czech embedded questions (see section 3.2), because, as already noted in the section on Czech, the fact that WHs move overtly in this language blocks further movement at LF.

35

(64) a. $[_{VP} \dots WH \dots WH \dots]$

  b. $[_{VP} WH_{<i,j>} [_{VP} \dots t_i \dots WH_j \dots]$

So, the prediction is that in languages with no overt movement one WH must move covertly, while the rest remains in situ.[22] Data presented by Watanabe (1992) indicate that this may well be the case. Watanabe shows that Subjacency effects can be observed in Japanese as well, despite the fact that all WHs are in situ at S-structure. Consider, for instance, the contrast in (65).

(65) a. ??John-wa [Mary-ga nani-o katta ka dooka] Tom-ni tazuneta no?
      John-TOP *Mary*-NOM *what*-ACC *bought whether Tom*-DAT *asked* Q
      'What did John ask Tom whether Mary bought'

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

Since the scope of the question is the matrix clause in these examples, a WH-expression must be moved to the top of the construction. The null hypothesis seems to be that Subjacency is a general constraint on movement, whether overt or covert.[23] If so, the ungrammaticality of (65a) can be explained if *what*, being the only Q-bearing element in the construction, must indeed be (covertly) moved. The grammaticality of (65b) can be explained as well: it follows from the presence of a WH-phrase in the matrix clause. This phrase can introduce the relevant Q-operator, so that *what* can remain in situ at LF and be interpreted in the way described by Reinhart.

---

[22] On apparent ECP effects with adjunct WHs in situ (or the lack thereof in certain languages) see Reinhart (1993), Aoun & Li (1993) and Cole & Hermon (1994).

[23] This is not what Watanabe concludes from this evidence. He assumes, following Huang (1982), that overt and covert movement fundamentally differ in that only the former is subject to Subjacency. This means that the Subjacency effect in (65a) gives evidence for S-structure movement. Watanabe then argues that at S-structure a null operator moves, while what we conceive of as the WH-expression remains in situ throughout the derivation (cf. also Aoun & Li 1993 on Chinese). Watanabe rejects analyses of the type assumed in the text (i.e. LF movement of one WH but not the others, with Subjacency holding for covert movement as well), on the basis of evidence suggesting that the spec-CP of embedded questions is already filled at S-structure. The evidence is based on the possible occurrence of Subjacency effects with certain overt movements in Japanese. This evidence is rather inconclusive, however. Watanabe notes (i) that Subjacency effects are lacking with the most straightforward type of overt movement in Japanese, namely scrambling, and (ii) that an alternative explanation can be given for most of the supposed S-structure Subjacency effects. But even if Watanabe is correct, this is not dramatic for our analysis. It would simply mean that Japanese is of the English type. We would only have a problem if there are no 'real' WH-in-situ languages at all.

36

To summarize, we have shown that the six possible rankings of Q-Marking, Q-Scope and Shortest Steps define four different patterns of question formation. All four patterns do actually occur. We take this as convincing evidence for the existence of these constraints, and for their operating in an optimality-theoretic fashion.

3.4 *French*

In French root questions a phenomenon can be observed that we have not discussed yet and which might seem problematic at first sight: optionality. It is possible that one WH moves to spec-VP* while the rest stays in situ, but it is also possible that all WHs remain in situ.[24]

(66) a. Qu' as-tu donné à qui
          what have you given to whom
      b. Tu as donné quoi à qui
          you have given what to whom

One way in which optionality can arise in optimality theory is when there happens to be more than one optimal candidate with respect to a given ranking; it is possible that

---

[24] If the subject is not a clitic, but a full NP like *Jean*, simple inversion as in (66a) is impossible (cf. ia). Rather, there either must be stylistic inversion as in (ib), or complex inversion (again involving a clitic) as in (ic).

(i) a. *Qui a Jean vu
         who has John seen
     b. Qui a vu Jean
         who has seen John
     c. Qui Jean a-t-il vu
         who John has he seen

Crucially, these examples do not incriminate an analysis in which the WH-movement in French is obligatorily accompanied by verb movement. On the contrary, this assumption makes possible an explanation of the fact that stylistic inversion and complex inversion do not occur in non-WH-sentences. (ia) shows that simple inversion in French is impossible when the subject is not a clitic. What is going on in (ib) and (ic), then, is that the subject is generated in a position such that it does not block verb movement (although this cannot be the sole trigger for stylistic inversion, since it is optionally available in embedded sentences as well). In (ib) it is in a right-peripheral position within VP (cf. Sportiche 1988, Friedemann 1991, and others); in (ic) it is degraded to an adjunct (adjoined to V*', cf. Rizzi & Roberts 1989) which is linked to the subject clitic. At this point we do not have an analysis for why full subjects block simple inversion (cf. Rizzi & Roberts 1989, Hulk 1993, for discussion).

candidates are different but nonetheless score equally on all constraints. However, given the discussion in section 3.1-3.3, this cannot be the case in (66). We discussed all possible rankings of the three relevant constraints and in all cases only one type of question formation turned out to be optimal. What must be the case, then, is that (66) is caused by the fact that in French two of the constraints are *unranked* with respect to each other. This leads to optionality as well. When two constraints A and B are unranked with respect to each other, then the candidate that is optimal when A outranks B and the candidate that is optimal when B outranks A are rated equally high (cf. Kager 1994). If these are different candidates, there will be a choice between the two. More specifically, the French paradigm can be explained by the constraint ranking in (67) (where '<>' indicates equal ranking).

(67)   Shortest Steps <> Q-Marking >> Q-Scope

Under the interpretation of (67) in which Shortest Steps outranks Q-Marking, the optimal candidate will be the 'Chinese' one, as the tableau in (68) shows (compare section 3.3).

(68)   *French WH in situ*

	Shortest Steps	Q-Marking	Q-Scope
WH V [t t WH]	*!*****		*
☞ [WH V WH]		*	**
V [WH t WH]	*!**	*	**
[WH [WH [t V t]]]	*!**	*	
WH V [WH [t t t]]	*!*******	*	
WH WH V [t t t]	*!*********		

Under the interpretation of (67) in which Q-Marking outranks Shortest Steps, the 'English' candidate is optimal, as can be observed in the tableau in (69) (compare section 2.3).

(69) *French WH-movement*

	Q-Marking	Shortest Steps	Q-Scope
☞ WH V [t t WH]		******	*
[WH V WH]	*!		**
V [WH t WH]	*!	***	**
[WH [WH [t V t]]]	*!	***	
WH V [WH [t t t]]	*!	*******	
WH WH V [t t t]		*******!***	

Note that our analysis explains that, although both WH-movement and supporting V-movement are optional in French, they cannot apply independently. Both (70a) and (70b) are ungrammatical.

(70) a. *Quoi tu as donné
          what you have given

      b. *As-tu donné quoi
          have you given what

The ungrammaticality of (70b) is an instantiation on a more general phenomenon. As we explained earlier, verb movement without movement of a WH will never lead to a more optimal candidate (see section 2.3). The ungrammaticality of (70a) also follows. The absence of inversion shows that the WH is in the 'Czech' position, adjoined to VP, but this structure is not optimal with respect to either of the two rankings of French.

Moreover, it follows that, if WH-movement takes place in French, superiority effects obtain. Under the 'English' ranking, which is the one that triggers WH-movement, Shortest Steps has the effect that the WH closest to spec-VP* must be moved:

(71) a. Qui a vu quoi
          who has seen what

      b. *Qu' a vu qui
          what has seen who

An apparent problem for the analysis proposed here is that in embedded questions the optionality of WH-movement disappears. Here, there must be WH-movement of one WH-expression to spec-CP:

(72) a. *Je me demande (que) tu as vu qui
          I wonder (that) you have seen who

      b. Je me demande qui (q̸u̸e̸) tu as vu
          I wonder who you have seen

Following Lasnik & Saito (1992), we would like to suggest that a simple explanation for the contrast between French main and embedded questions can be given once the notion of selection is taken into account. The difference between main and embedded questions is, of course, that the latter are selected by a matrix verb. A verb like *se demander* in (72), for example, selects a complement carrying a Q-feature. This type of selection is basically semantic in nature. Suppose, however, that there is a constraint which states that semantic selection must be reflected in syntax:

(73) *Selection*
      Selectional requirements must be satisfied at S-structure

In the case of embedded questions, Selection will be satisfied when the highest projection of the embedded sentence carries the Q-feature. In (72b) this is the case, because the WH-phrase occupies spec-CP. By spec-head agreement and percolation the Q-feature will be present on CP. In (72a), however, the Q-feature will not reach the top of the projection, since it is present on the complement of the verb.

We must now ask ourselves what the status of Selection is within the optimality theoretic theory of question formation outlined above. We would like to propose that it is not evaluated on a par with the other constraints, for reasons to be explained below. Instead, two sets of constraints must be distinguished. One set contains the constraints proposed in section 2.3, the other contains Selection and possibly other constraints. The output of the evaluation of the first set then functions as input for the evaluation of the second set. So, if the first set of constraints allows for more than one optimal structure, as in the case of French questions, the second set of constraints may pick one of these as the ultimately optimal candidate. Thus the optionality disappears. This is what happens in French embedded questions. Like in main questions the first set of constraints defines two optimal candidates. The constraint Selection, which plays no role in main questions, then filters out the one with WH-in-situ. In strict WH-in-situ languages like Chinese, the first set of constraints defines only one optimal candidate, due to the fact that Shortest Steps is not equally ranked with respect to Q-Marking, but outranks it. Hence, there is only one candidate to be considered by Selection. Given that this is so, it will be the optimal candidate, despite the fact that it violates this constraint.

One may wonder whether the difference that in French Selection is not violated while in Chinese it is, cannot be explained as well under the assumption that Selection is evaluated on a par with the other constraints. This is indeed possible. In this view,

Selection would be the highest ranked constraint in French, while in Chinese it would be ranked somewhere below Shortest Steps. As is easy to see, the facts then follow. However, if language variation is explained by different rankings of Selection with respect to the other constraints, unattested language types are predicted. Consider, in particular, the ranking, in (74).

(74)   Selection >> Shortest Steps >> Q-Marking >> Q-Scope

A language defined by (74) would obligatorily leave all WHs in situ in main clauses, while it would obligatorily move one WH to spec-CP in embedded clauses. The reason for this is obvious: Selection plays no role in main clauses, so that movement is suppressed by Shortest Steps; it does play a role in embedded clauses, so that WH-movement is triggered. To the best of our knowledge, no such language exists. We therefore feel that separate evaluation is empirically justified.

We think that there is some conceptual justification as well: there seems to be a parallel in phonology, as pointed out to us by René Kager (p.c.). In phonology, morphemes can impose selectional requirements on the prosodic shape of the base they attach to. For instance, if a language allows three prosodic patterns, a morpheme may require that its base has one of these patterns. However, in general a morpheme cannot require the base to have a pattern that is not part of the prosodic inventory of the language. That is to say, selectional requirements typically do not overrule the constraints determining the general prosodic make-up of the language, a phenomenon known as structure preservation (cf. Kiparsky 1985). This might be explained if these requirements operate on the output of the other constraints, like we propose for syntactic selectional requirements.

In sum, the optionality of WH-movement in French can be successfully analyzed in terms of equal ranking of Shortest Steps and Q-Marking. However, the possibility of equal ranking at first sight seems to reintroduce the problem of language typology. Recall that all possible rankings of Shortest Steps, Q-Marking and Q-Scope define an existing pattern of WH-question formation. As we saw for French, equal ranking has the effect that two of these patterns occur in the same language. If equal ranking is an option readily available in UG, one might expect many such co-occurrences of patterns of question formation. One might even expect a language in which all four patterns discussed in the previous sections are attested, namely when all constraints are ranked equally. In fact, however, optionality as in French seems to be rare and, to the best of our knowledge, languages allowing all four patterns do not exist. Apparently, what must be the case is that equal ranking is a marked option. This is in fact a plausible assumption, for reasons having to do with language acquisition.

Tesar & Smolensky (1993) address the question of how grammars are acquired in optimality theory. They propose a learning algorithm of which we can only give a rough

sketch here. Upon encountering a datum, the child concludes that this datum is the optimal form in the candidate set (in syntax: all outputs that target the same semantics and that are projected from the same set of lexical items). Crucially, the child then hypothesizes that all other outputs in the candidate set are less optimal. This hypothesis allows it to rank the constraints. Constraints that are violated by the encountered datum (i.e. the optimal output), but not by the other candidates, cannot be ranked among the highest ranked constraints: they must be outranked by constraints that are violated by the other candidates but not by the encountered datum. Consider for instance English WH-questions. Suppose a child encounters a sentence like *what will John read*. It will then hypothesize that other potential outputs of the underlying form (*John will read what*) are less optimal. Those potential other outputs include **what John will read*, in which a WH-expression has been fronted but the verb remains in situ, and **John will read what*, in which both elements remain in situ. Since the optimal candidate violates Shortest Steps to a greater extent than either of the two alternative outputs, this constraint must be ranked low. Since the two potential outputs, as opposed to the actual output, violate Q-Marking, this constraint must be ranked high. The child thus arrives at the correct partial ordering: Q-Marking >> Shortest Steps. Further data are required to order Q-Scope with respect to these two constraints.

The problem of acquiring a grammar with equal ranking should now be obvious. The learning strategy is based on the hypothesis that all potential outputs other than the actually attested form are suboptimal. Hence, a grammar will be acquired that rules out these other candidates. This means that if there is actually a second optimal candidate, due to equal ranking, part of the already acquired ranking must be unlearned. It seems plausible that this requires exposure to robust evidence, otherwise the child will ignore the evidence for a second optimal form like it ignores occasional mistakes in the input. This explains the relatively rare occurrence of equal ranking.

If this explanation is correct, we predict that, in the acquisition of WH-questions, French children will first go through a stage of optionality (the constraints are not yet ranked), then go through a stage in which the optionality displayed by the adult grammar is lacking (the constraints are ranked), and then learn the adult optionality (Shortest Steps and Q-Marking are 'disranked' again). This seems to be correct. Weissenborn (1993) reports that from 2;01;19 up to 2;03;21 WH-in-situ is as good as absent in Philippe (one example out of 114 WH-questions). From 2;06;13 onwards, WH-in-situ coexists with WH-movement (81 examples out of 199 WH-questions).[25]

In conclusion, the optionality in French question formation can be explained by assuming that Q-Marking and Shortest Steps are unranked with respect to each other. At

---

[25] A different matter is why French children (or Philippe at least) go through a stage of obligatory WH-movement instead of obligatory WH-in-situ. Our suggestion is that the evidence for WH-movement is more robust in the input because of its obligatory nature in embedded contexts.

the same time, we expect this type of optionality to be relatively rare, since it is difficult for the child to learn that two constraints are not ranked: the learning algorithm is designed to acquire total ranking.

4. Summary

Let us summarize the main proposal of the paper. We have argued that variation in the formation of WH-questions in such languages as English, Bulgarian, Czech, Chinese and French can be accounted for by a different ranking of three general constraints on movement and question marking, namely the following:

(75) a. *Shortest Steps*
          Minimize the distance between chain links
      b. *Q-Marking*
          A question must be overtly Q-marked
      c. *Q-Scope*
          [+Q] elements must c-command VP at surface structure

Several properties of English-type questions follow from the constraint ranking in (76).

(76) Q-Marking >> Shortest Steps >> Q-Scope

The high ranking of Q-Marking ensures that the pro⸱ ⸱r structure for Q-marking must be derived. Therefore, in English root WH-questions there must be verb movement plus movement of at least one WH to spec-VP*. (In embedded questions, we assume that a Q-marking complementizer is always present, though under certain cirumstances it is not spelled out; cf. Pesetsky 1994). In multiple questions the fate of WHs other than the one that moves to spec-VP*/CP is determined by the fact that Shortest Steps outranks Q-Scope: they remain in situ. Finally, Shortest Steps explains that the one WH that is moved is the one closest to spec-VP*/CP (superiority) (Golan 1993).

The properties of Bulgarian-type question formation follow from the constraint ranking in (77).

(77) Q-Marking >> Q-Scope >> Shortest Steps

Like in English, the proper Q-marking structure must be derived. So, one WH at least must be moved to spec-VP*. But since Q-Scope now outranks Shortest Steps, all WHs must be moved out of VP. We have argued that adjunction of the other WHs to VP blocks Q-marking, hence the ordering in (77) has the result that all WHs move to spec-VP*/CP.

The high ranking of Q-Marking also ensures that there must be verb movement. So, if the subject is not a WH itself, inversion occurs. Finally, we have shown that even though Shortest Steps is ranked lowest in (77), it still has its effects: it (indirectly) explains the effect reminiscent of superiority that within the WH-cluster the subject must come first. A grammar with identical properties is derived when the ranking of Q-Marking and Q-Scope is reversed in (77), because they are never crucially in conflict.

If the constraints are ranked as in (78), Czech-type questions are derived.

(78)   Q-Scope >> Shortest Steps >> Q-Marking

The high ranking of Q-Scope ensures that all WHs must be moved out of VP. However, they do not move to spec-VP* in this case. Given that Shortest Steps outranks Q-Marking, it is better to adjoin all WHs to VP than it is to apply verb movement plus movement of all WHs to spec-VP*. The low ranking of Q-Marking implies that there will be no verb movement, and hence no inversion, in questions in these languages. Finally, we have argued that it follows from (78) that the order of the WHs is free in this case (there are no superiority effects).

The two rankings in which Shortest Steps outranks the other two constraints have the effect that all WHs must remain in situ, as in Chinese and Japanese. Summarizing, the six possible rankings of the three constraints result in four different patterns of question formation (in two cases two different rankings result in the same pattern). As we have seen, these patterns are all attested, while we do not know of other patterns.[26] Amongst other things, it is predicted that there are no languages in which WH-questions are formed by verb movement while all WHs remain in situ, and that there are no languages in which one WH is moved without accompanying verb movement while the other WHs remain in situ.

There is some further language variation: in French root questions, either all WHs remain in situ, or one WH moves to spec-VP*, with accompanying verb movement. This

---

[26] In some languages, like Italian and Irish, multiple questions seem to be impossible. This might be incorporated into the system by adding a constraint Parse (WH), which is violated if a WH-phrase in the input is not syntactically realized (cf. Legendre, Smolensky & Wilson 1995). Italian/Irish might then be accounted for by the ranking in (i) (in the languages discussed in the text Parse (WH) is ranked high).

(i)   Q-Marking >> Q-Scope >> Shortest Steps >> Parse (WH)

The high ranking of Q-Marking ensures that there must be movement of one WH to spec-VP*/CP. The question is what happens to the other WHs in the input of a multiple question. Q-Scope for the other WHs can be satisfied in two ways: by moving them out of VP as well (see Bulgarian and Czech) or by not realizing them in syntax at all (in that case Q-Scope is satisfied vacuously). If Shortest Steps outranks Parse (WH), as in (i), the latter option must be chosen in the optimal candidate.

means it is possible to have two of the attested patterns occurring in a single language. This can be accounted for by assuming that two constraints can be unranked with respect to each other. Considerations of language acquisition suggest that this is a marked option.

If the analysis given here is correct, there is reason to believe that syntax and phonology are less different than is sometimes assumed.

References

Ackema, Peter: 1995, *Syntax below Zero*, PhD dissertation, Utrecht University.
Ackema, Peter, Ad Neeleman & Fred Weerman: 1993, 'Deriving Functional Projections', in A. Schafer (ed.), *Proceedings of NELS 23*, Amherst, GLSA, pp. 17-31.
Aoun, Joseph & Yen-hui Audrey Li: 1993, '*Wh*-elements in situ: Syntax or LF?', *Linguistic Inquiry* 24, 199-238.
Barrs, Andrew & Howard Lasnik: 1986, 'A Note on Anaphora and Double Objects', *Linguistic Inquiry* 17, 347-354.
Billings, Loren & Catherine Rudin: 1995, 'Optimality and Superiority: A new Approach to Overt Multiple-*wh* Ordering', Ms. Princeton/Rutgers and Wayne State College. To appear in *Formal Approaches to Slavic Linguistics* 3.
Blom, Alied & Saskia Daalder: 1977, *Syntaktische theorie en taalbeschrijving*, Muiderberg.
Chomsky, Noam: 1986, *Barriers*, Cambridge, MIT Press.
Chomsky, Noam: 1993, 'A Minimalist Program for Linguistic Theory', in K. Hale & S.J. Keyser (eds.), *The View from Building 20*, Cambridge, MIT Press, pp. 1-52.
Chomsky, Noam: 1994, 'Bare Phrase Structure', *MIT Occasional Papers in Linguistics* 5.
Chomsky, Noam: 1995, 'Categories and Transformations', Ms., MIT.
Cole, Peter & Gabriella Hermon: 1994, 'Is there LF *Wh*-Movement?', *Linguistic Inquiry* 25, 239-262.
Collins, Chris: 1994, 'Economy of Derivation and the Generalized Proper Binding Condition', *Linguistic Inquiry* 25, 45-61.
De Haan, Germen: 1979, *Conditions on Rules; The Proper Balance between Syntax and Semantics*, Dordrecht, Foris.
Di Sciullo, Anna-Maria & Edwin Williams: 1987, *On the Definition of Word*, Cambridge, MIT Press.
Fox, Danny: 1995, 'Economy, Scope and Semantic Interpretation', Talk given at the OTS Colloquium, Utrecht University.
Friedemann, Marc-Ariel: 1991, 'On the D-Structure Position of Subjects in French', in S. Barbiers, M. den Dikken & C. Levelt (eds.), *LCJL 3 Proceedings*, Leiden University, pp. 155-168.

Golan, Yael: 1993, 'Node Crossing Economy, Superiority and D-Linking', Ms., Tel Aviv University.

Grimshaw, Jane: 1993, 'Minimal Projection, Heads, and Optimality', Ms., Rutgers University, New Brunswick. To appear in *Linguistic Inquiry*.

Higginbotham, James & Robert May: 1981, 'Questions, Quantifiers and Crossing', *The Linguistic Review* 1, 41-80.

Huang, James: 1982, *Logical Relations in Chinese and the Theory of Grammar*, PhD dissertation, MIT.

Hulk, Aafke: 1993, 'Residual Verb-Second and the Licensing of Functional Features', *Probus* 5, 127-154.

Jensen, John & Margaret Stong-Jensen: 1984, 'Morphology is in the Lexicon!', *Linguistic Inquiry* 15, 474-498.

Kager, René: 1994, 'Ternary Rhythm in Alignment Theory', Ms., Utrecht University.

Kiparsky, Paul: 1985, 'Some Consequences of Lexical Phonology', *Phonology* 2, 85-138.

Koeneman, Olaf. 1995. 'Flexible Syntax and Sentence-Type Marking', MA Thesis, Utrecht University.

Kraskow, Tina: 1992, 'Multiple WH-Movement in Slavic', *The Slavic Syntax Newsletter* 2, 9-10.

Lapointe, Stephen: 1980, *A theory of grammatical agreement*, PhD dissertation, University of Massachusetts, Amherst.

Lasnik, Howard & Mamoru Saito: 1992, *Move α; Conditions on its Application and Output*, Cambridge, MIT Press.

Legendre, Géraldine, Paul Smolensky & Colin Wilson: 1995, 'When is Less More? Faithfulness and Minimal Links in *wh*-Chains', paper presented at the Workshop on Optimality in Syntax, MIT.

Lieber, Rochelle: 1980, *On the Organization of the Lexicon*, PhD dissertation, MIT.

May, Robert: 1977, *The Grammar of Quantification*, PhD dissertation, MIT.

May, Robert: 1985, *Logical Form; Its Structure and Derivation*, Cambridge, MIT Press.

Muysken, Pieter: 1982, 'Parametrizing the Notion Head', *Journal of Linguistic Research* 2, 57-75.

Neeleman, Ad: 1994, *Complex Predicates*, PhD dissertation, Utrecht University.

Pesetsky, David: 1994, 'Optimality Principles of Sentence Pronunciation', Ms., MIT.

Prince, Alan & Paul Smolensky: 1993, *Optimality Theory; Constraint Interaction in Generative Grammar*, Ms., Rutgers University/University of Colorado, Boulder.

Reinhart, Tanya: 1993, 'WH-in-Situ in the Framework of the Minimalist Program', Ms., Tel Aviv University/Utrecht University.

Riemsdijk, Henk van: 1978, 'On the Diagnosis of WH Movement', in S.J. Keyser (ed.), *Recent Transformational Studies in European Languages*, Cambridge, MIT Press.

Riemsdijk, Henk van: 1989, 'Movement and Regeneration', in P. Benincà (ed.), *Dialect Variation and the Theory of Grammar*, Dordrecht, Foris, pp. 105-136.

Rizzi, Luigi & Ian Roberts: 1989, 'Complex Inversion in French', *Probus* 1, 1-30.

Rudin, Catherine: 1988, 'On Multiple Questions and Multiple WH Fronting', *Natural Language and Linguistic Theory* 6, 445-501.

Sportiche, Dominique: 1988, 'A Theory of Floating Quantifiers and its Corollaries for Constituent Structure', *Linguistic Inquiry* 19, 425-449.

Tesar, Bruce & Paul Smolensky: 1993, 'The Learnability of Optimality Theory: An Algorithm and some Basic Complexity Results', Ms., University of Colorado, Boulder.

Toman, Jindrich: 1986, 'Cliticization from NPs in Czech and Comparable Phenomena in French and Italian', in H. Borer (ed.), *The Syntax of Pronominal Clitics*, New York, Academic Press, pp. 123-145.

Watanabe, Akira: 1992, 'Subjacency and S-Structure Movement of WH-in-Situ', *Journal of East Asian Linguistics* 1, 255-291.

Weissenborn, Jürgen: 1993, 'Matrix Infinitives and Economy in Language Development: Comments on Wexler', Talk given at the Conference on the Robustness of the Language Faculty, Utrecht University.

DRAFT, May 1995

Utrecht University
Research Institute for Language and Speech (OTS)
Trans 10
3512 JK Utrecht
The Netherlands

ackema@let.ruu.nl / neeleman@let.ruu.nl