Chapter 2

On *Wh*-Head Movement   Caterina Donati

2.1   Introduction

The merit of "On *Wh*-Movement" (Chomsky 1977) was that it explained in a principled way and unified under a single theory—movement— a number of properties previously unobserved, or unrelated, or un- accounted for. I will start by considering with a critical and "minimalist" eye another property of *wh*-movement that was simply assumed as a primitive in the theory, and left unexplained: the fact that *wh*-movement operates on a phrase. More generally, the theory of movement as it devel- oped in the 1970s and 1980s took for granted the existence of two config- urationally different movements, head movement and phrasal movement, which would never overlap and never compete in their domain of applica- tion, in what can really be seen as a remnant of a "construction-driven" approach to syntax.

This approach can no longer be maintained if the minimalist attitude is to be taken seriously. More precisely, an approach assuming as a primitive the existence of two kinds of movement is incompatible with a feature-driven approach. As is well known, in Chomsky 1995, 2000, 2001, movement is reanalyzed as a complex operation crucially triggered and governed by features: an uninterpretable feature α, the probe, searches locally for a matching feature β, the goal (Agree), a copy of which is merged in a local configuration for checking (Merge Copy or Internal Merge). What the computational procedure really sees is a feature, and the phrasal properties of (standard) *wh*-movement need to be derived.

The chapter is organized as follows. After first excluding simple feature movement as a syntactic option in section 2.2, in section 2.3 I address the issue of the status of head movement, reviewing with a critical eye Chom- sky's recent attempt to exclude it from narrow syntax. I conclude that

head movement does exist as a syntactic option and that it is indeed the unmarked pied-piping operation, being minimal. Next, I systematically review the conditions on movement spelled out in the theory starting from "On *Wh*-Movement," searching for a property of head movement that might make it unavailable in standard *wh*-movement contexts, hence triggering the less minimal, more costly operation of phrasal movement (section 2.4). Finding no such property, I propose in section 2.5 that the choice between phrase and head as the target of the movement operation is simply due to the phrase structure status of heads and phrases, together with the Condition on Uniformity of Chains (Chomsky 1995). In section 2.6, I provide two case studies, free relatives and comparatives, showing that *wh*-movement does not necessarily involve phrases and that the fact that it does in the classical case is just a consequence of the definition of what a head/phrase is, together with the interpretation of these constructions.

Finally, in section 2.7 I examine some cases that look like residuals of specific constraints on head movement, showing that they are effects of a phonological requirement that can interfere with any movement operation. In section 2.8, I offer some conclusions.

2.2   No Feature Movement

A preliminary question, explicitly addressed in Chomsky 1995, is why the complex operation described above as movement (Agree + Internal Merge) cannot simply occur at the level of features, and some kind of (generalized) pied-piping is necessarily involved—in other words, why features cannot move. Chomsky's first tentative answer is that features cannot move owing to a phonological constraint:

For the most part—perhaps completely—it is properties of the phonological component that require such pied-piping. Isolated features and other scattered parts of words may not be subject to its rules. (Chomsky 1995, 262–263)

The problem with this proposal is that it implies that post-Spell-Out movement, being free from any phonological constraint, should display no pied-piping. There is good evidence, however, of the existence of pied-piping—that is, of phrasal movement—at LF.

A simple case providing evidence for covert phrasal movement is antecedent-contained deletion: in (1), where the deleted VP is contained within its antecedent, the only way to recover ellipsis without descending

into infinite regress is to allow quantifier raising of the phrase *every city Bill does*.[1]

(1) John will visit [every city Bill does]
     [every city Bill does] John will visit [t]

A more convincing alternative is to identify the constraint banning feature movement with a property of Merge. When it applies to the numeration, Merge indeed is restricted to operate on lexical items, on bundles of features. It is thus natural and minimal to assume that this restriction holds as well when Merge is part of a movement operation. Let us then assume that there is a single operation Merge in the syntax, which applies freely and is constrained to operate on words (Matushansky 2006).[2]

Assuming that features cannot move (or better, be merged), some sort of pied-piping is required. The minimal amount of material a feature should pied-pipe to form a mergeable object is a word, that is, a head. This derives head movement. What about XP-movement? Being antiminimal, why should it exist at all?

2.3   The Status of Head Movement

The easiest way to solve this puzzle is to proceed as we did for feature movement, banning it from syntax. This is what Chomsky (2001) does: he assumes that head movement is not a syntactic operation but part of the "phonological branch of the computation." If this is true, then the minimal amount of material a feature can pied-pipe becomes a phrase, and the phrasal properties of (standard) *wh*-movement end up being derived for free.

This move is untenable, however, for several reasons: it is based on wrong assumptions, both theoretical and empirical; it is incompatible with other aspects of Chomsky's theory (e.g., phases); and it brings undesirable consequences.

First, this proposal is based on a wrong theoretical assumption, namely, that head movement violates crucial syntactic requirements—and in particular, that head movement does not extend the target (hence being countercyclic). This is true for the standard head adjunction configuration given in (2).

(2)

$$XP$$

$$X^0 \quad \ldots \quad YP$$

$$Y^0 \quad X^0 \quad \overline{Y^0}$$
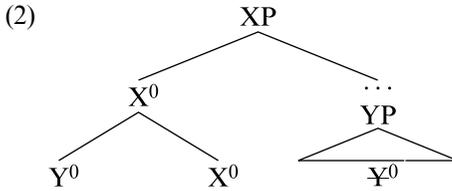
The traditional reason for this configuration is related to the standard cases of head movement usually considered in the literature: in V-to-T movement, or in V-to-C movement, the two heads conflate and behave like a single constituent, hence the assumption that they form a sort of "derived lexical item" represented in the head adjunction configuration in (2).
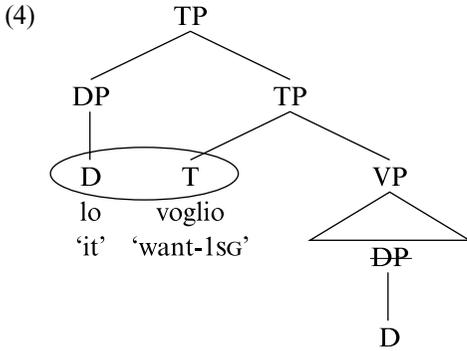
But this is not the only configuration head movement can in principle produce. Suppose we have a head X endowed with an uninterpretable feature α that needs to be deleted. In principle, nothing prevents a copy of a head Y endowed with the matching feature β from merging with the root of the structure, as in (3).

(3)

$$Y^0 \quad XP$$

$$X^0 \quad \ldots \quad YP$$

$$\overline{Y^0}$$
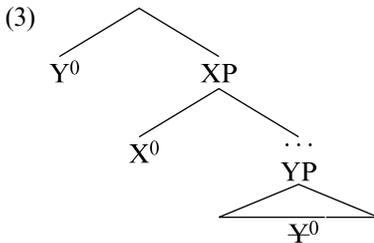
The configuration in (3) is obtained by merging a new item to the root of the tree, hence complying with the Extension Condition. Given (3), the head conflation effect correlated in many cases with this configuration can be the result of an independent process, perhaps phonological, which in principle has nothing to do with head movement: this process—call it "affixation"—is something that can happen to two adjacent heads independently of how and why they ended up being adjacent.

Separating head movement from affixation is a good idea in light of several empirical facts: first, there are cases of affixation that do not correlate with head movement, but with phrasal movement (see below); second, there are cases where head movement does not yield any effect of this kind (section 2.6); third, there are cases where affixation does not involve moved heads, but merged ones (section 2.7).

As an illustration of the first cases, consider (4) as a possible derivation for a clitic construction in Italian.

(4)

```
                    TP
                   /  \
                DP      TP
                |      /  \
          (  D    T  )    VP
             lo   voglio  /\
            'it' 'want-1SG'  D̶P̶
                              |
                              D
```

Clitic pronouns in Romance have always been problematic given standard assumptions about phrasal and head movement: clitics are clearly maximal projections in the position where they are generated, but they end up being affixed to an inflectional head. This tension is at the root of standard and influential analyses such as Kayne's (1989), where the clitic starts out as a phrase but moves and adjoins as a head, changing its phrase structure status in the course of the derivation (violating a condition like that on the uniformity of chains; Chomsky 1995). The derivation in (4) is by far more minimal: here the phrase structure status of the clitic never changes during the derivation, and the clitic moves as a phrase. The D-to-T incorporation is an independent process akin to phonology.[3]

Summarizing so far, nothing about head movement prevents it from being cyclic, and assuming the existence of cyclic head movement appears to be empirically convenient.

But also inherent in banning head movement from syntax is a troublesome implication, which appears to be too strong: being phonological, head movement should never feed LF, hence never have any interpretive effect. This may well be true if we consider movement of predicative heads, which need to reconstruct for interpretability (Heim and Kratzer 1998). But even in these cases, the issue depends crucially on the status we assign to the notion of extended projection (Grimshaw 1991): the extended projections of V, or of N, are typically defined as head movement domains, and they do look like interpretable objects as a whole.[4] On the other hand, the LF effects of head movement become simply unquestionable once we consider movement of quantificational heads, which do acquire their scope through this mechanism; we will consider a case of this type in section 2.6.

Returning once again to Chomsky's proposal, there is also a technical problem with its implementation if it is combined with other aspects of

the more recent framework: the idea that head movement is phonological clashes with the details of the derivation-by-phases approach to computation. Suppose the system is building a simple structure like (5).

(5)

$$TP$$

T^0 $vP \rightarrow$ phase

v^0 $VP \rightarrow$ sent to Spell-Out

V^0
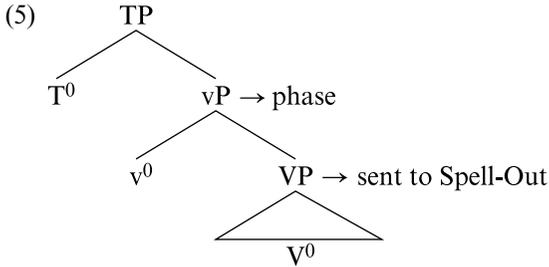
Suppose moreover that V^0 must move to T^0 to check some uninterpretable feature. At the level of the first strong phase, vP, the sister of v (VP) is sent to Spell-Out. But the edge of the current phase, v^0, is not yet accessible to phonology. In other words, the launching site and the landing site of this head movement operation are not accessible to phonology in the same phase. A solution could be that V-to-T movement happens when the next phase (CP) is sent to PF; but then the derivation must deal with two phases at once, hence raising the computational burden, contrary to the whole spirit of phase theory (Matushansky 2006).

Finally, pushing head movement onto a phonological branch of the computation has a broader undesirable consequence for the theory: it complicates the grammar by positing a ''second syntax'' in the PF branch. It also raises at least two questions. For one thing, if both phrasal movement and head movement are driven by features, it seems at least strange to locate them on two different computations. For another, consider that Chomsky (2000) assumes that PF also involves some kind of phrasal movement (e.g., extraposition; see also Lebeaux 1990), so that the problem of the duplicity of movement is reduplicated within this second syntax.

The conclusion is that there is no principled way to exclude head movement from narrow syntax, and a simple stipulation opens a number of problems. Head movement is indeed an operation available to narrow syntax; moreover, it is the default movement option, being the minimal pied-piping strategy. The next step, then, is to find some property of head movement that makes it unavailable in standard *wh*-movement contexts, hence triggering the less minimal, more costly operation: phrasal movement.

2.4   Head Movement versus Phrasal Movement

We are now looking for the property that triggers phrasal movement instead of head movement in the classical *wh*-constructions. The first place to look is in the definition of movement, among the conditions that constrain its various components. Let us review what these constraints on movement look like, given minimalist assumptions and the feature approach, to see if there are any that might be responsible for the choice between the two movements.

The traditional account relies on conditions concerning Agree, the search procedure—call them locality or closeness. Head movement, the default option, is limited by the Head Movement Constraint (HMC; Travis 1984). Whenever the goal is not (a feature of) the head of the complement of the probe, the nondefault option is chosen, and the goal pied-pipes its phrase; the HMC has recently been reframed in these terms by Pesetsky and Torrego (2001, 363), as follows:

(6) *Head Movement Generalization*
    Suppose a head H attracts a feature of XP as part of a movement
    operation.
    a. If XP is the complement of H, copy the head of XP into the local
       domain of H.
    b. Otherwise, copy XP into the local domain of H.

This is not a viable solution for both theoretical and empirical reasons. On the theoretical side, there is no room in a feature approach for such a constraint. Notice, interestingly, that Pesetsky and Torrego arrive at (6) by assuming that phrasal movement is the default option, blocked in case (a) because it results in remerging. But the system cannot work this way. If movement is triggered by a search procedure of a feature (Agree), locality must be defined in terms of features with no reference to the amount of material that gets moved with the feature. Such is the case with the Minimal Link Condition (7) and its more recent version (8).

(7) *Minimal Link Condition*
    K attracts α only if there is no β, β closer to K than α, such that K
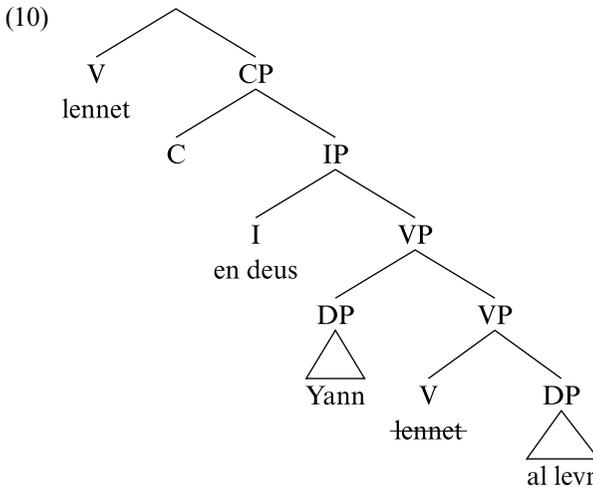   attracts β.
    (Chomsky 1995, 311)

(8) Locality conditions yield an intervention effect if probe α matches
    inactive β that is closer to α than matching γ, barring Agree (α, γ).
    (Chomsky 2001, 4)

In other words, locality is a condition on the search procedure, which is a component of movement defined at the feature level. At that level, the distinction between phrase and head plays no role and indeed does not exist at all. Therefore, the criterion we seek for choosing between head movement and phrasal movement cannot be found among the locality conditions.[5]

On the empirical side, let us look at the well-known though controversial ''long head movement'' phenomena recorded in the literature (e.g., Lema and Rivero 1990; Manzini 1994; Roberts 1994; Carnie 1995; Borsley, Rivero, and Stephens 1996). For example, the simple Breton sentence in (9) appears to instantiate the derivation in (10), where the verb raises to a left-periphery head, skipping the head I—hence violating the HMC, or the generalization (6).

(9) Lennet en   deus Yann al   levr   (Breton)
     read   3M.SG has Yann the book
     'Yann has read the book.'
     (Roberts 1998)

(10)

```
                    ┌──────────┴────────┐
                    V                    CP
                  lennet          ┌──────┴──────┐
                                  C             IP
                                        ┌───────┴────────┐
                                        I                VP
                                     en deus      ┌───────┴──────┐
                                                  DP             VP
                                                  △        ┌──────┴──────┐
                                                 Yann      V            DP
                                                        ̶l̶e̶n̶n̶e̶t̶          △
                                                                      al levr
```

The same derivation yields an ungrammatical result if the sentence contains a different kind of intervener, such as negation.

(11) *Lennet n'en   deus ket Tom al   levr
       read   NEG 3M.SG has NEG Tom the book

The explanation proposed by Roberts (1994) and others is that Relativized Minimality–type constraints such as (7) and (8) apply ''across the board,'' to any movement dependency, no matter whether it involves a

head or a phrase. In other words, what counts for locality is the class of interveners and the class of attractees, not the amount of moved material. More data of this sort are given in (12) through (14).

(12) Citao sam   knjigu   (Croatian)
     read   I-have book
     'I have read the book.'
     (Wilder and Ćavar 1994)

(13) Procel sum   knigata   (Bulgarian)
     read   I-have the-book
     'I have read the book.'
     (Borsley, Rivero, and Stephens 1996)

(14) Dar-te   he   un exemplo   (Old Spanish)
     give-you I-will an example
     'I will give you an example.'
     (Borsley, Rivero, and Stephens 1996)

The interpretation of such data in terms of long head movement is controversial and still being discussed, but the point is that these are exactly the data we predict given a feature approach to movement: if locality is a condition on the search procedure component of movement, then it is to be defined in terms of features, or classes of features. We thus predict that the same kind of constraints will govern both head movement and phrasal movement. But if locality is not responsible for the choice between head and phrasal movement, we must look at other movement constraints.[6]

A second class of constraints on movement relates to the goal, which must be active (Chomsky 2000) and not embedded in an island. But these constraints do not seem to be relevant for present purposes, since they apply to any movement operation.

Finally, a third class of constraints is conceivable, relating to the other operation involved in Move, namely, Internal Merge. The Condition on Uniformity of Chains (CUC) can be defined as such a constraint.

(15) *The Condition on Uniformity of Chains*
     A chain is uniform with regard to phrase structure status.
     (Chomsky 1995, 253)

(I will show in the next section that this constraint is only apparently irrelevant for the head/phrase distinction we are seeking.) Also conceivable is an economy condition on Internal Merge, for which what is to be merged is the minimal amount of material necessary for convergence: (16).

(16) Merge just enough material for convergence.

This is indeed the notion of minimality I referred to earlier, when I said that the minimal amount of pied-piped material is a head, giving head movement as the unmarked option.

The only conclusion we can draw from what we have seen so far is that there is no difference whatsoever between head movement and phrasal movement, except that the former is more minimal than the latter. To find a way to derive phrasal movement as an option and its selection in standard *wh*-constructions, we must turn to some other component of the grammar.

2.5   Head versus Phrase

If the difference between head and phrasal movement does not reside in the movement module, the only alternative is to rely on phrase structure theory. There is indeed a primitive difference between heads and phrases, which we will capitalize on to derive the complementarity we need: in a nutshell, a head projects; a phrase is a projection. Together with the independent conditions on Internal Merge spelled out earlier (the CUC (15) and the economy condition (16)), this primitive difference is enough to derive in a principled way (1) the complementarity between head movement and phrasal movement, (2) the choice of phrasal movement in standard *wh*-constructions, and (3) the existence of minimally different *wh*-constructions involving head movement.

But let us proceed step by step, trying to be really minimal. Suppose a probe α on a head A at the root attracts a goal β. Given that a feature cannot be merged, some extra material needs to be merged (call it B), and the offending feature is deleted (17). The operation Merge (A, B) is asymmetric in essence, so that given the configuration in (17) one of two things must happen: either A or B must project, yielding the two configurations (18) or (19). The element that projects is the head of the resulting phrase.

(17)   (18)   B   (19)   A

        B   A   B   A   B   A
   [β]   [α]
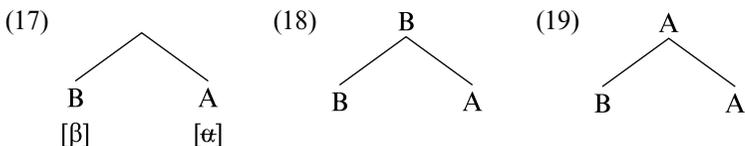
Let us combine these trivial assumptions with what we know about movement. Suppose a moved item and its copy must be the same with re-

spect to their phrase structure status (the CUC). This means that when a head, which by definition projects, gets moved, it must project; but when a phrase, itself by definition a projection, gets moved, it does not project. As a result, whenever a feature moves as a head, all the features it is associated with (and notably the categorial feature) project. Therefore, head movement changes the feature composition of the target. When the very same feature moves as a phrase, this does not happen, and the target remains unchanged.

The economy condition on Internal Merge ensures that the two movements never overlap: if the grammar always copies enough material for convergence, it will select head movement unless convergence at LF (interpretation) chooses differently.[7]

To see how this account works, let us look at (standard) *wh*-constructions. In interrogatives, the complementizer head selecting the structure contains a feature, call it *[wh]*, that needs to be checked, acting as a probe. The goal corresponds to a *wh*-element embedded in the clause, (a copy of) which needs to be merged in a local configuration with the probe. Since features cannot be merged (see section 2.2), the minimal option is for the *wh*-feature attracted by C to pied-pipe the *wh*-word alone (head movement). But this minimal option does not yield a convergent derivation: moving the *wh*-feature as a head means projecting all the features associated with it, and notably its categorial feature (D). This would turn the interrogative clause into a complex DP as in (20), which is not interpretable as an interrogative clause at the interface.[8] This is why the more costly derivation (21) is selected.

(20)　　　DP　　　　(21)　　　CP

D　　　　C　　　DP　　　C
[wh]　　[wh]　　[wh]　　[wh]
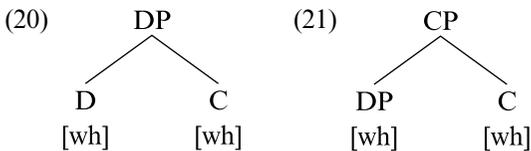
Summarizing so far, there is only one operation, Move, which is triggered by a feature and defined at that level, and which merges just enough material for convergence. Then, convergence at LF decides whether "enough material" is a head (which retains its projection property throughout the derivation, given the CUC) or a phrase (which remains a projection throughout the derivation). In standard *wh*-constructions, such as interrogatives, LF convergence selects the less minimal option, that of moving the entire phrase, preserving the simple CP categorial status of the clause.

2.6   Free Relatives and Comparatives

Good evidence that the phrasal status of *wh*-movement is only a by-product of the environments where it occurs (interrogative CPs) would then be an instance of *wh*-movement embedded in a different environment, notably in a context compatible with DP selection and showing no phrasal pied-piping.[9]

This is exactly what we find with free relatives. Consider the minimal pairs in (22a–b) and (23a–b) in Italian and English, respectively.

(22) a. *Ho   mangiato quanti   biscotti hai   preparato [t]
           have-1SG eaten   how-many cookies have-2SG prepared
           '*I have eaten what cookies you have prepared.'
     b. Mi chiedo quanti   biscotti hai   preparato [t]
           me wonder how-many cookies have-2SG prepared
           'I wonder how many cookies you have prepared.'
     c. Ho   mangiato quanto hai   preparato [t]
           have-1SG eaten   what   have-2SG prepared
           'I have eaten what you have prepared.'

(23) a. *I shall visit [*what* town] you will visit [t]
     b. I wonder [*what* town] you will visit [t]
     c. I shall visit [*what*] you will visit [t]

At first sight, free relatives (a,c) and interrogatives (b) differ in that the *wh*-element needs to be bare in the former but not in the latter.[10] This anti-pied-piping restriction correlates with a different interpretation of the two constructions: as is well known, free relatives are complex nominal structures, while interrogatives are simple clauses. This is shown straightforwardly by the selection possibilities of the two structures: free relatives like (22a,c) or (23a,c) are compatible with a DP-selecting verb like *mangiare* (24a) or *visit* (25a); interrogatives like (22b) or (23b) are compatible only with non-DP-selecting verbs like *chiedersi* (24b) or *wonder* (25b).

(24) a. Mangerò i   biscotti
           I-will-eat the cookies
     b. *Mi chiedo i   biscotti
           I   wonder the cookies

(25) a. I shall visit Rome
     b. *I wonder Rome

This correlation can be expressed through a generalization like (26).

(26) A simple *wh*-structure excludes pied-piping exactly in those cases in which it occurs in a nominal position.

Given what we already know about heads and phrases, we can suggest a principled explanation for (26). What happens in free relatives is that the *wh*-element moves as a head. This means that, besides checking the *wh*-feature on C as usual, it endows the clause with the D-feature required for interpretation, as in (20).

Consider (27) for a more detailed comparison of free relatives (27a), full relatives (27b), and interrogatives (27c). The basic structural differences are minimal, but they correlate with important interpretive and syntactic differences. Relative clauses are characterized by the fact that they are complex nominal clauses (a DP embedding a CP: Kayne 1994; Bianchi 1999), the difference between free and full relatives being that D is moved in the former and merged in the latter. Interrogatives and free relatives involve the same uninterpretable feature [wh] on C, which is checked through phrasal movement in the former and through head movement in the latter. The crossed arrows in (27a) and (27c) correspond to instances of movement violating the CUC: a phrase projecting (27a);[11] a head not projecting (27c).

(27)
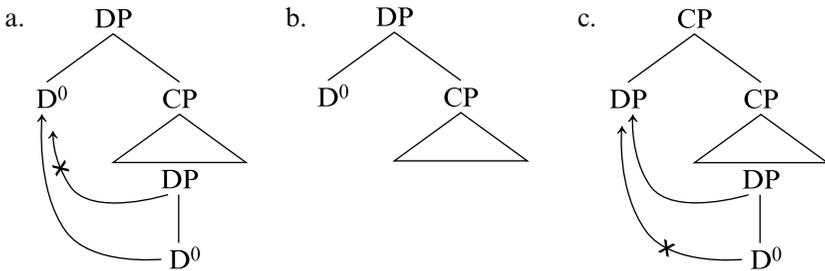
A second piece of evidence that we are on the right track comes from a different instance of the generalization in (26), represented by comparative clauses. As we will see, comparative clauses are *wh*-constructions displaying all the typical properties associated with *wh*-movement, together with clear complex nominal properties. We thus predict that comparatives should disallow phrasal pied-piping and involve *wh*-head movement. This prediction is confirmed, as the following review of the comparative construction's properties will show.

First, comparatives are *wh*-constructions, displaying all the typical properties associated with movement since Chomsky 1977. For example, they are sensitive to island effects, as (28) and (29) show for Italian and

English. In Italian, the comparative is moreover introduced by an overt *wh*-element.

(28) a. *Complex Noun Phrase Constraint*
            *I ate more cookies than I met a man who ate [t]
      b. Wh-*Island Constraint*
            *I ate more cookies than I wonder who ate [t]

(29) a. *Complex Noun Phrase Constraint*
            *Ho   mangiato più   biscotti di quanti ho   incontrato
            have-1SG eaten   more cookies of WH   have-1SG met
            un uomo che ha mangiato [t]
            a   man   who has eaten
      b. Wh-*Island Constraint*
            *Ho   mangiato più   biscotti di quanti mi chiedo chi
            have-1SG eaten   more cookies of WH   me wonder who
            ne   abbia mangiati [t]
            of-them has   eaten

Second, comparatives display clear complex nominal properties, as has been observed primarily on the semantic side. In particular, they give rise to the kind of scope ambiguities typically associated with nominals, as the following classic example from Russell 1905 shows:

(30) I thought your yacht was larger than it is
      a. [[than it is] [I thought your yacht was larger]]   (consistent)
      b. I thought [[than it is] [your yacht was larger]]   (inconsistent)

The easiest and most standard way to account for the ambiguity of (30) is to derive it from the scope possibilities of the comparative clause, so as to include or exclude the belief verb from the comparison, as informally represented in (30a–b) (von Stechow 1984).

The complex nominal character of comparatives is also visible in their interpretation, which is that of amount or degree *descriptions* such as (free) relatives; in this way, they differ from corresponding *wh*-interrogatives, which are simple amount or degree *predicates* (e.g., Heim 1985; Donati 2000). Again, this is represented informally in (31).

(31) Mary ate more cookies than Paul did
      $\exists y$ [[y > ιx [Paul ate x-many cookies] $\land$ [Mary ate y-many cookies]]

On the syntactic side, comparatives are strong islands for extraction, just like all complex nominal clauses, but unlike standard *wh*-constructions.

(32) a. *When do you eat more cookies in the morning than Paul does
         [t] [t]?
     b. *What do you eat the soup more quickly than Paul does [t] [t]?

Given that comparative clauses are *wh*-structures and they display these nominal properties, the prediction is clear: comparatives should involve *wh*-head movement. This is exactly what the data in (33) and (34) show, from Romanian and Bulgarian, respectively.

(33) a. *Maria e cu   mult mai deşteaptă *decît   de frumoasă* e
         Maria is with much more clever   of-WH of beautiful is
         Zamfira [t]
         Zamfira
         '*Maria is much more clever than beautiful Zamfira is.'
     b. Maria e cu   mult mai deşteaptă *decît* e Zamfira [t] de
         Maria is with much more clever   of-WH is Zamfira   of
         frumoasă
         beautiful
         'Maria is much more clever than Zamfira is beautiful.'
     c. *Cît de frumoasă* e Zamfira [t]?
         WH of beautiful is Zamfira
         'How beautiful is Zamfira?'
         (Grosu 1994)[12]

(34) a. *Ivan izpi   povece vino ot-*kolkoto*   bira Maria izpi [t]
         Ivan drank more   wine of-how-much beer Maria drank
         '*Ivan drank more wine than beer Maria drank.'
     b. Ivan izpi   povece vino ot-*kolkoto*   Maria izpi [t] bira
         Ivan drank more   wine of-how-much Maria drank beer
         'Ivan drank more wine than Maria drank beer.'
     c. *Kolko*   bira izpi   Maria [t]?
         how-much beer drank Maria
         'How much beer did Maria drink?'
     d. *Kolko*   izpi   Maria [t] bira?
         how-much drank Maria   beer
         '*How much did Maria drink beer?'
         (Izvorski 1995)

(33) and (34) show that the very same *wh*-element (*cît* in Romanian, *kolko* in Bulgarian[13]) moves as a head in comparatives (a–b), but as a phrase in interrogatives (c). This provides good evidence that the *wh*-element

involved in the embedded comparative is a Q head, which endows the clause with the Q-feature required for interpretation. The same evidence holds also in English and Italian.[14] In English, however, the facts are partially obscured since the *wh*-element is null and the contrast with interrogatives is not minimal.

(35) a. *Mary ate more cookies than candies she ate
      b.   Mary ate more cookies than she ate candies
      c.   How many candies did she eat?
      d. *How many did she eat candies?

In Italian, as we will see in the next section, the movement pattern illustrated so clearly by the Romanian and Bulgarian data is complicated by a gap in the paradigm that requires an explanation.

   All the data discussed in this section provide strong evidence for the following analysis: Comparatives and free relatives, on the one hand, and interrogatives, on the other hand, display a minimal syntactic contrast that correlates with a semantic difference. While all involve *wh*-movement, triggered by the same feature and constrained by the same restrictions, free relatives and comparatives move a bare head, which changes the CP configuration domain into a DP/QP category, turning the clause into a complex nominal; interrogatives move a phrase, which leaves the CP configuration domain unchanged as a simple (interrogative) clause. This is what we expect if movement is an operation merging just enough material for convergence, and if heads and phrases differ only in their projection ability.

2.7 The Freezing Effect of Agreement

As I mentioned, Italian comparatives require a final glance. In Italian, while there is clear evidence that the *wh*-element *quanto* moves as a head (see (36)), the situation is complicated by some gaps in the paradigm: in particular, *quanto* never strands its complement (37a), except when it is cliticized with *ne* (37c) or when it shows no agreement with its adjectival complement (37b).

(36) a. *Maria ha mangiato più   biscotti di *quante   caramelle* Paolo
          Maria has eaten   more cookies of WH-F.PL candies   Paolo
          abbia mangiato [t]
          has   eaten
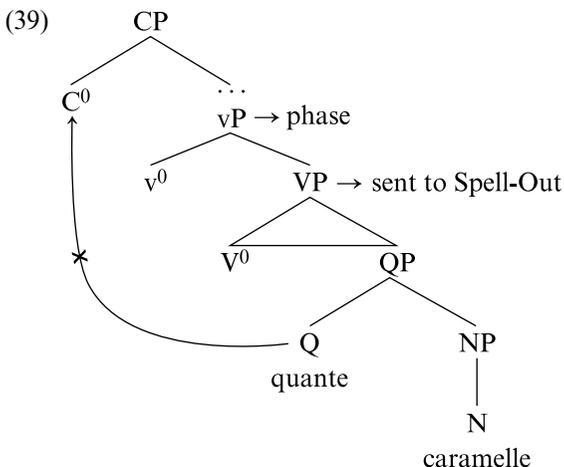   '*Maria has eaten more cookies than candy Paolo ate.'

b. *Maria è più   intelligente di *quanto bella*   sia [t]
   Maria is more intelligent of WH   beautiful is
   '*Maria is more intelligent than beautiful she is.'

(37) a. *Maria ha mangiato più   biscotti di *quante*   Paolo abbia
      Maria has eaten   more cookies of WH-F.PL Paolo has
      mangiato [t] *caramelle*
      eaten   candies
      'Maria has eaten more cookies than Paolo has eaten candies.'

   b. Maria è più   intelligente di *quanto* sia [t] *bella*
      Maria is more intelligent of WH   is   beautiful
      'Maria is more intelligent than she is beautiful.'

   c. Maria ha mangiato più   biscotti di *quanti*   *ne*   abbia
      Maria has eaten   more cookies of WH-M.PL of-them has
      mangiati [t] Paolo
      eaten   Paolo
      'Maria has eaten more cookies than Paolo has.'

If we compare these data with those of Bulgarian, the generalization
seems to be that agreement within the QP has a sort of a freezing effect,
blocking head movement of the Q alone and the corresponding stranding
of the remnant of the phrase. The same effect seems at work in the free
relative cases discussed above: both in English and in Italian (38), the D
head can never strand its nominal complement.[15]

(38) a. *Mangerò *quanti*   vorrai   [t] biscotti
      I-will-eat how-many you-will-want cookies
   b. *I shall visit *what* you will visit [t] *town*

This kind of gap in the paradigm is very important for the hypothesis
under discussion here, because it seems like a residue of a construction-
specific property of head movement, which would contradict our attempt
at unification. We will see, however, that this constraint has nothing to do
with head movement per se and that the same kind of effect occurs with
phrasal movement. We will therefore end up confirming the claim that
movement is a unitary operation. But let us proceed step by step.

   Consider in (39) the (simplified) structure of a comparative construc-
tion containing an agreeing QP, like *quante caramelle* in (37a).

(39)

```
            CP
          /    \
        /       \
      C⁰         ···
       ↑       vP → phase
       |      /    \
       |    v⁰      VP → sent to Spell-Out
       |           /    \
       *         V⁰       QP
       |                 /  \
       |                Q    NP
        _____/|    |
                     quante  N
                          caramelle
```
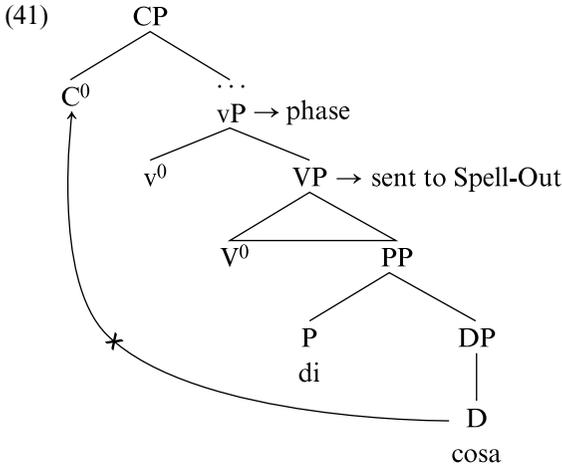
Suppose Q-N agreement means Q has an uninterpretable feature that needs to be checked by a feature of N. Head movement is not triggered here because it is useless: the head-head checking configuration is already there, merged in the base. Since agreement feeds pronunciation, this configuration must be conserved until Spell-Out. This means that at the level of the first strong phase (vP), Q and NP must be adjacent. But at that point, VP, the sister of v, becomes inaccessible to the computation (Phase Impenetrability Condition[16]). There is therefore no way of moving Q without N: if it moves before Spell-Out, agreement cannot feed pronunciation; after that, it violates the PIC. The restriction on movement at play in (37) is due not to a peculiarity of head movement, but to the effect of a phonological adjacency requirement.

Crucially, a similar effect is attested with phrasal movement, confirming that it has nothing to do with head movement per se. As is well known, preposition stranding is prohibited in those languages, like French and Italian, where P and D tend to conflate in a P+D suppletive form (Law 1998; Salles 1997). In (40), this restriction against preposition stranding is shown at play in an Italian interrogative.

(40) *Cosa hai   parlato [di t]?
      what you-have talked   of

Consider the corresponding simplified structure in (41).

(41)

```
              CP
           ╱      ╲
        C⁰        ...
                  vP → phase
                ╱    ╲
              v⁰      VP → sent to Spell-Out
                    ╱    ╲
                  V⁰      PP
                        ╱   ╲
                      P      DP
                      di      │
                              D
                             cosa
```

If P and D need to conflate, they need to be adjacent at Spell-Out. But at the first strong phase, when vP is sent to Spell-Out, VP becomes no longer accessible to the computation: the DP *cosa* cannot make a long movement alone, and preposition stranding is barred.

2.8   Conclusions

The theory of movement as it emerged from "On *Wh*-Movement" is still the core of the current theory of movement. This chapter has focused on what appears to be a stipulation of such a theory in light of the minimalist framework—the limitation of *wh*-movement to phrases—and has critically reviewed the more general division of movement into two distinct subtypes: head movement and phrasal movement.

Having critiqued Chomsky's recent attempt to ban head movement from syntax and reviewed the different constraints on movement and their legitimacy in a minimalist framework, I conclude that there is only one movement, triggered by features and defined at that level. The difference in the amount of material pied-piped in different environments is due to an economy principle on Merge that declares, "Merge just enough material for convergence." Given that principle, the existence of two distinct movements can be derived from the irreducible difference that holds by definition between heads and phrases: heads project, phrases are projections. A natural implication of this conclusion is that there is no principled reason for *wh*-movement to be restricted to phrases in any context, as the two cases of *wh*-head movement discussed in this chapter illustrate.

Notes

Different versions of this chapter have been presented at Going Romance 2002, the workshop "On *Wh*-Movement," and GLOW 2003. Many thanks to the audiences for their important comments, questions, and suggestions.

1. There is a highly detailed and well-motivated recent literature dedicated to covert phrasal movement: see Sauerland 1998, Fox 1999, Cecchetto 2000, Nissenbaum 2000, and Pesetsky 2000, among others. See Brody 1995 for a different view also discussing antecedent-contained deletion cases.

2. See Pesetsky 2000 for a different view, where both feature movement and phrasal movement are available to syntax as different operations.

3. The phenomena related to Romance cliticization are outrageously simplified in the text, which is meant only to show that the standard analysis may be complicated by a mistaken theoretical assumption (i.e., that affixation *is* head movement) that we might want to eliminate. (4) is not meant to be a full-fledged alternative analysis of cliticization, of course, an analysis that should for example take into account the differences among clitics standardly reduced to the distinction between a class of syntactic clitics and a class of phonological ones.

4. See also Zwart 2001 for a discussion of some clear interpretive effects of instances of verb movement.

5. In fact, there is a sense in which a head is indeed always more remote than its phrase: fewer nodes separate the probe from the phrase bearing the goal than separate it from the corresponding head. But this notion of locality is too restrictive: it amounts to excluding head movement as a whole, which is not what we want.

6. There is an alternative I will not explore here, as it is less minimal than the one I choose to pursue. Suppose head movement and phrasal movement are systematically triggered by two intrinsically different classes of features; then it is conceivable that they are sensitive to different interveners and display different locality constraints. See Matushansky 2006 for a proposal along these lines.

7. The economy condition in (16) also makes an interesting prediction concerning movement of a feature associated with a given category to a target associated with the same feature. Since in this case the projection property of heads does not affect the categorial status of the target, the minimal option should be compulsory, and phrasal movement systematically excluded. More precisely, we predict derivations moving a DP into Spec,DP, or a PP into Spec,PP, or a CP into Spec,CP to be ruled out as violations of the economy condition in (16). This means that standard analyses of constructions like the following need to be revised:

(i) *Possessives*
    John's wife
    (DP in Spec,DP; see, e.g., Kayne 1994)

(ii) *Exclamatives*
    Into the dungeon with the traitor!
    (PP in Spec,DP; see Barbiers 1995)

(iii) *Clitic left-dislocations*

Che fosse un assassino, l'ho   sempre saputo
that was an assassin   it(CL)-have-1SG always known
(CP in Spec,CP; see, e.g., Cecchetto 1999)
'That he was an assassin, I have always known it.'

Addressing the details of such a revision goes far beyond the limits of this chapter. Notice that some idiosyncratic properties of these constructions would benefit from an analysis in which the dislocated constituent is adjoined to some extra functional head: in particular, this would explain the impossibility of recursion in both exclamative PPs and clitic left-dislocations, and the exclamative mood of examples like (ii). Thanks to Peter Svenonius for pointing out these facts to me.

8. As an anonymous reviewer points out, it is not true that a nominal cannot be interpreted as an interrogative. This is exactly what happens with so-called concealed questions, as discussed in Grimshaw 1979 and illustrated in (i).

(i) a.   I know what answer he gave
    b.   I wonder what answer he gave
    c.   I know the answer he gave
    d. *I wonder the answer he gave

Under certain predicates (e.g., *know* but not *wonder*), a simple DP can replace a full-fledged interrogative. The existence of such a class of nominals does not affect the conclusions drawn here: without entering into the debate about what the structure of concealed questions is, it will be enough to say that the complex DPs resulting from head movement discussed here are not of the appropriate form—that is, are not concealed questions.

9. Again, the fact that some environments selecting for interrogative CPs also admit special DPs as concealed questions is not addressed here. I will simply concentrate on standard nominals. See note 8.

10. There is an exception to this generalization that readily comes to mind: both English and Italian have a class of free relatives that appear to allow some amount of pied-piping.

(i) Mangerò   *qualunque biscotto* preparerai [t]
    will-eat-1SG whatever cookie will-prepare-2SG
    'I will eat whatever cookie you will prepare.'

(ii) I shall visit [*whatever* town] you will visit [t]

The fact that this movement pattern correlates with the presence of some extra material in the head of the clause, the suffix -*ever*/-*unque*, suggests an obvious solution: namely, that these relatives are only apparently "free" (i.e., defective) and in fact correspond to full relative clauses. As such, they are generated as the complement of an external determiner (Kayne 1994; Bianchi 1999), the universal quantifier -*ever*/-*unque*.

(iii) I shall visit [DP[D ever] [CP[DP what town] [. . .]]]

For some reason perhaps related to its universal value (Larson 1988), -*ever*/-*unque* triggers the raising and head adjunction of the *wh*-determiner, yielding (iv).

(iv) $[_{DP}[_D \text{ what}_i [_D \text{ ever}]] [_{CP}[_{DP} t_i \text{ town}] [\dots]]]$

See Battye 1989 for a series of empirical arguments demonstrating that these relatives ought to be treated as ''pseudo-free relatives.'' See also Kayne 1994:154 for a similar analysis proposed on totally different grounds.

11. There is more than I have said in the text, of course: while the definition of heads as projecting features is certainly correct, that of phrases as nonprojecting features leaves adjunction out of the picture. This could be a problem for the analysis proposed here if a derivation such as (i) were allowed.

(i)
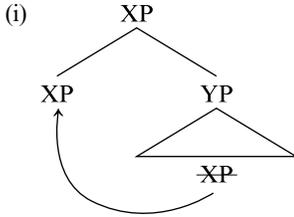
Here, the phrase XP is moved and it projects in its landing site, turning the root into a right adjunct. There are, however, good reasons to exclude such a derivation on independent grounds. One concerns the legitimacy of the chain formed in (i), which violates a strict version of the CUC, since XP is maximal in its base position, but nonmaximal in its landing site (see Chomsky 1995, 257–259, for discussion). Another reason concerns the structure in (i), which is a case of right-adjunction and questionable as such (Kayne 1994). Thanks to Howard Lasnik for pointing this potential problem out to me.

12. Notice that Grosu (1994) proposes a different analysis of the movement pattern illustrated in these data.

13. Notice that the *wh*-element *kolko* is introduced by a special preposition *ot* in comparatives; other such prepositions are found in Italian (*di*; see (36)–(37)) and Spanish (*de*; see note 14, (iii)–(iv)). A similar analysis can be given of *de* in *decît* in Romanian. In Donati 2000, I discuss the nature of this kind of particle in detail, showing that it does not form a constituent together with the *wh*-element it precedes, but instead governs the entire comparative clause.

14. As an anonymous reviewer points out, this crosslinguistic generalization does not hold in Spanish, at least if we consider the construction discussed by Price (1990, 43) and illustrated in (i) and (ii).

(i) a.   Mi padre vende más libros que discos compra mi madre
         my father sells   more books than records buys   my mother
    b. *Mi padre vende más libros que compra mi madre discos
        my father sells   more books than buys   my mother records
        'My father sells more books than my mother buys records.'

(ii) a.   La mesa es más larga que ancha es la   puerta
          the table is more large than wide   is the door
     b. *La mesa es más larga que es la   puerta ancha
         the table is more large than is the door   wide
         'The table is larger than the door is wide.'

In these examples, a null operator appears to pied-pipe its complement, contrary to what my analysis of head movement predicts. I have no proposal to make about this construction (but see Kennedy 2002), although its strongly emphatic flavor may be invoked to justify a more complex structure for this type of comparative. Notice moreover that the preferred construction for comparatives in Spanish, illustrated in (iii) and (iv), indeed exhibits the movement pattern predicted by the analysis, with a bare element *lo* moving alone, stranding its complement.

(iii) a. Mi padre vende más libros de lo que mi madre compra (en) discos
          my father sells   more books of Op that my mother buys   (as) records

b. *Mi padre vende más libros de lo discos que (mi madre) compra
          my father sells   more books of Op records that (my mother) buys
          (mi madre)
          (my mother)
          'My father sells more books than my mother buys records.'

(iv) a. La mesa es más larga de lo que la puerta es ancha
          the table is more large of Op that the door   is wide

b. *La mesa es más larga de lo ancha que (la puerta) es (la puerta)
          the table is more large of Op wide   that (the door) is (the door)
          'The table is larger than the door is wide.'

Thanks to Daniela Capra for these data.

15. This possibility appears to be realized in other languages. For example, so-called internally headed relative clauses in Japanese (e.g., Shimoyama 1999) and nominalized PE-clauses in Italian Sign Language (LIS; Cecchetto, Geraci, and Zucchi, to appear) can be analyzed as involving the movement of a D head alone, stranding its nominal complement. Interestingly, no agreement surfaces in the DP either in Japanese or in LIS. See Branchini and Donati 2005 for details.

16. Chomsky (2001, 13) states the Phase Impenetrability Condition as follows (emphasis original):

(i) [In a phase α with head H,] the domain of H is not accessible to operations outside α; only H and its *edge* are accessible to such operations.

References

Barbiers, Sjef. 1995. Another case of scrambling. *Linguistics in the Netherlands 1995*, 13–24.

Battye, Adrian. 1989. Free relatives, pseudo-free relatives and the syntax of CP in Italian. *Rivista di Linguistica* 1, 219–250.

Bianchi, Valentina. 1999. *Consequences of antisymmetry: Headed relative clauses*. Berlin: Mouton de Gruyter.

Borsley, Richard, Maria Luisa Rivero, and John Stephens. 1996. Long head movement in Breton. In Richard Borsley and Ian Roberts, eds., *The syntax of the Celtic languages: A comparative perspective*, 53–74. Cambridge: Cambridge University Press.

Branchini, Chiara, and Caterina Donati. 2005. A snippet for the typology of relativization strategies: Relatives in Italian Sign Language. Paper presented at GLOW in Asia 2005, Nehru University, Delhi.

Brody, Michael. 1995. *Lexico-Logical Form*. Cambridge, Mass.: MIT Press.

Carnie, Andrew. 1995. Nonverbal predication and head movement. Doctoral dissertation, MIT.

Cecchetto, Carlo. 1999. A comparative analysis of left and right dislocation in Romance. *Studia Linguistica* 53, 40–67.

Cecchetto, Carlo. 2000. Doubling structures and reconstruction. *Probus* 12, 1–34.

Cecchetto, Carlo, Carlo Geraci, and Alessandro Zucchi. To appear. Strategies of relativization in LIS. *Natural Language and Linguistic Theory*.

Chomsky, Noam. 1977. On *wh*-movement. In Peter W. Culicover, Thomas Wasow, and Adrian Akmajian, eds., *Formal syntax*, 71–132. New York: Academic Press.

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

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In Roger Martin, David Michaels, and Juan Uriagereka, eds., *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, 89–155. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz, ed., *Ken Hale: A life in language*, 1–52. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 2004. Beyond explanatory adequacy. In Adriana Belletti, ed., *Structures and beyond: The cartography of syntactic structures, vol. 3*, 104–131. Oxford: Oxford University Press.

Donati, Caterina. 2000. *La sintassi della comparazione*. Padova: Unipress.

Fox, Danny. 1995. Condition C effects in ACD. In Rob Pensalfini and Hiroyuki Ura, eds., *Papers on minimalist syntax*, 105–120. MIT Working Papers in Linguistics 27. Cambridge, Mass.: MIT, Department of Linguistics and Philosophy, MITWPL.

Fox, Danny. 1999. Reconstruction, binding theory, and the interpretation of chains. *Linguistic Inquiry* 30, 157–196.

Grimshaw, Jane. 1979. Complement selection and the lexicon. *Linguistic Inquiry* 10, 279–326.

Grimshaw, Jane. 1991. Extended projection. Ms., Brandeis University.

Grosu, Alexander. 1994. *Three studies in locality and Case*. London: Routledge.

Heim, Irene. 1985. Notes on comparatives and related matters. Ms., University of Texas.

Heim, Irene, and Angelika Kratzer. 1998. *Semantics in generative grammar*. Oxford: Blackwell.

Izvorski, Roumyana. 1995. A DP-shell for comparatives. In Antonietta Bisetto, Laura Brugè, João Costa, Rob Goedemans, Nicola Munaro, and Ruben van de Vijver, eds., *Proceedings of Console III*, 99–121. The Hague: Holland Academic Graphics.

Kayne, Richard. 1989. Facets of Romance past participle agreement. In Paola Benincà, ed., *Dialect variation and the theory of grammar*, 85–103. Dordrecht: Foris.

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

Kennedy, Chris. 2002. Comparative deletion and optimality in syntax. *Natural Language and Linguistic Theory* 20, 553–621.

Larson, Richard. 1988. Scope and comparatives. *Linguistics and Philosophy* 11, 1–26.

Law, Paul. 1998. A unified analysis of preposition stranding in Romance and Germanic. In Pius Tamanji and Kiyomi Kusumoto, eds., *Proceedings of NELS 28*, 219–234. Amherst: University of Massachusetts, GLSA.

Lebeaux, David. 1990. Relative clauses, licensing and the nature of the derivation. In Juli Carter, Rose-Marie Déchaine, Bill Philip, and Tim Sherer, eds., *Proceedings of NELS 20*, 318–332. Amherst: University of Massachusetts, GLSA.

Lema, José, and Maria Luisa Rivero. 1990. Long head movement: ECP vs. HMC. In Juli Carter, Rose-Marie Déchaine, Bill Philip, and Tim Sherer, eds., *Proceedings of NELS 20*, 333–347. Amherst: University of Massachusetts, GLSA.

Manzini, M. Rita. 1994. Locality, minimalism, and parasitic gaps. *Linguistic Inquiry* 25, 481–508.

Matushansky, Ora. 2006. Head movement in linguistic theory. *Linguistic Inquiry* 37.1.

Nissenbaum, Jonathan. 2000. Investigations of covert phrase movement. Doctoral dissertation, MIT.

Pesetsky, David. 2000. *Phrasal movement and its kin*. Cambridge, Mass.: MIT Press.

Pesetsky, David, and Esther Torrego. 2001. T-to-C movement: Causes and consequences. In Michael Kenstowicz, ed., *Ken Hale: A life in language*, 355–426. Cambridge, Mass.: MIT Press.

Price, Susan. 1990. *Comparative constructions in Spanish and French*. London: Routledge.

Rivero, Maria Luisa. 1991. Long head movement and negation. *The Linguistic Review* 8, 319–351.

Roberts, Ian. 1994. Two types of head-movement in Romance. In David Lightfoot and Norbert Hornstein, eds., *Verb movement*, 207–242. Cambridge: Cambridge University Press.

Roberts, Ian. 1998. Have/Be raising, Move F, and Procrastinate. *Linguistic Inquiry* 29, 113–125.

Russell, Bertrand. 1905. On denoting. *Mind* 14, 479–493.

Salles, Heloisa. 1997. Prepositions and the syntax of complementation. Doctoral dissertation, University of Wales, Bangor.

Sauerland, Uli. 1998. The meaning of chains. Doctoral dissertation, MIT.

Shimoyama, Junko. 1999. Internally headed relative clauses in Japanese and E-type anaphora. *Journal of East Asian Linguistics* 8, 147–182.

Stechow, Arnim von. 1984. Comparing semantic theories of comparison. *Journal of Semantics* 3, 1–77.

Travis, Lisa. 1984. Parameters and effects of word order variation. Doctoral dissertation, MIT.

Wilder, Chris, and Damir Ćavar. 1994. Long head movement? Verb movement and cliticization in Croatian. *Lingua* 93, 1–58.

Zwart, C. Jan-Wouter. 2001. Syntactic and phonological verb movement. *Syntax* 4, 34–62.