Late Merge of Degree Clauses [1]

Rajesh Bhatt & Roumyana Pancheva

University of Texas at Austin & University of Southern California

The focus of this paper is the place of degree clauses (i.e. *than/as-* phrases) in the overall architecture of comparatives. We propose that degree clauses are merged late, after the degree head *-er/as* has moved to a scope position. [2], [3] The position of merge of the degree clause is the position in which the degree clause is pronounced. Degree clauses do not move by themselves nor do they move covertly with the degree head. The movement of the degree head is covert; it has no effect on how the generated structure is realized at PF. The covert nature of this movement is not the result of a post Spell-Out timing; temporally it *precedes* the merge of the degree clause which does affect PF. Rather, the movement of the degree head is covert because the lower copy of the chain is pronounced instead of the head of the chain. This may be so because of general properties of QR or it may be the result of morphological well-formedness conditions on the realization of the degree head affix.

Our proposal allows us to refine, and importantly, motivate, Williams' generalization with respect to degree constructions. Williams (1974) noted a correlation between the scope of the DP out of which a constituent has been extraposed, and the adjunction site of the extraposed expression. Fox and Nissenbaum (1999) and Fox (2002) provide an analysis of extraposition which ensures that the scope of a source DP is at least as high as the attachment of the extraposed relative clause. Adopting the mechanism used by Fox and Nissenbaum (1999) to derive extraposition, we show that the scope of *-er* is exactly as high as the surface position of the degree clause. The 'exactly as high' part does not follow from the derivational mechanism. Instead we show that it follows from the semantics of degree heads

and Trace Conversion, a mechanism by which movement structures involving copies are interpreted (cf. Fox (2001), and Fox (2002)). Finally, the question of what can be merged late arises. It has been previously proposed that adjuncts but not arguments can be merged countercyclically (cf. Lebeaux (1990), Chomsky (1993), Fox and Nissenbaum (1999), Fox (2002)). We show that the fact that degree clause complements can be merged late unlike the complements of restrictors of determiners follows from the manner in which Trace Conversion proceeds.

The possibility of late merge for degree clauses has several implications. An operation that does not affect phonology - the movement of *-er* - precedes an operation that does affect phonology - the merge of the degree clause. This argues against a model of grammar where syntax has two components, preceding and following a single point in the derivation such as Spell-Out. Rather, it suggests a single-component grammar (cf. Fox and Nissenbaum 2000, Fox (2002)) or variants thereof, such as grammar based on multiple TRANSFER (cf. Chomsky (2001)).[4] Our proposal also provides indirect support for what has come to be known as the Phonological Theory of QR (cf. Bobaljik (1995), Bobaljik (2002); Pesetsky (2000), a.o.), namely the idea that QR is a covert operation simply because at PF it is the lower copy of the quantificational DP that is pronounced.

Before we present our proposal in detail, we consider the issue of constituency in degree constructions - an issue that has proven to be far from trivial. Because of conflicting requirements on the relationship between *-er*, the gradable predicate (e.g., *tall*), and the degree clause, the overall structure of comparatives has remained elusive. A number of divergent analyses have been proposed (a representative list of the various approaches includes Bresnan (1973), Bresnan (1975), Larson (1988), Moltmann (1992), Kennedy (1997b), Lechner (1999), Heim (2000)). Below we present the arguments that have motivated the alternative analyses of

the position of merge of the degree clause. Readers familiar with the architectural issues may want to skip sections 1 and 2.

1 Constituency in Degree Constructions

Comparatives exhibit a range of properties that cannot be easily accommodated into a single structure. In particular, there is convincing syntactic and semantic evidence that the degree clause is the complement of *-er*. However, there is some morphological evidence that *-er* forms a constituent with the gradable predicate (e.g., *taller*) to the exclusion of the degree clause. Moreover, not only are *-er* and the degree clause non-adjacent in the majority of cases, but more often than not they may not even appear together as a constituent that excludes the gradable predicate (e.g., *tall*). These diverging properties suggest that the degree clause is obligatorily extraposed. The question then arises as to why this is so. We address this question in §7.

1.1 The Degree Clause is an Argument of the Degree Head

1.1.1 Selectional restrictions despite non-adjacency

One of the clearest syntactic pieces of evidence that the degree clause forms a constituent with the degree head are the selectional restrictions between the two. As shown in (1), differential comparatives (i.e., comparatives with *more, less* and *fewer*) co-occur with a degree clause introduced by *than*, whereas equatives (i.e., *as*-comparatives) have a degree clause introduced by *as*.

(1)   a. Cleo ate *more* apples *than/*as/*that* Matilda did.

      b. David is *less* worried *than/*as/*that* Monica is.

      c. Simone drank *fewer* beers *than/*as/*that* Alex did.

d. Anastasia is *as* tall *as/ *than/ *that* Daniel is.

The variety of differential comparatives can be reduced to a single case of *-er* in a direct relationship with the *than* degree clause. Bresnan 1973, a.o., analyzes *more* as the degree head *-er* plus *many* or *much*. Similarly, she proposes that *less* is *-er* + *little*, and its counterpart *-er* + *few* is overtly expressed as *fewer*. In other words, just like *fewer people* is derived from *[-er few] people*, *more people* has an underlying structure *[-er many] people*, *more coffee* has the structure *[-er much] coffee*, and *less coffee* is derived from *[-er little] coffee*. Bresnan proposes that adjectival comparisons, e.g. *more interesting*, also involve *many/much*, and even synthetic forms such as *happier* have the structure [[*-er* much] happy]. We adopt the essence of Bresnan's analysis. [5] Given this analysis, it is possible to reduce the co-occurrence facts to the following:

(2) a. i. *-er* (+ many/much = *more*) ... *than*

   ii. *-er* (+ little = *less*)... *than*

   iii. *-er* (+ few = *fewer*)... *than*

   b. *as* (+ many/much/little/few)... *as*

In other words, there are selectional restrictions between *-er* and *than*, and between *as* and *as*. Selectional restrictions are the hallmark of head-argument relationships. It is thus reasonable to conclude that the degree clause is the syntactic argument of the degree head, as have Bresnan (1973), Bresnan (1975), Carlson (1977), a.o.[6] [7]

1.1.2 Semantic constituency despite non-adjacency

In addition to the syntactic selectional effects described above, there is semantic evidence that the degree head and the degree clause form a constituent to the

exclusion of the gradable predicate.

It has often been assumed that gradable adjectives denote a relation between individuals and degrees. Under this view, measure phrases such as *6 feet* are seen as the referential degree argument of predicates such as *tall*. If degrees can be explicitly referred to (cf. (3)), it is also to be expected that they can be quantified over (cf. (4)).

(3)   a. John is 6 ft tall.

        b. John is that (much) tall.

(4)   a. John is taller than 6 feet.

        b. John is [$_{AP}$ [$_{DegP}$ *-er* than 6 ft] tall]

        c. [$_{DegP}$ *-er* than 6 ft]$_1$ John is [$_{AP}$ t$_1$ tall]

In other words, *-er* and the degree clause form a semantic constituent, a degree phrase that is interpreted as a degree quantifier argument of the matrix gradable predicate (cf. Cresswell (1976), von Stechow (1984), Heim (1985), Heim (2000) a.o.). Under fairly standard assumptions, quantificational expressions undergo QR; similarly, the degree quantifier [*-er* + the degree clause] may be analyzed as moving to a scope position within the clause from where it binds the degree variable in argument position, as in (4c). Such an analysis is developed by Heim (2000).

Admittedly, the argument around (4c) is theory-dependent: to the extent that quantificational expressions may be composed with predicates without QR, there will be no evidence that *-er* and the degree clause move and take scope together, separately from *tall*. There is, however, additional evidence, coming from considerations of ellipsis resolution, that *-er* and the degree clause can scope together and thus form a semantic constituent.

Carlson (1977) observes that antecedent contained deletion (ACD) in relative clauses cannot be resolved if the head of the relative clause is a weak noun phrase. Diesing (1992) notes that this is also the case for bare plurals (e.g. 5a).[8] Standard accounts of ACD resolution posit QR of the whole noun phrase - the head together with the relative clause. After QR, the ellipsis site is no longer antecedent-contained and the matrix VP contains the necessary A'-gap to serve as a suitable antecedent for the ellipsis in the relative clause. Within the context of Diesing (1992)'s proposal that weak NPs do not undergo QR, the unacceptability of ACD with weak NPs follows.

Carlson, however, further notes that comparatives license ACD (see (5b, c)). In light of the account of (5a), the grammaticality of (5b, c) is puzzling: the weak noun phrases *more trees* and *higher trees* (the latter a bare plural) should not undergo QR and thus ACD in the degree clause should not be resolved.

(5) a. *John was climbing trees that Bill was.

     b. John was climbing more trees than Bill was.

     c. John was climbing higher trees than Bill was.

Wold (1995) suggests a solution to the above puzzle.[9] He proposes that what undergoes QR in (5b, c) is the degree quantifier formed of *-er* and the degree clause, as shown in (6a).[10] This movement removes the ellipsis site from the antecedent and also leaves a gap of the appropriate type. Thus the antecedent now contains the necessary constituent for copying (e.g., the underlined phrase in (6b). Similar considerations apply in the case of (5c).

(6) LF for (5b):

     a. [*-er* [(than) *wh* Bill was $\triangle$]]

         [John was climbing $d-$many trees]

b. [-*er* [(than) $\lambda\, d$ Bill was <u>climbing $d-$many trees</u>]]

[$\lambda\, d$ John was <u>climbing $d-$many trees</u>]

The fact that *-er* and the degree clause can take scope together can account not only for ACD resolution, as discussed above, but also for ambiguities with respect to intensional predicates. The following sentence from Heim 2000 is two-ways ambiguous.

(7) (This draft is 10 pages long.) The paper is required to be exactly 5 pages longer than that.

    a. required $>$ -er: required [[exactly 5 pages -er than that][the paper be t-long]]

    $\forall w \in$ Acc: max $\{$d: long$_w$(p,d)$\}$ = 15 pages

    b. -er $>$ required: [exactly 5 pages -er than that] [required [the paper be t-long]]

    max $\{$d: $\forall w \in$ Acc: long$_w$(p,d)$\}$ = 15 pages

Under one of its reading, (7) means that the paper is exactly 15 pages long in every acceptable world (= 7a). On its other reading, (7) says that the paper is exactly 15 pages long in those acceptable world where it is shortest. In other words, the paper must be at least 15 pages long (= 7b). (7) is ambiguous in exactly the way that that is predicted if the *-er* can move to a position above or below *required*.

Heim (2000) points out that a similar ambiguity arises with possibility operators and exactly differentials, as well as with, as pointed out by Stateva (1999), with *less*-comparatives.

(8) (This draft is 10 pages long.)

    a. The paper is allowed to be exactly 5 pages longer than that.

      i. allowed > -er: $\exists w \in$ Acc: max $\{$d: long_w(p,d)$\}$ = 15 pages

     ii. -er > allowed: max $\{$d: $\exists w \in$ Acc: long_w(p,d)$\}$ = 15 pages

  b. The paper is required to be less long than that.

      i. required > less: $\forall w \in$ Acc: max $\{$d: long_w(p,d)$\}$ $<$ 10 pages

     ii. less > required: max $\{$d: $\forall w \in$ Acc: long_w(p,d)$\}$ $<$ 10 pages

  c. The paper is allowed to be less long than that.

      i. allowed > less: $\exists w \in$ Acc: max $\{$d: long_w(p,d)$\}$ $<$ 10 pages

     ii. less > allowed: max $\{$d: $\exists w \in$ Acc: long_w(p,d)$\}$ $<$ 10 pages

(from Heim (2000)

Assuming that *-er* and the degree clause can take scope together, separate from the matrix degree predicate makes available an elegant analysis for ellipsis resolution. The ambiguities that surface with comparatives in intensional contexts also receive a natural explanation. This success in handling ellipsis resolution and ambiguities in intensional contexts can be taken as evidence in favor of the degree head and the degree clause forming a constituent together that excludes the AP.

To summarize, there are very good syntactic and semantic reasons to posit that *-er* first forms a constituent with the degree clause, and not with the matrix gradable expression. Moreover, given the selectional restrictions between *-er* and *than*, and between *as* and *as*, it is reasonable to posit that the degree clause is an argument of the degree head.

However, there is also evidence pointing to the opposite conclusion, that it is *-er* and the gradable predicate that form a constituent, to the exclusion of the degree clause. The next section presents this evidence in some detail.

1.2 Constituency of the Degree Head and the Matrix Gradable XP

1.2.1 Suppletive forms

One of the reasons for positing constituency for *-er* and the gradable predicate has been the existence of fully and partially suppletive forms, as in (9) and (10).

(9)   a. [*-er* good] → *better*

      b. [*-er* bad] → *worse*

(10) [*-er* tall] → *taller*

   It needs to be remembered, though, that the existence of suppletive forms is compatible with *-er* and the adjective being sisters, but does not require such constituency. If the the structure is in fact [[*-er* t] tall], where *t* is the trace of the extraposed degree clause, merger of -er and *tall* can still obtain, given that the two will be linearly adjacent at the point of vocabulary insertion. For instance, Embick and Noyer (2001) argue that, because the forms in (9) are lexically conditioned, and those in (10) are prosodically conditioned, they are formed after vocabulary insertion. Thus, according to them, such forms are derived by an operation that is sensitive to conditions of linear adjacency, and not solely of syntactic structure.

1.2.2 Obligatory non-adjacency between *-er* and Degree Clause

Much stronger evidence against constituency for *-er* and the degree clause seems to come from the fact that in general,[11] it is not possible for the two to appear together (cf. (11-14)).

(11) a. *Ralph is [more than Flora is] tall.

        cf. Ralph is taller than Flora is.

b. *Ralph has [more than Flora does] cars.

       cf. Ralph has more cars than Flora does.

(12) a. *Ralph is [more than her] tall.

          cf. Ralph is taller than her.

      b. *Ralph has [more than her] cars.

          cf. Ralph has more cars than her.

(13) a. *Ralph is [more than he is fit] tall.

          cf. Ralph is more tall than he is fit.

      b. *Ralph has [more than he has bikes] cars.

          cf. Ralph has more cars than he has bikes.

(14) a. *Ralph is [more than fit] tall.

          cf. Ralph is more tall than fit.

      b. *Ralph has [more than bikes] cars.

          cf. Ralph has more cars than bikes.

The above sentences may actually be reduced to just two cases. If (12) is an-
alyzed as a clausal comparative (like (11)) to which Comparative Ellipsis has ap-
plied, the two will be amenable to a uniform analysis. Similarly, the sentences in
(14) may be derived from the ones in (13) through ellipsis. In any event, even if
we need an explanation for two rather than four cases, it is still true that these
sentences represent a serious challenge to the analysis of degree clauses as argu-
ments of *-er*. The proponents of this analysis could not claim that the need for
ellipsis resolution is the reason for obligatory extraposition of the degree clause.
Some of the prohibited degree clauses do not contain ellipsis (cf. (13)). For the
ones that do, extraposition alone cannot be the answer to ellipsis resolution. Ex-
traposition may solve antecedent containment, but will not, by itself, create a gap

of the needed type in the matrix clause, to provide the necessary antecedent for the ellipsis site. For instance, if the degree clause in (11a) is to be extraposed, the matrix clause will contain the expression *[[-er t] tall]* whereas the one needed to satisfy the parallelism requirement on ellipsis resolution is *t-tall*. It could be argued that additional movement of *-er* will result in the creation of a proper antecedent, yet the same effect could be achieved by movement of [*-er* + degree clause], without extraposition. Thus, the unacceptability of the above sentences remains a mystery.

In sum, it appears to be the case that *-er* and the degree clause can never form a constituent at the position in which *-er* is initially merged. This is a very strong argument for the constituency of *-er* and the degree predicate to the exclusion of the degree clause. An alternate analysis that assumes that *-er* and the degree clause form a constituent to the exclusion of the degree predicate needs to countenance obligatory extraposition of the degree clause and furthermore explain why this extraposition is obligatory. We believe that the arguments for treating the degree clause as an argument of *-er* are convincing. Therefore we will assume that there is, in effect, obligatory extraposition of the degree clause and in §7 we will provide an explanation for why this extraposition is obligatory.

2   The Architecture of Degree Constructions

Because of the conflicting requirements on the constituency in degree constructions, their overall architecture has remained an unresolved and much debated issue. Below we present the two main approaches to the relationship between *-er* and the degree clause.

2.1 The Classical View

The classical view - the earliest generative analysis of comparatives - assigns constituency to *-er* and the degree clause. In particular, *-er* and the degree clause form a degree quantifier, which is a syntactic specifier to the gradable predicate, as in (15), (cf. Chomsky (1965), Selkirk (1970), Bresnan (1973), Heim (2000)).[12][13]
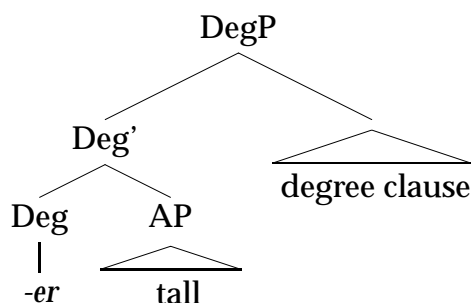
(15)
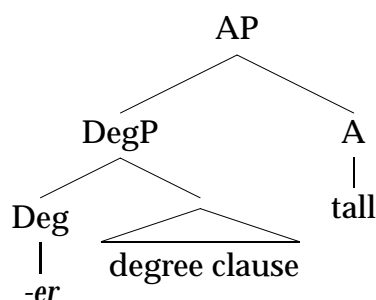
As we can expect, given the discussion in section 1.1, the classical analysis runs into several problems. Most importantly, in this analysis the degree clause has to obligatorily extrapose (cf. (11-14)). This movement is not motivated by anything apart from the need to derive the correct surface word order.[14] Analogous extraposition of relative clauses, for instance, is never obligatory. Furthermore, LF-movement of the degree clause together with *-er* (cf. Heim (2000), and the discussion around examples (4) and (5)) would necessitate an obligatory reconstruction of the degree clause after the extraposition. This raises a range of issues. Given that currently reconstruction is taken to involve not literally putting back the moved element in the position of origin, but simply interpretation of its copy, the LF-movement of [*-er* + degree clause] would require moving the copy of the extraposed degree clause at LF. But if that was a general option in the grammar, a well-known generalization about A′-movement would be directly contradicted: overtly A′-moved elements have frozen scope, i.e., they cannot undergo further movement at LF (cf. Aoun et al. (1981)). If overt A′-movement could be

'undone' by LF-movement of the copy, frozen scope would not obtain. An LF-movement of the copy account also predicts that the surface position of the degree clause will be irrelevant, in fact invisible, to semantic interpretation. But in fact, as we will discuss later, the surface position of the degree clause marks the exact scope of *-er*.

2.2   A Common Alternative

The alternative to the classical analysis posits non-constituency of *-er* and the degree clause. A representative structure is illustrated in (16) (cf. Abney (1987), Larson (1988), Corver (1990), Corver (1993), Kennedy (1997b), Kennedy (2002), for variants on this structure.)

(16)

```
                        DegP
                       /    \
                 Deg'         \
                /    \       /___\
            Deg      AP    degree clause
             |      /__\
            -er      tall
```

The alternative analysis, while avoiding some of the problematic cases of extraposition, still has problems of its own. The *-er* and the degree clause do not form a constituent that excludes the AP. Hence one would not expect the two to be able to move together without the AP. However, as we have discussed, for interpretive purposes, *-er* and the degree clause can take scope together, separate from the AP (see (5), (7)).

There are alternatives to the classical view which posit a coordination structure for comparatives (Hankamer (1973), Napoli (1983), Moltmann (1992), for subset of cases Corver (1990), Corver (1993), and aspects of Lechner (1999).).

These do not have the problems of extraposition, but they still have difficulties with assigning scope for *-er* and the degree clause, and moreover, they run into additional problems pertinent to the posited coordination structure for comparatives.

3 Late Merge of the Degree Clause Resolves the Conflict

We propose that the conflicting pieces of evidence regarding constituency in degree constructions discussed above can be reconciled if it is posited that the degree clause is merged late - still as an argument to the degree head, but not at the point at which the degree head enters the derivation. Thus, at the initial steps in the derivation of a comparative, *-er* is a sister to the gradable predicate, whereas later in the derivation, it becomes a sister to the degree clause (hence the contradictory behavior and the widely divergent analyses).

Specifically, we propose that the degree clause is merged countercyclically, after *-er* moves covertly to its scope position. The selectional restrictions between *-er* and the degree clause (*-er...than*, *as...as*) obtain because the degree clause is merged as an *argument* to the QR-ed and right-adjoined *-er*; i.e., the selectional restrictions can be stated in a local head-to-head configuration. There is no need to posit rightward A′-movement of the degree clause, thus avoiding empirical and conceptual problems. The fact that *-er* and the degree clause receive scope together - for ellipsis resolution and with respect to intensional predicates - follows directly, given that the degree clause is merged to *-er* precisely at *-er's* scope position. Furthermore, this approach allows us to derive the fact that there is a correlation between the surface position of the degree clause with the semantic scope of the degree head.

14

The lack of adjacency effects between the degree clause and the degree head follows from the fact that only the tail of the *-er*-chain is pronounced. So even though the degree clause forms a constituent with *-er*, this is not reflected in the phonology. And given that the (pronounced) copy of *-er* is in a configuration where it is syntactically local and adjacent to the AP, a PF mechanism can naturally be used to derive suppletive effects. The pronunciation of the copy of *-er* may be a property of QR in general, as held by the Phonological Theory of QR (Bobaljik 1995, Bobaljik (2002), Pesetsky 2000, Fox and Nissenbaum 1999, Fox 2001). It also appears to be the case that a morphological condition on *-er* is responsible for the pronunciation of the copy, i.e., *-er* is an affix, in that it needs to be spelled out together with a category of measure phrases such as *many/much* or adjectives.[15] This point applies to *as, so, too* as well, which are morphologically prefixes, rather than suffixes like *-er*.

This proposal in its essentials follows Fox & Nissenbaum (1999)'s analysis of relative clause extraposition. Developing Lebeaux (1990)'s proposal that relative clauses can be countercyclically merged, Fox & Nissenbaum propose that relative clause extraposition involves countercyclic merger of the relative clause to an unpronounced copy of a QP that has undergone QR. Our analysis not only extends the idea of countercyclic merge to the domain of comparatives, it also shows that not only adjuncts but complements as well can be merged late.
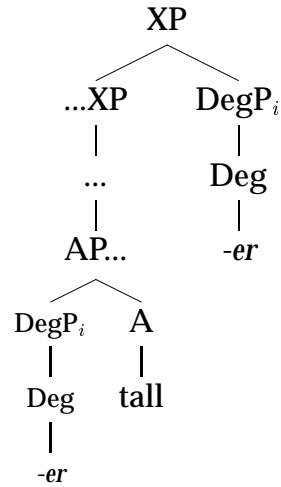
Let us illustrate the proposal in some more detail. We assume that *-er* is the head of a DegP, and (for ease of presentation) that the DegP is an argument of the gradable predicate, as in the classical analysis. We further adopt (for further convenience) that the DegP is a specifier of the gradable predicate.[16]

(17)

```
            AP
          /    \
       DegP     A
         |      |
        Deg    tall
         |
        -er
```
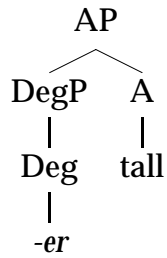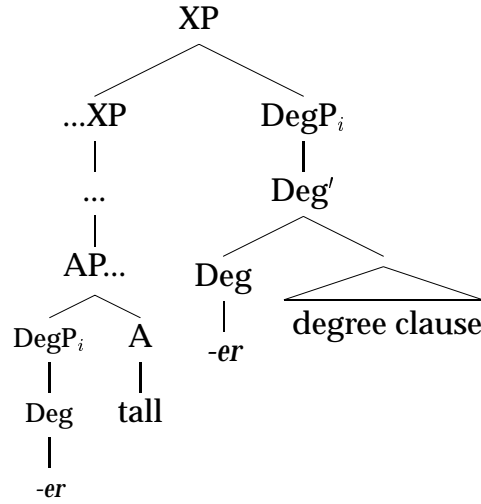
Being a quantificational expression, the DegP headed by *-er* undergoes QR, leaving behind a copy. Following Fox and Nissenbaum (1999), we assume that it right-adjoins in a scope position i.e. it adjoins to a node of type t (indicated as XP in the trees in (18) and (19)). The degree clause is then merged as an argument to the QR-ed *-er* (see (18) and (19) for an illustration).

(18)

```
                    XP
                  /    \
              ...XP    DegP_i
                |        |
               ...      Deg
                |        |
              AP...     -er
             /    \
         DegP_i    A
           |       |
          Deg     tall
           |
          -er
```

(19)

XP
...XP   DegP$_i$
...   Deg$'$
AP...   Deg
DegP$_i$ A   -er   degree clause
Deg   tall
-er

The degree head *-er* is interpreted in its scope position (with the copy of *-er* interpreted as a degree variable), but is pronounced in its base position.

Given that the merge of the degree clause is countercyclic, it targets a position that is not the root. However, this is precisely the remerge operation that also happens in the case of head-movement, thus should not be seen as problematic (at least not more so than head-movement itself).

Our analysis predicts that the degree clause can end up at different heights in the tree, depending on the scopal position of *-er*. In other words, assuming that in (20) *-er* can QR to the AP (given predicate-internal subjects this is the lowest position where the DegP can be interpreted), the infinitival IP, or the matrix IP, the degree clause can be merged as an argument of *-er* at these three locations, corresponding to the different interpretations. The LFs in (20a-c) reflect the possible positions of merge of the degree clause.

(20)   Bill wants to be taller than John is.

    a.   John wants PRO$_j$ to be [$_{aP}$ [$_{aP}$ t$_j$ [$_{AP}$ t$_i$ tall] [-*er$_i$* [than Bill is *tall*]]]]

    b.   John wants [$_{IP}$ [$_{IP}$ PRO to be [$_{AP}$ t$_i$ tall]] [-*er$_i$* [than Bill is *tall*]]]

    c.   [$_{IP}$ [$_{IP}$ John wants [$_{IP}$ PRO to be [$_{AP}$ t$_i$ tall]]] [-*er$_i$* [than Bill is *tall*]]]

The following example illustrates a case where the degree clause is discontinuous not only from the degree head but also from the degree predicate. In cases like this, both the classical analysis and its alternative (as discussed in Section 2) have to posit rightward A$'$-movement of the degree clause. In our analysis such rightward A$'$-movement is not necessary. The degree clause is merged where it appears, namely higher than the tense-level adverbial. The only movement is the QR of the DegP headed by *-er*. This is illustrated in (21).

(21) Nicole made more money last year than Tom did.

  a. 'Nicole made more money last year' is generated.

  b. QR of *-er* takes place. The lower copy is pronounced and so this step is 'covert.'
      Overt structure: [[Nicole made more money last year] *-er*]
      LF: [[λd Nicole made d-much money last year] -er]

  c. The degree complement is merged with *-er*
      Overt structure: [[Nicole made more money last year] [*-er* than Tom did Δ]]
      LF: [[λd Nicole made d-much money last year]
              [-er [λd Tom made d-much money last year]]]

Our approach is similar in spirit to Gueron & May's (1984) analysis of extraposition in that the constituency between the extraposed constituent and the phrase from which it is taken to be extraposed (the source DP) is established at a scope position. Gueron & May propose that in cases of extraposition from DPs, following the overt movement of the extraposed constituent, the source DP raises at LF to a position governing the extraposed expression. Thus both for Gueron & May and for us, the establishment of the relationship involves QR of the source of the

extraposition (DegP in the case of degree constructions). In the model of grammar assumed by Gueron & May, QR takes place at LF and hence this relationship is also established at LF. Since we adopt a single cycle model of grammar where QR (and 'covert' movement in general) can be followed by overt movement, the establishment of this relationship does not have to wait until LF. Another difference between our approach and Gueron & May's, in fact an advantage of our approach, is that the degree clause is merged as the complement of the degree head, allowing for straightforward semantic composition. If the degree clause was merged first, then subsequent movement of *-er* could not establish a head-complement relationship, thus creating a problem for the syntax-semantics mapping.

The remaining questions are why late merge of the degree clause is obligatory and why the surface position of the degree clause marks the scope of *-er.* In other words, why can the degree clause not be merged as the complement of *-er* in the base position of *-er* and be pronounced in that position and further why can the *-er* and the degree head once merged in a scope position not move further covertly? We believe the answer to these questions follows from a recent proposal made by Fox (2001)/Fox (2002) concerning the interpretation of copies and the fact that comparative quantifiers are not conservative. Merging the degree clause as the complement of *-er* in its base position and subsequently moving it with *-er* to a scope position would yield a semantically illegitimate object. Therefore late merge of the degree clause with *-er* in a scope position is the only way of achieving a semantically wellformed LF. A further covert movement of *-er* with its complement to a higher scope position would again yield a semantically illegitimate object. This is what blocks movement of *-er* with its degree clause complement.

This section illustrated the basics of our proposal regarding the architecture of degree constructions. The solution that we offered to the problem of constituency explores a possibility which has not been previously considered, and which, importantly, relies on no special assumptions. The two crucial aspects of the proposal - countercyclic merge and comparative operator movement - are operations that have been independently proposed and are, arguably, well-justified (the former outside of the domain of comparatives). Our contribution is to relate these two ideas in a way that resolves much of the previously contradictory evidence concerning the structure of comparatives and that directly relates the surface position of the degree clause with the scope of *-er*. The rest of the paper is devoted to presenting detailed evidence in support of our proposal for late merge of degree clauses.

4 Aspects of DegP Movement

4.1 Constraints on DegP Movement

Our proposal crucially relies on the idea that comparatives involve movement of *-er* (or of a larger constituent consisting of the *-er* and a differential phrase, henceforth DegP). The DegP is a quantificational argument of the degree predicate (e.g., *tall*) and as such it undergoes QR (cf. Heim 2000). There have been arguments, however, against positing such movement in the LF-syntax of comparatives. One influential argument against movement of the DegP is due to Kennedy (1997) and is developed further in Heim (2000). Kennedy noted that putative movement of the DegP predicts the existence of readings that are not in fact attested. Consider (22).

(22)   John is 4' tall. Every girl is exactly 1" taller than that.

a. $\forall > $ -*er*:

$\forall x$ [girl(x) $\rightarrow$ max {d: tall(x,d)} = 4' + 1"]

b. -*er* $> \forall$:

max {d: $\forall x$ [girl(x) $\rightarrow$ tall(x,d)]} = 4' + 1"

(ex. 22 from Heim (2000))

(22) has the reading indicated in (22a), where the DegP does not move over the universal quantifier. According to this reading, every girl is exactly 1 inch taller than John. The reading indicated in (22b) is, however, unavailable. If it was available, we could judge (22) as true in the scenario where the shortest girl is exactly 4'1" but the other girls are taller. Similar points are made by (23).

(23)　　a.　(The frostline is 3 and a half feet deep.) Mary set every post exactly 2 feet deeper than that. (= ex. 25 from Heim (2000))

　　　　b.　(John gave every candidate an A.) ?Mary was less impressed with every candidate than that. (= ex. 26 from Heim (2000))

Both (23a, b) lack the reading corresponding to the DegP taking scope over the QP. For example, (23) is not judged true in the scenario where Mary set just the least deep post exactly 2' below the frostline, but set the other posts deeper.

The following generalization emerges from the above facts.

(24)　The Heim-Kennedy Constraint:

If the scope of a quantificational DP contains the trace of a DegP, it also contains that DegP itself. (= ex. 27 from Heim (2000))

Based on his observation that a DegP can never scope over a QP, and that more generally putative degree quantifiers always take the narrowest possible scope, Kennedy (1997) concludes that there is no such thing as DegP movement. In fact,

DegP movement is not even an option under the structure he assumes where there is no degree quantifier that can be moved around.

We take the Heim-Kennedy Constraint to reveal an important aspect of the syntax and semantics of degree constructions, but following Heim (2000), we do not take the generalization to show that there is no DegP movement. We view it as a filter on abstractions over degree variables that rules out degree abstractions of the form in (25).

(25)   * $\lambda d \ldots \mathrm{QP} \ldots d \ldots$

It is not clear why there is a constraint like (25), but there seem to be similar constraints on extractions involving *why*, which create abstractions over reason variables. Such extractions cannot cross over a quantifier or negation (cf. Beck (1996)).

(26)   a. Why did John not bring potato salad? (why > not, *not > why)

b. Why did everyone bring potato salad? (why > everyone, *everyone > why)

The task of determining what the Heim-Kennedy Constraint follows from is left for future work.

4.2   The Occasional Semantic Indetectability of DegP Movement

The careful reader will have noted that the examples used by us to demonstrate the semantic effects of DegP Movement had a very particular form. These examples are repeated here as (27).

(27)   (This draft is 10 pages long.)

a. The paper is required to be exactly 5 pages longer than that.

b. The paper is allowed to be exactly 5 pages longer than that.

c. The paper is required to be less long than that.

d. The paper is allowed to be less long than that.

They all involve exactly differentials or *less*-comparatives with a referential *than* clause. We briefly discuss the reasons for the particular choices we made. We used a referential *than* clause to be able to focus on the scope of the DegP with respect to the intensional predicate and to abstract away from the additional question of *de re* vs. *de dicto*. We take *de re* interpretation to not be a matter of scope i.e. a *de re* interpretation is compatible with either wide or narrow scope for the DegP. However, a *de dicto* interpretation of the DegP necessitates narrow scope for the DegP. Consider (28), where the *than*-clause is not referential in form.

(28) John wants to be richer than Bill is.

a. *than*-clause *de re*, -er > want:
Bill is d_1-rich. John wants to be d_2-rich. $d_2 > d_1$.

b. *than*-clause *de re*, want > -er:
Bill is d_1-rich. John's desire: to be richer than d_1.

c. *than*-clause *de dicto*, want > -er:
John's desire: to be richer than Bill

The system under discussion assigns the three distinct LFs (28a-c) to (28). The *de dicto* narrow scope interpretation is quite clearly distinction from the two *de re* interpretations and so we can set it aside. One simple way of doing so is by using a referential *than*-clause. The question that arises now is whether we can distinguish between the narrow and wide scope *de re* readings. Intuitively, it seems that we can. The wide scope *de re* reading seems to involve the expression of a non-comparative desire, while the narrow scope *de re* reading seems to involve the expression of a comparative desire.

Unfortunately, our intuitions turn to be unreliable here. As Heim (2000), footnote 17 points out, despite their form LFs like (28b) do not (necessarily) involve comparative desires. To see this point clearly let us consider the semantic representations assigned to the *de re* LFs in (28).

(29)   (Bill's financial worth is 1 billion dollars.)

    a. *than*-clause *de re*, -er > want:

       max $\{$d: $\forall w \in$ Boul: rich(j,d)$\}$ > \$ 1bn.

    b. *than*-clause *de re*, want > -er:

       $\forall w \in$ Boul: [max $\{$d: rich$_w$(j,d)$\}$ > \$ 1bn.]

Given the semantics assumed above for *want*, it turns out that the wide scope *de re* LF and the narrow scope *de re* LF are semantically equivalent. (29a) is true if in all of John's desire worlds, John is rich to a degree that exceeds a billion dollars. In other words, even in the desire world where John is poorest, he has more than a billion dollars. (29b) is true if in all of John's desire worlds, John's total worth exceeds a billion dollars. Once again, even in the desire worlds where John is poorest, he has more than a billion dollars. (29a) and (29b) are semantically equivalent and we cannot use truth-conditional judgements to distinguish between them. Similar problems arise with intensional operators like *required* and *allow* in (30).

(30)   (The draft is 10 pages long.)

    a. The paper is required to be longer than that.

      i. -er > required: max $\{$d: $\forall w \in$ Acc: long$_w$(p,d)$\}$ > 10

      ii. required > -er: $\forall w \in$ Acc: [max $\{$d: long$_w$(p,d)$\}$ > 10]

    b. The paper is allowed to be longer than that.

      i. -er > allowed: max $\{$d: $\exists w \in$ Acc: long$_w$(p,d)$\}$ > 10

ii. allowed > -er: $\exists w \in \text{Acc}$: [max $\{$d: $\text{long}_w(\text{p,d})\} > 10$]

The wide scope and the narrow scope readings of the examples in (30) are se-
mantically equivalent, in exactly the same way in which the wide and the narrow
scope readings of (29) were equivalent. Adding an exactly differential or moving
to a *less* comparative makes the wide scope and the narrow scope readings truth
conditionally distinct.

There is also another complication that needs to be kept in mind when we
are trying to detect the ambiguity of a DegP with respect to an intensional op-
erator using truth conditional judgements. The semantics of certain intensional
operators combines with DegP movement in such a way that the high scope and
narow scope readings are not truth-conditionally distinct. This seems to be the
case with 'neg raising' predicates like *should*, *supposed*, and *want*. Heim (2000)
suggests a preliminary semantics that relates the 'neg raising' nature of the se-
mantics of these predicates and the semantic non-effect of DegP movement with
these predicates.

The moral of the story in this section is that DegP movement does not always
have truth-conditional effects. If we want truth-conditional effects, we need to
provide the right ingredients - an appropriate predicate (*require, allow, need, be
able*) together with an *exactly* differential or a *less* comparative. However, the
absence of truth-conditional effects does not indicate an absence of DegP Move-
ment. Truth conditions are just one way of detecting DegP Movement. We will
see that Ellipsis Resolution and Condition C can be used to diagnose DegP Move-
ment even when truth conditional judgements do not help us to distinguish be-
tween wide and narrow scope.

5 Evidence from Extraposition

As a descriptive term, extraposition refers to the phenomenon of a constituent of a phrase appearing discontinuous from the phrase and to its right. This phenomenon has been offered a number of analyses. Perhaps most commonly, it is proposed that that the discontinuous constituent is merged with its source phrase and is then rightward-moved to its surface position - the A′-movement analysis. Another analysis has the discontinuous constituent moving to the left and then being stranded by the source phrase which moves leftward even higher - the remnant-movement analysis. Finally, there are proposals that the discontinuous constituent is base generated in its surface position, and various mechanisms are employed to ensure it is interpreted together with its 'source' phrase - the base-generation analysis. Below, we use 'extraposition' largely pre-theoretically, as a description of the cases where the degree clause appears discontinuous from the degree predicate, separated from it by sentential material such as, e.g., adjuncts (as in example (21) above). As we stated above, we take such degree clauses to be merged in their surface positions, i.e., we employ a version of the base-generation analysis.[17] We consider the interaction of 'extraposition' with (i) Binding Theory Condition C effects, and (ii), scope of the comparison. Both of these interactions provide evidence that argues in favor of merging degree clauses late, post-QR of the DegP headed by *-er*.

5.1 Extraposition and Condition C

It is well-known that standard cases of A′-movement cannot reverse Condition C effects. *Wh*-movement in (31a) and QR in (31b) remove the DP containing the R-expression *John* from the c-command domain of the 3rd person singular masculine pronoun, at Spell Out or LF, respectively. Yet a Condition C violation still

obtains in these cases suggesting obligatory reconstruction (or interpretation of the lower copy) of the A′-moved DP, or at least of the part of the DP containing the R-expression.

(31)   a. ?? Which student of John$_i$'s did he$_i$ visit?

       b. ??/* Mary introduced him$_i$ to every friend of John$_i$'s.

It has been noticed that adjuncts behave differently from arguments with respect to the interaction of A′-movement and Condition C. Whereas A′-moved arguments (as in (31)) behave as if interpreted at the base position, adjuncts do not. In (32b), the R-expression *John* is contained in an adjunct, and no condition C violation obtains. If *student* and its argument are said to obligatorily reconstruct in (32a), a uniform analysis would require that *student* in (32b) reconstruct obligatorily as well. As pointed out by Lebeaux (1990), the uniform analysis can be maintained, if it is further posited that adjuncts can be merged countercyclically, unlike arguments. If adjuncts are not present in the in-situ position of the *wh*-phrase, they cannot be reconstructed (they leave no lower copy which can be interpreted).[18]

(32)   a. ?? Which student of John$_i$'s did he$_i$ visit?

       b.   Which student that John$_i$ likes did he$_i$ visit?

The Condition C facts discussed above extend to degree clauses. It is expected that if the degree clause can be merged late, Condition C may be obviated in certain cases. But before we examine whether the predictions concerning the obviation of Condition C are borne out, it is useful to discuss the effect of an interfering factor: the preference for minimal attachment. In general, there seems to be a parsing preference that rules out string-vacuous semantically-vacuous dislocations. This can be seen in the contrast between (33a) and (33b) (from Taraldsen

(1981)).

(33) a. I gave him$_i$ a book yesterday [that John$_i$ liked].

b. ??I gave him$_i$ a book [that John$_i$ liked] (yesterday).

The absence of Condition C effects in (33a) can be taken to show the existence of an attachment site for the relative clause that is outside of the c-command domain of *him*. In principle, this high attachment site should be available to the relative clause in (33b) also. The oddness of (33b) suggests that such a site is either unavailable or at least not easily available. With Fox (2001), we take the unavailability (or relative inaccessibility) of a high attachment site in (33b) as reflecting a parsing preference for low attachment. A precise formulation of this preference is beyond the scope of this paper but it effectively mandates that when there are a number of semantically equivalent and string equivalent structures, the structure with the lowest attachment is the only one available. Thus this parsing preference does not rule out the semantically non-equivalent high attachment in *Mary saw the boy [with a telescope]*, but it does rule out high attachment of the relative clause in *I gave him a book [that John liked]*.

Degree clauses display similar minimal attachment effects. If the high attachment of the degree clause would be string-vacuous and would not yield a new scopal interpretation, the high attachment is inaccessible.[19] Hence we find Condition C effects in (34a). If the high attachment is made non-string-vacuous by for example the presence of an adverb as in (34b), Condition C is indeed obviated.

(34) a. ?? I will tell him$_i$ a sillier rumor (about Ann) than Mary told John$_i$.

b. I will tell him$_i$ a sillier rumor (about Ann) tomorrow than Mary told John$_i$.

In addition, there is a contrast between degree clauses and complements of nominals. Fox and Nissenbaum (1999) note that like the complements discussed earlier in this section, extraposition of complements of nominals does not obviate Condition C effects.

(35)   * I will tell him$_i$ a silly rumor tomorrow that Mary likes John$_i$.

They further note that the extraposition of adjuncts obviates Condition C but not the extraposition of complements.

(36)   a. I gave him$_i$ an argument yesterday [that supports John's$_i$ theory].

        b. ??/*I gave him$_i$ an argument yesterday [that John's$_i$ theory is correct].

Fox and Nissenbaum (1999) give the contrast in (36) an explanation similar to the explanation given by Lebeaux (1990) to the contrast in (32): the adjunct in (36a) can be merged late in its surface position (after QR of the DP *an argument* to the right) and hence it does not have to leave behind a copy that would trigger Condition C. (36b) involves a complement which must be merged early. Hence it must leave behind a copy and it is the presence of this copy that triggers Condition C effects.

We take the contrast between (34b) and (35) as demonstrating that late merge in its surface position is an option for degree clauses. Now, given that degree clauses are taken to be complements of the degree head, this outcome is perhaps surprising. Crucially, however, evidence that arguments may not be merged late has only been provided for arguments of lexical predicates. If our analysis is right, the argument/adjunct distinction is not the relevant one for the possibility of late merge. Likely, the necessity for theta-licensing forces arguments of lexical heads to be merged together with their predicates. Adjuncts and degree clauses need not be theta-licensed, and are thus free to come into the structure

late. As far as we are aware, the only other proposal that functional heads and their arguments may be introduced separately from one another has been made by Sportiche (see Sportiche (1997), Sportiche (1999)). According to his proposal, determiners and their NP arguments are merged independently: the NPs are merged in the theta-domain of predicates (e.g., verbs) whereas the determiners are merged later, in the inflectional domain associated with the predicate.

Another difference between the cases at hand involving late merge of degree clause complements and the cases of late merge of complements disallowed by Lebeaux (1990)/Fox and Nissenbaum (1999) is that degree clause complements are complements of the quantificational head that moves, while the familiar complement cases discussed by Lebeaux (1990)/Fox and Nissenbaum (1999) involve complements of restrictors of determiners. In §7, we will see that Fox (2002)'s mechanism for interpreting traces will be sensitive to this distinction. Late merge of complements of restrictors of determiners will yield an illegitimate object, while late merge of complements of determiners/-*er* will yield a semantically well-formed object.

5.2 Extraposition and Scope of -*er*

5.2.1 The Extraposition-Scope Generalization

The work of Williams (1974) established a correlation between the surface position of an extraposed clause and the scope of its source XP. The following is Williams (1974)'s statement of this generalization.

(37) If two scope items x and y with their determining clauses are represented in Deep Structure as: $[[x \text{ S}_1] \ldots [y \text{ S}_2]]$

and if extraposition yields the structure: $\ldots x \ldots y \ldots \text{S}_1 \text{ S}_2 \ldots$

then semantically $(y(x))$;

if it yields $\ldots x \ldots y \ldots S_2\, S_1 \ldots$

then semantically $(x(y))$; (from Williams (1974):194-195).

A similar correlation between the site of extraposition and the scope of the source of extraposition is implicit in the analysis of extraposition developed by Guéron and May (1984). They propose that at LF the source XP undergoes QR to a location where the source XP and the extraposed clause are both adjuncts of the same maximal projection. Most recently Fox and Nissenbaum (1999) and Fox (2002) have provided the following articulation of the extraposition-scope correlation.

(38) *Williams's generalization*:

When an adjunct β is extraposed from a "source DP" α, the scope of α is at least as high as the attachment site of β (the extraposition site). (from Fox (2002):ex. 19)

We will argue for a stronger version of the correlation between extraposition and scope than the one expressed in (38) at least as far as degree expressions are concerned.[20]

(39) *The Extraposition-Scope generalization* (for degree expressions):

When a degree clause β is extraposed from a degree head α, the scope of α is *exactly as high* as the merge site of β.

First we will illustrate the generalization as articulated in (38) i.e. show the *at least as high* part. Then we will show that the scope of the degree head is *exactly as high* as the degree clause. Finally an explanation for the generalization will be provided in §7.

The Extraposition-Scope Generalization is illustrated by the contrast between (40) and (41). (40) is ambiguous. The ambiguity is related to the scope of *every*

book. If *every book* undergoes short QR to the edge of the VP, it can serve as an antecedent for the ellipsis. This yields Reading 1 (= 40a), where I read the entire set of books before you read them. The QP *every book* can also be long-QR-ed to the edge of the IP, in which case it is not an antecedent for the ellipsis. This yields Reading 2 (= 40b), where for each book it is the case that I read it before you read that book.

(40) I read every book before you did (ambiguous)

    a. Reading 1: *before* > *every book*, the QP moves only to the edge of the VP.

        i. I $[_{VP}$ [every book]$_i$ $[_{VP}$ read t$_i$]] [before you did $[_{VP}$ Δ]]

        ii. I PAST $[_{VP}$ [every book]$_i$ $[_{VP}$ read t$_i$]]

                        [before you did $[_{VP}$ [every book]$_i$ $[_{VP}$ read t$_i$]]]

       At t$_1$, $\forall x$, book x, I read x, and at t$_2$, $\forall x$, book x, you read x, and t$_1$ < t$_2$

       [21]

    b. Reading 2: *every book* > *before*, the QP moves to the edge of the IP.

        i. $[_{IP}$ [every book]$_i$ $[_{IP}$ I $[_{VP}$ read t$_i$] [before you did $[_{VP}$ Δ]]]]

        ii. $[_{IP}$ [every book]$_i$ $[_{IP}$ I PAST $[_{VP}$ read t$_i$]

                        [before you did $[_{VP}$ read t$_i$]]]]

       $\forall x$, book x, I read x at t$_1$ and you read x at t$_2$ and t$_1$ < t$_2$

Now consider (41) and (42). They differ only in whether or not the relative clause has been 'extraposed' higher than the *before*-clause. (41), without 'extraposition', is ambiguous in exactly the same way as (40). When the QP undergoes short QR to the edge of the VP, we get Reading 1, and when it undergoes long QR to the edge of the IP, we get Reading 2. (42), where 'extraposition' has applied, is unambiguous, in contrast to (41).

(41) I read every book that John had recommended before you did.

    a. Reading 1: *before* > *every book*

    b. Reading 2: *every book* > *before*   (ambiguous)

(42) I read every book before you did that John had recommended.

    a. *Reading 1: *before* > *every book*

    b. Reading 2: *every book* > *before*   (unambiguous)

The absence of the reading corresponding to VP-scope for *every book* in (42) exemplifies the Extraposition-Scope Generalization. In (42), the relative clause has been extraposed beyond the *before* clause. The absence of the reading associated with narrow scope shows that the source DP *every book* is forced to take scope higher than the *before* clause.

Within Fox and Nissenbaum (1999) and Fox (2002)'s proposal, the Extraposition-Scope Generalization follows in the following manner. In (41), the relative clause is merged with the source DP in its base position. It can then undergo short (to the edge of the VP) or long (to the edge of the IP) QR, yielding the observed ambiguity. In contrast, in (42), the relative clause is merged late in a position that is higher than the attachment site of the *before*-clause, after the source DP has QR-ed to this position. It follows that only the reading corresponding to the LF where the source DP takes scope over the *before*-clause is available.

We will now demonstrate that the Extraposition-Scope Generalization holds for comparatives too. In particular, the surface position of the extraposed (=late-merged) degree clause marks the scope of the comparison. The Extraposition-Scope Generalization for comparatives receives a natural explanation if we allow for QR of the DegP headed by *-er*, followed by late merge of the degree clause with the degree head. Therefore we take the Extraposition-Scope Generalization

for comparatives as evidence for our proposal that degree clauses can be merged late.

We begin with a simple example where a comparative *-er* and a *before*-clause should, in principle, be able to enter into a scopal relationship. We see, however, that *-er* cannot scope over the *before*-clause. This is a manifestation of a wider phenomenon observed by Kennedy (1997), to which we will return later. Importantly for our present concerns, the impossibility for *-er* to scope over the *before*-clause, entails that the degree clause cannot be merged higher than the *before*-clause. This is so, because in our analysis, the degree clause is merged as a complement to *-er* only **after** *-er* has moved to its scope position. Thus, when the degree clause is extraposed higher than the *before*-clause, ungrammaticality should result. With this in mind, consider the case of (43) and (44). Example (43) is grammatical, but is not ambiguous. The only reading it has is the one where the *before*-clause has higher scope than the comparison, as in (43a). It is true in a situation where Mary and you are in a climbing competition, and Mary passes the 1000 ft mark before you pass it. The reading, resulting from *-er* having higher scope than the *before*-clause, as in (43b), is not available. If it were, (43) would be judged true in a situation where you were the first to pass the 1000 ft mark, as long as then Mary reached another height, say 1100 ft, before you. Intuitions are clear that (43) is false in such a situation. Given the unavailability of Reading 2 in (43b), the ungrammaticality of (44), where the degree clause has been extraposed, is significant.[22] According to our proposal, the degree clause is merged at the scope position of *-er*. Therefore, the only way for the degree clause to be extraposed higher than the *before*-clause is for *-er* to first QR over the *before*-clause. Yet, given the absence of Reading 2 of (43), this is not possible. Our analysis predicts a syntactic problem with the generation of the degree clause and indeed we find (44) to be ungrammatical.

(43) Mary climbed higher than 1000 feet before you did. (unambiguous)

    a. Reading 1: *before* > *-er*

        i. Mary $[_{VP}[_{VP}$ climbed $[_{AP}$ t_i high$]]$ $[_{DegP}$ *-er* [than 1000 ft]$]]$ [before you did $[_{VP}$ Δ $]]$

        ii. Mary PAST $[_{VP}[_{VP}$ climb $[_{AP}$ t_i high$]]$ $[_{DegP}$ *-er* [than 1000 ft]$]]$ [before you did $[_{VP}[_{VP}$ climb $[_{AP}$ t_i high$]]$ $[_{DegP}$ *-er* [than 1000 ft]$]]]$

        The earliest time t_1 [max d [Mary climbed to d at t_1] $d > 1000$ feet] the earliest time t_2 [max d [you climbed to d at t_2] $d > 1000$ feet] $t_1 < t_2$

    b. *Reading 2: *-er* > *before*, the Heim-Kennedy Constraint

        i. $[_{IP}$ $[_{IP}$ Mary $[_{VP}$ climbed $[_{AP}$ t_i high$]]$ [before you did $[_{VP}$ Δ $]]]$ $[_{DegP}$ *-er* [than 1000 ft]$]$

        ii. $[_{IP}$ $[_{IP}$ Mary PAST $[_{VP}$ climb $[_{AP}$ t_i high$]]$ [before you did $[_{VP}$ climb $[_{AP}$ t_i high$]]]$ $[_{DegP}$ *-er* [than 1000 ft]$]]$

        max d [the earliest time t_1 [Mary climbed to d at t_1] the earliest time t_2 [you climbed to d at t_2] $t_1 < t_2$] $d > 1000$ feet

(44) *Mary climbed higher before you did than 1000 feet.

    a. *Reading 1: *before* > *-er*, position of degree clause presupposes high scope for *-er*

    b. *Reading 2: *-er* > *before*, the Heim-Kennedy Constraint

Now that we have illustrated the interaction between *-er* and a *before*-clause, and its implications for the licensing of an extraposed degree clause, we can look at a more complicated example which provides further support to the proposal that comparatives involve *-er*-movement and that the degree clause is merged in the scope position of *-er*. As will become clear, a comparative DP (e.g., *more books*

than X) can scope either above or below a *before*-clause. If the degree clause is extraposed higher than the *before*-clause, only a high scope reading is available for the comparative DP, similarly to the case of (46). Furthermore, the *before*-clause cannot intervene between the scope of *-er*, as marked by the position of the degree clause, and the degree predicate containing the base position of *-er*, in conformity with the Heim-Kennedy Constraint.

Consider the contrast between (45) and (46). These sentences are structurally identical to (40), except that they contain the weak DP *more books than Mary published in her life*.[23] Example (46) involves extraposition of the degree clause higher than the *before*-clause. Correspondingly, this example lacks the reading of (45) where the comparative DP has scope lower than the *before*-clause (Reading 1).

(45) John read more books than Mary published in her life before you did.

    a. Reading 1: *before* > *-er ... d-many books*

    b. Reading 2: *-er ... d-many books* > *before*

    c. *Reading 3: *-er* > *before* > *d-many books*, the Heim-Kennedy Constraint

(46) John read more books before you did than Mary published in her life.

    a. *Reading 1: *before* > *-er ... d-many books*

    b. Reading 2: *-er ... d-many books* > *before*

    c. *Reading 3: *-er* > *before* > *d-many books*, the Heim-Kennedy Constraint

The LFs behind Readings 1 and 2 are illustrated below:

(47) a. Reading 1 (available in (45, unavailable in (46)):

      before > *-er ... d-many books*

      i. J. [$_{VP}$read [[t$_i$ many books] [*-er* [than M. published in her life]]$_i$]] [before you did [$_{VP}$ Δ]].

ii. J. PAST [$_{VP}$read [[t$_i$ many books] [-*er* [than M. published in her life]]$_i$]]

[before you did [$_{VP}$ read [[t$_i$ many books] [-*er* [than M. published in her life]]$_i$]]].

John read more books than Mary published in her life before you read more books than Mary published in her life.

b. Reading 2 (available in both (45) and (46)):

-er ... d-many books > before[24]

i. J. [[$_{VP}$ read t$_i$] [before you did [$_{VP}$ Δ]]]

[[t$_j$ many books] [-*er* [than M. published in her life]]$_j$]$_i$

ii. J. [[$_{VP}$ read t$_i$] [before you did [$_{VP}$ read t$_i$]]]

[[t$_j$ many books] [-*er* [than M. published in her life]]$_j$]$_i$

The number of books that John read before you read them exceeds the number of books that Mary published in her life.

In contrast to (45), (46) is unambiguous. It only has the reading shown in (47b), where the comparative DP scopes over the *before*-clause. This contrast exemplifies the Extraposition-Scope Generalization. In our account of it, in (45), the degree clause is merged low within the source DP that contains the degree head it is associated with. The whole DP can take scope with respect to the *before*-clause. If it scopes below *before*, we have the reading in (47a), and if it takes scope over *before*, we have the reading in (47b). However, in (46), the degree clause is merged late at a position higher than the *before* clause. By our assumptions, this means that the degree head is in a position above *before*. Yet, given the Heim-Kennedy Constraint, the *before*-clause cannot intervene between the degree quantifier (the degree head and the degree clause) and the degree predicate (*d-many books*). Therefore, the whole DP, containing the degree quantifier must have scope higher than the *before*- clause. Consequently, the reading that obtains is the one given in (47b). Reading 1, with a low scope for the DP, is not available.

The above discussion illustrated that the scope of *-er* is at least as high as the surface position of the degree clause. We now show that the scope of *-er* is *exactly* as high as the surface position of the degree clause. The initial evidence comes from structures that involve two extrapositions. To set the scene consider (48).[25]

(48)   a. So many people ate faster yesterday [than we had expected] [that we were all done by 9pm].

       b. *So many people ate faster yesterday [that we were all done by 9pm] [than we had expected].

(48b) is ungrammatical due to a constraint that result clauses must follow comparative clauses. We do not understand the exact nature of this constraint but see Williams (1974):206 and Guéron and May (1984):29 for discussion.

In (48a), the two degree abstractions do not have to cross.

(49)   Degree abstractions in the LF of (48a):

so [that....] λd_1 [[d_1-many people] λx [-er [than....] λd_2 [x ate d_2 fast]]]

If we consider a structure like (50) where the two degree abstractions cross, we get ungrammaticality irrespective of the relative order of the *than*-clause and the result clause.

(50)   a. ???/*More people ate so fast yesterday [than we had expected] [that we were all done by 9pm].

       b. *More people ate so fast yesterday [that we were all done by 9pm] [than we had expected].

The ungrammaticality of (50b) follows from the ordering constraint that holds between *than*-clauses and result clauses. We take the ungrammaticality of (50a) to follow from the fact that (50a) involves crossing degree abstractions, something we take is ruled out by the Heim-Kennedy constraint.

(51) Degree abstractions in the LF of (50a):

  so [that....] λd_1 [-er [than....] λd_2 [d_2-many people ate d_1 fast]]]

Further QR of *-er* with its *than*-clause does not help in rescuing the structure. We get the structure in (52).

(52) -er [than....] λd_3 [so [that....] λd_1 [d_3 λd_2 [d_2-many people ate d_1 fast]]]

Like (51), (52) also involves crossing degree abstractions. The only difference is that in (51) the abstraction associated with *so* crosses over *-er*, while in (52) , the abstraction associated with *-er* crosses over *so.* In these examples, the surface position of the extraposed degree clauses marks exactly the scope of the corresponding degree heads. However, these examples involve a confound in that further covert movement of the degree quantifier creates a violation of the Heim-Kennedy constraint. We need a case where potential further movement of the degree quantifier would result in a legitimate LF and then show that the corresponding reading is unavailable.

Such a case is provided by the following examples. They involve an intensional predicate and a degree quantifier such that the two LFs corresponding to the two possible scopes of the intensional predicare and the degree quantifier are semantically distinct (see §4.2 and Heim (2000) for details).

(53) a. Degree clause is in the embedded clause:

  John is required [to publish fewer papers this year [than that number] in a major journal] [to get tenure].

  Simplified LF: *required* > [*fewer* [than n]]

  required [fewer [than n] λd [PRO publish d-many papers]]

  b. Degree clause is outside the matrix clause:

  John is required [to publish fewer papers this year in a major journal]

39

[to get tenure] [than that number].

Simplified LF: [*fewer* [than n]] > *required*

fewer [than n] λd [*required* [PRO publish d-many papers]]

(54)   a. Degree clause is in the embedded clause:

John is required [to publish exactly five more papers this year [than that number] in a major journal] [to get tenure].

Simplified LF: *required* > [*exactly 5 more* [than n]]

required [[exactly 5 -er] [than n] λd [PRO publish d-many papers]]

   b. Degree clause is outside the embedded clause:

John is required [to publish exactly five more papers this year in a major journal] [to get tenure] [than that number].

Simplified LF: [*exactly 5 more* [than n]] > *required*

[exactly 5 -er] [than n] λd [*required* [PRO publish d-many papers]]

The structures where the degree clause is in the embedded clause have only the reading corresponding to *required* taking scope over the comparison (cf. 53, 54a). This is why (53a) has the pragmatically odd reading that if John publishes more than a certain number of articles, he will not get tenure. Likewise when the degree clause is outside the embedded clause, only the reading corresponding to the comparison taking scope over *required* is available. This is the pragmatically reasonable reading which limits from above the minimum number of articles John needs to publish to get tenure. It does not penalize John for being productive.

    The availability of the *-er* > *required* reading in the (53/54b) examples shows that the structure involving a degree abstraction that crosses *required* is semantically well-formed. The absence of this reading in (53/54a) shows that the scope of *-er* is marked exactly by the surface position of the degree clause i.e. the degree quantifier in (53/54a) cannot move further. [26]

That the Extraposition-Scope Generalization obtains in comparatives provides, we believe, strong support for our proposal that extraposed degree clauses are overt indicators of the scope of the comparison and that they are merged late after QR of the degree head.

5.2.2 Extraposition and Russell's Ambiguity

Russell's ambiguity (cf. Russell (1905)) is exhibited by sentences such as (55). On one reading, sentence (55) is true if John is mistaken about Mary's actual height, e.g., he thinks she is 6 feet tall, whereas in fact she is 5'8". On another reading, John has a contradictory thought, namely that Mary's height exceeds itself.

(55)   John thinks that Mary is taller than she is.

It is sometimes claimed that Russell's ambiguity is purely a matter of scopal interaction between the comparison and the attitude verb. Under the scope approach, the sensible reading is to be represented as, e.g., *The degree to which John thinks that Mary is tall is greater than the degree to which Mary is tall*, with the comparison outside the scope of *think*; whereas the contradictory reading is to be represented as *John thinks that the degree to which Mary is tall is greater than the degree to which Mary is tall*, with the comparison within the scope of *think* (e.g., Postal (1974), Cresswell (1976), Hellan (1981)). As pointed out by von Stechow (1984), Russell's ambiguity can be handled in terms of a de-dicto/de-re distinction. On the de-re reading, the comparison can still be within the scope of *think* as long as the degree clause is interpreted with respect to the actual world (which von Stechow represents as the degree clause itself having a scope outside of *think*, with *-er* within the scope of *think*). Such a non-scopal approach to Russell's ambiguity is required in cases like (56).

(56)   If Mary smoked less than she does, she would be healthier.

41

Scoping the DegP out of the *if*-clause is problematic because *if*-clauses are strong islands. In addition, such a movement yields incorrect truth conditions. von Stechow (1984)'s solution is related to an idea that we have discussed before: *de re* interpretations are compatible with wide scope, as well as narrow scope, with respect to an intensional operator. *de dicto* interpretations, however, require narrow scope.[27]

Let us now consider the interaction of the Russell's ambiguity cases with extraposition (cf. 57). (57a) displays the familiar Russell's ambiguity. This ambiguity remains even when the embedded clause is extraposed higher than *tomorrow*, as in (57b). Yet, when just the degree clause is extraposed, as in (57c), only the 'sensible' *de re* reading that corresponds to a mistaken but coherent claim is available.

(57) a. John will claim that Mary is taller than she is tomorrow. (incoherent claim; coherent mistaken claim)

  b. John will claim tomorrow that Mary is taller than she is. (incoherent claim; coherent mistaken claim)

  c. ?John will claim that Mary is taller tomorrow than she is. (coherent mistaken claim)

Let us look closer into the syntax behind the possible readings in (57a-c). The incoherent claim reading of (57a, b) corresponds to a narrow scope *de dicto* interpretation of the DegP while the coherent mistaken reading corresponds to a *de re* interpretation of the DegP. Following the discussion in the preceding sections, we cannot tell from the interpretation whether the *de re* interpretation involves narrow scope of the DegP or wide scope. From the existence of the *de dicto* reading, which has to be narrow scope, we know that narrow scope is a possibility and therefore we can conclude that a narrow scope *de re* reading is also available.

42

We return to the question of whether (57a, b) have an additional wide scope *de re* reading after addressing (57c).

Given that the degree clause in (57c) is merged in the scope position of *-er*, and given that in the surface it appears extraposed higher than *tomorrow*, it must be the case that *-er* has QR-ed to the matrix clause. Predictably, only the high scope *de-re* reading is available. The derivation is shown below:

(58)   John will claim that Mary is taller tomorrow than she is.

    a. 'John will claim that Mary is taller tomorrow' is generated.

    b. QR of *-er* takes place to the matrix clause. The lower copy is pronounced and so this step is 'covert.'

        Overt structure: [[John will claim that Mary is taller tomorrow] *-er*]

        LF: [[λd John will claim that Mary is d-tall tomorrow] *-er*]

    c. The degree complement is merged with *-er*

        Overt structure: [[John will claim that Mary is taller tomorrow] [*-er* [than she is]]]

        LF: [[λd John will claim that Mary is d-tall tomorrow] [*-er* [than λd she is d-tall]]]

The fact that extraposition of just the degree clause (as opposed to the whole embedded clause) leads to a loss of the low scope readings of Russell's ambiguity is directly predicted by our proposal.

Most speakers find (57c) to be degraded in comparison to (57a, b). We believe that the oddness of 57c) is related to the fact that it involves QR of a DegP out of a finite clause. QR out of finite clauses is either unavailable or marked and this seems to also be the case with DegPs. In general, the derivation of a wide scope *de re* structure in cases like (57a-c) will involve QR out of a finite clause. (57a, b) have alternate derivations which do not involve QR out of a finite clause. In contrast,

the only derivation available to (57c) involve QR out of a finite clause. This is why an acceptability contrast is perceived between (57a-b) on the one hand and (57c) on the other.

To conclude this section, we have seen several cases where extraposition of the degree clause constrains interpretation. If an R-expression in the degree clause has an antecedent pronoun in the main clause c-commanding the overt degree head and degree predicate, a Condition C violation may be avoided, if the degree clause is extraposed so that its surface position is outside of the c-command domain of the pronoun (see section 5.1). This, together with the fact that A′-movement of arguments does not rescue Condition C violations, suggests that the degree clause does not appear in its surface position as a result of A′-movement, but rather that it is merged there. Extraposition of the degree clause also marks the scope of *-er* - with respect to quantificational adverbials (see section 5.2.1). Thus extraposition of the degree clause higher than an intensional verb blocks low scope readings of the DegP with respect to the intensional verb (see section 5.2.2).

6   Evidence from Ellipsis Resolution

According to our proposal, the surface position of the degree clause functions as a marker of the scope of the DegP headed by *-er*. We will now show that the Sag-Williams Ellipsis-Scope Generalization follows from our proposal (§6.1) as does a related correlation between Condition C and scope (§6.2). Finally in §6.3, cases of trapped ACD that are problematic for alternative analyses will be discussed and it will be shown that our proposal makes the right predictions for these.

6.1 The Ellipsis Scope Generalization

Sag (1976) and Williams (1974) noted that there is a correlation between ellipsis and scope. This correlation has since come to be known as the Sag-Williams Ellipsis-Scope Generalization.

(59) Ellipsis-Scope Generalization: The scope of a DegP containing elided material must contain the antecedent of the ellipsis.

(60) illustrates the Ellipsis-Scope Generalization. The reading in (60b) is missing because the scope of the comparison is contained inside the antecedent of the ellipsis (= *tell her to work hard*).[28]

(60) Mary's father tells her to work harder than her boss does.

    a. *tell* > *-er*, elided material = work d-hard

       Mary's father tells Mary: work harder than your boss works.

    b. * *tell* > *-er*, elided material = tell her to work d-hard

       Mary's father tells Mary: work harder than your boss tells you to work.

    c. *-er* > *tell*, elided material = work d-hard

       Mary's father tells Mary: work d_1-hard;

       Mary's boss works d_2-hard; $d_1 > d_2$.

    d. *-er* > *tell*, elided material = tell her to work d-hard

       Mary's father tells Mary: work d_1-hard;

       Mary's boss tells Mary: work d_2-hard; $d_1 > d_2$.

If the problem of antecedent containment is resolved by decreasing the size of the ellipsis, the missing reading becomes available (cf. 61a).

(61) Mary's father tells her to work harder than her boss tells her to.

a. *tell* > *-er*, elided material = work d-hard

Mary's father tells Mary: work harder than your boss tells you to work.

b. *-er* > *tell*, elided material = work d-hard

Mary's father tells Mary: work d_1-hard; Mary's boss tells Mary: work d_2-hard; $d_1 > d_2$.

The fact that once the problem of antecedent containment is resolved, the reading missing (= 60b) becomes available (= 61a) buttresses the Sag-Williams Ellipsis-Scope Generalization. The reading isn't missing because of potential pragmatic or other ill-formedness. It is missing due to a failure of ellipsis resolution.

Our analysis of the Ellipsis Scope Generalization follows essentially that of Heim (2000) - the difference being that for her the degree clause and the degree head are merged together in the base position of *-er* and move together as a unit at LF. (62) demonstrates how the Sag-Williams Ellipsis-Scope Generalization follows from our analysis. Following Fox & Nissenbaum (1999), we assume that QR takes place to the right. In (62a, b), we first generate the embedded clause *her to work harder*. This is followed to QR of *-er* to the right, followed by the merging of the Degree Clause as the complement of *-er*. The resulting syntactic object is then combined with the remaining lexical items to yields the structures in (62a, b). Choosing the embedded VP to resolve the VP ellipsis in the Degree Clause yields yields the LF in (62a). The reading indicated by the LF in (62b) is unavailable because there is no VP in the structure that can license the required ellipsis resolution.

(62) Mary's father tells her to work harder than her boss does.

a. *tell* > *-er*, elided material = work $d-$hard

Degree Clause is merged low, at the level of the embedded clause

46

LF: Mary's father tells her [[λd PRO to work d-hard]

[-er [λd her boss *works* d-*hard*]]]

b. * *tell* > *-er*, elided material = tell her to work $d-$hard

Degree Clause is merged low, at the level of the embedded clause

LF: Mary's father tells her [[λd PRO to work d-hard]

[-er [λd her boss *tells her to work* d-*hard*]]]

c. *-er* > *tell*, elided material = work $d-$hard

Degree Clause is merged high, at the level of the matrix clause

LF: [[λd Mary's father tells her to work d-hard]]

[-er [λd her boss *works* d-*hard*]]

d. *-er* > *tell*, elided material = tell her to work $d-$hard

Degree Clause is merged high, at the level of the matrix clause

LF: [[λd Mary's father tells her to work d-hard]]

[-er [λd her boss *tells her to work* d-*hard*]]

In (62c, d), we first generate *Mary's father tells her to work hard*, followed by QR of *-er* to the matrix clause. The Degree Clause is now merged as a complement of *-er*. Depending upon which VP we take as the antecedent of the elliptical VP, we get (62c) (= embedded VP), or (62d) (= matrix VP).

6.2 Correlations with Condition C

Further examination of the examples used to motivate the Sag-Williams Ellipsis-Scope Generalization reveals that there also exists a correlation between coreference judgements and the scope of the comparison. A pronoun that c-commands the embedded VP can corefer with a name in the Degree Clause only if the comparison takes scope over the matrix predicate. This correlation is demonstrated by (63).

47

(63) Her father tells her$_i$ to work harder than Mary$_i$'s boss does.

    a. *tell $>$ *-er*, elided material = work d-hard, coreference impossible

    b. * tell $>$ *-er*, elided material = tells her to work d-hard, ruled out by conditions on ellipsis resolution

    c. *-er* $>$ tell, elided material = work d-hard, coreference possible

    d. *-er* $>$ tell, elided material = tells her to work d-hard, coreference possible

As discussed in the previous section, the reading in (63b) is missing due to a failure of ellipsis resolution. If we remove the problem of ellipsis resolution, we see that the correlation found between the scope of comparison and the possibility of pronominal coreference that we see in (63) reemerges.

(64)   a. Her father tells her$_i$ to work harder than Mary's$_i$ boss tells her to.

       i. * tell $>$ *-er*, coreference impossible

       ii. *-er* $>$ tell, coreference possible

    b. Her father tells Mary$_i$ to work harder than her$_i$ boss tells her to.

       i. tell $>$ *-er*, coreference possible

       ii. *-er* $>$ tell, coreference possible

We see that the only reading available for (64a) is the reading where the comparison scopes over the matrix predicate. In contrast both readings are available in (64b).

In Heim (2000)'s analysis, the impossibility of coreference in (63a) and the missing reading in (64a.i) would reduce to a Condition C violation. Condition C is taken to operate at LF and coreference between a pronoun that c-commands the embedded VP and a name is the degree clause is only possible if the Degree Head and the Degree Clause are QR'd out of the c-command domain of the relevant pronoun. This is shown in (65).

(65) Her father tells her$_i$ to work harder than Mary$_i$'s boss does.

    a. * tell > *-er*, elided material = work d-hard, coreference impossible, Condition C violation

        LF: Her father tells her [[-er [λd Mary's boss *works* d-*hard*]]

                [λd PRO to work d-hard]]

    b. * tell > *-er*, elided material = tells her to work d-hard, ruled out by conditions on ellipsis resolution

        LF: Her father tells her [[-er [λd Mary's boss *tells her to work* d-*hard*]]

                [λd PRO to work d-hard]]

    c. *-er* > tell, elided material = work d-hard, coreference possible, no Condition C violation

        LF: [[-er [λd Mary's boss *works* d-*hard*]]

                [λd Her father tells her to work d-hard]]

    d. *-er* > tell, elided material = tells her to work d-hard, coreference possible, no Condition C violation

        LF: [[-er [λd Mary's boss *tells her to work* d-*hard*]]

                [λ Her father tells her to work d-hard]]

At this point, it is worth noting that QR by itself does not bleed Condition C. This is shown by (66).

(66) (from Fox (2000), pg. 175)

    a. ??/*You bought him$_i$ every picture that John$_i$ liked.

    b. *He$_i$ bought you every picture that John$_i$ liked.

Fiengo & May (1994) and Fox (2000) note that there is an environment where QR does bleed Condition C - when QR resolves antecedent containment.

(67) (from Fox (2000), pg. 180)

a. You sent him$_i$ the letter that John$_i$ expected you would.

b. You introduced him$_i$ to everyone that John's$_i$ mother did.

c. You reported him$_i$ to every cop that John$_i$ was afraid you would.

Since the environments discussed by Heim involve antecedent containment, it makes sense that the QR there bleeds Condition C. However, we find that Condition C is bled even when there is no problem of antecedent containment as in (68=64).

(68) Her father tells her$_i$ to work harder than Mary's$_i$ boss tells her to.

    a. * tell $>$ *-er*, coreference impossible, Condition C violation:

       LF: Her father tells her$_i$ [[-er [λd Mary's boss tells her to *work d-hard*]]

                [λd PRO to work d-hard]]

    b. *-er* $>$ tell, coreference possible, no Condition C violation:

       LF: [[-er [λd Mary's boss tells her to *work d-hard*]]

              [λd her father tells her to work d-hard]]

In (68) we find QR bleeding Condition C even though there is no antecedent containment to be resolved. Given what we know about QR, this is a puzzle. [29]

The Condition C-Scope Generalization follows from our proposal. The merge location of the Degree Clause marks the scope of the comparison. The Condition C effects merely reflect the structural location of the Degree Clause. This is shown in (69). As before, we assume that the Degree Head QR's to the right.

(69) Her father tells her$_i$ to work harder than Mary$_i$'s boss does.

    a. * tell $>$ *-er*, elided material = work d-hard, coreference impossible, Condition C violation

       Degree Clause is merged low, at the level of the embedded clause

LF: Her father tells her [[λd PRO to work d-hard]

[-er [λd Mary's boss *works d-hard*]]]

b. * tell > *-er*, elided material = tells her to work d-hard, ruled out by conditions on ellipsis resolution

Degree Clause is merged low, at the level of the embedded clause

LF: Her father tells her [[λd PRO to work d-hard]

[-er [λd Mary's boss *tells her to work d-hard*]]]

c. *-er* > tell, elided material = work d-hard, coreference possible, no Condition C violation

Degree Clause is merged high, at the level of the matrix clause

LF: [[λd Her father tells her to work d-hard]

[-er [λd Mary's boss *works d-hard*]]]

d. *-er* > tell, elided material = tells her to work d-hard, coreference possible, no Condition C violation

Degree Clause is merged high, at the level of the matrix clause

LF: [[λd Her father tells her to work d-hard]

[-er [λd Mary's boss *tells her to work d-hard*]]]

QR is not responsible for moving the Degree Clause out of the c-command domain of the relevant pronoun. In the readings where there is no Condition C violation, the Degree Clause is just merged high, out of the c-command domain of the relevant pronoun. The only element moved by QR is the Degree Head. Since QR is not responsible for removing the offending Degree Clause out of the c-command domain of the relevant pronoun, it is no longer a puzzle why the bleeding of Condition C is not dependent on the presence of antecedent containment. Our analysis of (68) is shown in (70).

(70)   Her father tells her$_i$ to work harder than Mary's$_i$ boss tells her to.

a. * tell > -*er*, coreference impossible, Condition C violation:

Degree Clause is merged low, at the level of the embedded clause

LF: Her father tells her$_i$ [[λd PRO to work d-hard]

[-er [λd Mary's boss tells her to *work d-hard*]]]

b. -*er* > tell, coreference possible, no Condition C violation:

Degree Clause is merged high, at the level of the matrix clause

LF: [[λd Her father tells her to work d-hard]

[-er [λd Mary's boss tells her to *work d-hard*]]]

6.3 Trapped Ellipsis

In all the cases that we have discussed so far, the degree clause was right peripheral. It was therefore plausible to assume that the degree clause could be merged at a variety of positions on the right frontier of the tree yielding the same right peripheral positions. In these right peripheral positions, there is no antecedent containment and therefore if the degree clause contains an ellipsis site, the ellipsis in the degree clause can be resolved.

However, as (71) shows degree clauses can also appear in positions that are not right peripheral.

(71) a. John gave more books [than he read last summer] to Mary.

b. John gives more people [than you have ever met] expensive presents.

c. John wanted more people [than I have ever met] to come to the party.

Further, such degree clauses can involve ellipsis (cf. 72).

(72) a. John gave more presents [than Bill did] to Mary.

b. John gave more teachers [than Bill did] presents.

c. John wanted more people [than I did] to come to the party.

We refer to the examples in (71) as involving 'trapped' degree clauses and to the examples in (71) as involving 'trapped ellipsis'. Cases like (71) and (72) raise potential problems for our proposal. So far we have always merged the degree clause with the degree head after the covert movement of the degree head. In other words, the degree clause has always been merged in a scope position. However, this is not forced by our analysis. If we assume, as is plausible, that there are DP-internal scope positions (see §8.5 of Heim and Kratzer (1998)), then it is also possible to merge the degree clause more locally within the DP after DP-internal QR of the degree head as suggested by the structure in (73).

(73) John gave [[-er$_i$ books] [-*er*$_i$ [than he read last summer]]] to Mary.

The structure in (73) takes care of (71). As for the problem of antecedent containment in (72), we could adopt Heim (2000)'s analysis of degree constructions and use QR as she does to resolve antecedent containment. However, for reasons that will become evident in the course of this discussion, we believe that this is the wrong way to go. Instead, the derivation of the examples in (72) involves a different strategy. This strategy involves rightward movement of the material to the right of the degree clause. That rightward movement of the relevant constituent is possible is shown by (74).

(74) (from Fox (2002))

  a. I gave a book yesterday to Mary.

  b. I gave everyone yesterday a book about linguistics.

  c. I wanted this man with all my heart to come to the party.

Our basic proposal concerning the analysis of (72) is that the Degree Clauses are, in fact, attached to the right frontier of the tree. It is just that there has been an additional rightward movement of syntactic material over the Degree Clause.

The derivations for (72) are indicated in (75).

(75) a. John gave more presents [than Bill did] to Mary.

    i. John gave more presents to Mary.

    ii. QR of Degree Head:

       [[John gave more presents to Mary] *more*].

    iii. Merge of Degree Clause:

       [[John gave more presents to Mary]

             [*more* [than Bill did]]]

    iv. Rightward Movement of *to Mary*:

       [[[John gave more presents *to Mary*]

             [*more* [than Bill did]]] [to Mary]]

  b. John gave more teachers [than Bill did] presents.

    i. John gave more teachers presents.

    ii. QR of Degree Head:

       [[John gave more teachers presents] *more*]

    iii. Merge of Degree Clause:

       [[John gave more teachers presents]

             [*more* [than Bill did]]]

    iv. Rightward Movement of *presents*:

       [[[John gave more teachers *presents*]

             [*more* [than Bill did]]] [presents]]

  c. John wanted more people [than I did] to come to the party.

    i. John wanted more people to come to the party.

    ii. QR of Degree Head:

       [[John wanted more people to come to the party] *more*]

iii. Merge of Degree Clause:

[[John wanted more people to come to the party]

[*more* [than I did]]]

iv. Rightward Movement of *to come to the party*:

[[[John wanted more people [*to come to the party*]]

[*more* [than I did]]] [to come to the party]]

(words in italics are unpronounced copies)

The derivations in (75) do not involve antecedent containment. If we adopt them, then the problem of antecedent containment raised by (72) is only an apparent problem. This means that we do not need to assume that the Degree Head and the Degree Clause QR together to resolve antecedent containment.

The examples in (71) do not involve even apparent antecedent containment and hence derivations like the ones offered in (75) are not forced for them. Such derivations are, however, still available. It is also worth noting that in certain cases there may be more than one way of resolving antecedent containment. In (75c), for example, if we assume that ECM subjects can overtly object shift to a position above the position where the ECM assigner is interpreted, then an alternate derivation that does not involve rightward movement becomes available. Under this derivation, the ECM subject moves to a position where it is not inside a minimal VP headed by *want.* The comparative clause containing ellipsis is introduced in this position. Since there is no antecedent containment in this position, the ellipsis can be resolved without any associated rightward movement. Note that this derivation would involve adjoining the comparative clause directly to a Deg head adjoined to a DP and not, as has been the case elsewhere, a Deg head adjoined to a clausal projection.

Our analysis of trapped ellipsis, as in (72), crucially depends upon the avail-

ability of rightward movement to the syntactic material to the right of the trapped ellipsis. Thus our analysis makes a clear prediction: if rightward movement is not an option, it should not be possible to resolve trapped ellipsis, and we should have ungrammaticality.

Fox (2001) notes that one environment where rightward movement is not permitted involves finite VP's. Embedded finite VP's are unable to rightward move out of a VP that dominates it. The evidence for the unavailability of this movement comes from the fact that matrix adverbs cannot intervene between the embedded finite VP and its subject.

(76)   (from Fox (2001))

    a. *I hope that Bill [with all my heart] will come to the party.

    b. * I told you that Bill [when we met] will come to the party.

The prediction then is that if the Degree Clause containing the trapped ellipsis is followed by a finite VP, the trapped ellipsis will not be resolved. As (77) shows, this prediction is borne out.

(77)   *Mary desires that more people than Bill does take syntax.

Since it is not possible to rightward move the embedded finite VP *take syntax* out of the matrix finite VP, the Degree Clause is really, not just apparently, trapped. Consequently, there is no way to resolve the ellipsis.

Note that it is not possible to use the embedded VP as a potential antecedent for the ellipsis in the Degree Clause. This is so because the VP *take syntax* does not contain any degree variable. The resulting LF shown in (78) would be illegitimate as it would involve vacuous quantification.

(78)   LF: *Mary desires that [[-er [λd Bill *take syntax*]]

                  [λd d-many people take syntax]]

The only suitable antecedent would be something like *desire that d-many people take syntax*. However, such an antecedent is not available to (77). If there is no 'trapped ellipsis', then such an antecedent becomes available and the corresponding sentence in (79) becomes grammatical.

(79)   Mary desires that more people take syntax than Bill does.

    a. Mary desires that more people take syntax.

    b. QR of *more*:

       [[Mary desires that more people take syntax] *more*]

    c. Merge of Degree Clause:

       [[Mary desires that more people]

               [*more* [than Bill *desires people take syntax*]]]

    d. LF: [[λd Mary desires that d-many people]

               [*-er* [λd Bill *desires d-many people take syntax*]]]

(79) demonstrates that the problem with (77) is not related to the scope possibilities available to the Degree Head in (77). Rather, the problem is one of ellipsis resolution and once that is taken care of as in (79), we have grammaticality.

The contrast between the ungrammatical (77) and the grammatical (79) is instructive because it points out that QR alone cannot resolve antecedent containment (see Fox (2002) for details). A theory like Heim (2000) where QR is the mechanism for resolving antecedent containment makes the wrong predictions here. Heim's system would involve reconstructing the extraposed Degree Clause into the complement position of the Degree Head, yielding essentially (77). Under this account the contrast between (77) and (79) is truly a mystery. Given that QR is indeed available in (79), it would be stipulative to rule it out in (77).

7 Late Merge and the Non-Conservativity of Degree Quantifiers

So far we have provided evidence that degree clauses are merged late as complements to degree heads after the degree heads have undergone QR. We have also shown that the surface position of the degree clause marks exactly the semantic scope of the degree head. Up until now we have not answered the following questions. Why can the degree clause not be merged with *-er* in *-er*'s base position - in other words, why must the degree clause be merged late? A related question is why the degree clause, a complement, can be merged late given the argument-adjunction distinction discussed by Lebeaux (1990) and Fox and Nissenbaum (1999). Finally, where does the 'exactly as high' part of the Extraposition-Scope Generalization come from? We have seen that the 'at least as high' generalization follows from the mechanism of countercyclic merger developed by Fox and Nissenbaum (1999)/Fox (2002), but no comparable explanation has been provided so far for the 'exactly as high' part.

In what follows we will show that answers to all of these questions follow from the mechanism by which movement is interpreted and the semantic properties of comparative quantifiers, in particular their non-conservativity.

7.1 Background: The Interpretation of Copies

The copy theory of movement (cf. Chomsky (1993)) raises questions about how lower links of chains are interpreted. To this effect, Fox (2001)/Fox (2002) proposes the mechanism of Trace Conversion.

(80) Trace Conversion (from Fox (2001), Fox (2002))

    a. Variable Insertion:

(Det) Pred $\rightarrow$ (Det) [Pred λy $(y = x)$]

b. Determiner Replacement:

(Det) [Pred λy $(y = x)$] $\rightarrow$ the [Pred λy $(y = x)$]

The mechanism of Trace Conversion creates interpretable LF objects out of movement structures involving copies. This mechanism has two components: Variable Insertion and Determiner Replacement. Together they convert the copy of a DP (= [D NP]) to a definite description with a free variable in it. This free variable is bound by a higher copy. The process of Trace Conversion is exemplified in (81).

(81) a. Mary talked to every boy.

[every boy] Mary talked to [every boy]. $\rightarrow$

[every boy] λx [Mary talked to the [boy $\lambda y(y = x)$]. $\rightarrow$

[every boy] λx [Mary talked to the [λy [boy$(y) \wedge y = x$]]. $\rightarrow$

[every boy] λx [Mary talked to **the boy** x]

b. Which boy did Mary visit?

[which boy] Mary visited [which boy]?

[which boy] λx [Mary visited **the boy** x]?

Fox (2001) notes that the LFs in (81a) yield the expected meaning because of the conservativity of natural language quantifiers.

(82) $Q(A, B) =$

$= Q(A)(A \cap B)$ (by conservativity)

$= Q(A)(A \cap [\lambda x : A(x).B(x)])$ (by assumptions about Presupposition Projection)

$= Q(A)([\lambda x : A(x).B(x)])$ (by conservativity)

$= Q(A)(B(\textbf{the}[A\ x]))$

(from Fox (2001))

He points out that non-conservative quantifiers, given Trace Conversion, could only have trivial meanings.

(83)   a.   Only Norwegians are tall. ($\rightarrow$, QR)

     b.   [Only Norwegians] [[Only Norwegians] are tall] ($\rightarrow$, Trace Conversion)

     c.   [Only Norwegians] λx [**the Norwegian** x are tall]

     d.   Only Norwegians are such that those Norwegians that are all. (**Tautology**)

7.2   The Semantics of *-er*

While Trace Conversion yields the expected meaning when applied to conservative quantifiers, the result of applying it to non-conservative quantifiers yields a trivial meaning.[30] Note now that *-er* is a non-conservative quantifier. (84) presents the semantic denotation of *-er*.

(84) $[\![\text{-er}]\!](A)(B) = 1$ **iff** $A \subset B$

     (where A, B, are sets of degrees)

The non-conservativity of *-er* stems from the fact that its semantics state that its first argument is a proper subset of its second argument.

(85)   John is taller than Bill is.

     (= The set of degrees to which Bill is tall is a **proper** subset of the set of degrees to which John is tall.)

     (= There is a degree d s.t. John is d-tall and Bill is not d-tall.)

(86) $[\![\text{-er}]\!](A)(B) \;\not\Longleftarrow\Longrightarrow\; [\![\text{-er}]\!](A)(A \cap B)$

     a. $[\![\text{-er}]\!](A)(B) = 1$ **iff** $A \subset B$: **contingent**

b. $\llbracket$-er$\rrbracket(A)(A \cap B) = 1$ **iff** $A \subset A \cap B$: **contradiction**

In (87), we consider the case where the first argument of *-er*, i.e. the degree clause, is merged with *-er* in its base position. The two then undergo QR together leaving behind a copy. The LF after Trace Conversion is as in (87c).

(87) a. Before QR:

[....**[-er [A]]** ....]

(John is [-er [than Bill is tall]] tall.)

b. After QR:

[-er [A]] [....**[-er [A]]** ....]

(**[-er [than Bill is tall]]** [John is **[-er [than Bill is tall]]** tall].)

c. Trace Conversion:

[-er [A]] λd [....**[the [A** d**]** ....]

(**[-er [than Bill is tall]]** λd [John is **[the [**λd_1 **Bill is** d_1**-tall]** d**]** tall].)

The LF in (87c) stands for a contradiction. The second argument of *-er* in (87c) can only be subset of the first argument. Yet given the semantics of *-er*, the first argument must be a proper subset of the second.

7.3 Late Merge avoids the contradiction

We now consider the option of merging the complement of *-er* late i.e. after *-er* has moved to a scope position. This option is shown in (88).

(88) a. Before QR:

[....**[-er]** ....]

(John is [-er] tall.)

b. After QR:

[-er] [....**[-er]** ....]

([-er] [John is [-er] tall].)

c. Late Merge of Degree Clause:

[-er A] [....[-er] ....]

([-er [than Bill is tall]] [John is [-er] tall].)

d. Trace Conversion:

[-er A] λd [....[the d] ....]

([-er [than Bill is tall]] λd [John is [the d] tall].)

Since the degree clause is not merged in the base position of *-er* in (88), the contradiction that emerged in (87) does not arise. The derivation in (88) yields a contingent LF with the expected truth-conditions. We now have an explanation for why the degree clause cannot be merged with *-er* in the base position of *-er*. Merge in the base position is ruled out because it leads to a contradictory LF. [31] The explanation from conservativity also provides an answer to the question of why *-er* and degree clause do not undergo further covert movement.[32] This is because any movement of *-er* with the degree clause would create a contradictory meaning.

(89) Cases Ruled Out:

a. Early Merge of Degree Clause in the base position of *-er* followed by QR of [*-er*+DegClause]

b. Late Merge of Degree Clause after QR of *-er* to a scope position, followed by further QR of [*-er*+DegClause]

We have now derived the following generalization.

(90) Degree Clauses can **only** be merged in their ultimate scope position.

7.4   Why is Late Merge of certain complements possible?

We now address the question of why the degree clause, a complement, can be merged late. Lebeaux (1990) and Fox and Nissenbaum (1999) showed that *wh*-moved and extraposed complements and adjuncts display an asymmetry with respect to Condition C. In particular, while *wh*-movement/extraposition of adjuncts obviates Condition C violations, the movement/extraposition of complements does not. They analyze this asymmetry by proposing that whereas adjuncts could be merged late complements had to be merged with their selecting head.

In considering this question, let us consider the contrast in (91). Given the logic of the argument, the disjoint reference effect in (91a) shows that the complement of *rumor* cannot be merged late. Similarly, the absence of a disjoint reference effect in (91b) shows that the complement of *-er* can be merged late.

(91)   a.   ?? Which rumor that John$_i$ liked Mary did he$_i$ later deny?

   b.   I will tell him$_i$ a sillier rumor (about Ann) tomorrow than Mary told John$_i$.

The difference between (91a) and (91b) follows from the way Trace Conversion works. The complement in (91a) is the complement of the restrictor of the determiner which is affected by the Determiner Replacement part of Trace Conversion. In contrast, the complement in (91b) is the complement of *-er*, which is element replaced by Determiner Replacement. Trace Conversion replaces *-er* and *which* with **the**, but leaves *rumor* untouched. Now if we do late merge of the complement of *rumor*, we have an illegitimate LF.

(92)   LF with Late Merge:
   [Which rumor that John liked Mary] λx [he denied [**the** rumor x]]

higher *rumor* is of type eet, lower *rumor* is of type et.

For (92) to be well-formed type-theoretically, the higher copy of *rumor* would have to be of type eet while the lower copy would be of type et. Since two copies of an element cannot be distinct in types, (92) cannot be semantically well-formed. No corresponding problem arises with *-er* because the lower copy of *-er* is replaced by the Determiner Replacement part of Trace Conversion.[33]

8   Conclusion

This paper has presented an approach to the analysis of comparative constructions that allows us to simultaneously capture two generalizations that seem to pull in opposite directions. One is that the degree clause is a complement of the degree head and the other is that the two cannot surface as sisters. Furthermore, the correlation between the surface location of the degree clause and the interpretation of the degree construction follows from our proposal. This synthesis is achieved by adopting a single cycle-model of grammar (cf. Bobaljik 2002, Pesetsky 2000, Chomsky 2001, Fox & Nissenbaum 1999, Fox 2002 a.o.). This model of grammar makes it possible to interleave overt and covert movement. Thus covert movement of the degree head can precede overt Merge of the degree clause with the degree head.

In addition to capturing the role played by the degree clause in determining the interpretation of a degree construction, our proposal is able to capture the Extraposition-Scope Generalization, the Ellipsis-Scope Generalization, and the related Condition C-Scope Generalization. In an advance on earlier proposals, it is able to predict the contrast between (93a) and (93b).

(93)   a.   *John desires that more people than I do take syntax.

b. John desires that more people take syntax than I do.

Our analysis relates the ungrammaticality of (93a) to the non-resolution of an antecedent contained ellipsis. We find parallel examples with canonical cases of antecedent contained deletion (cf. 94 from Tiedeman (1995) and Wilder (1995) via Fox (2002) a.o.).

(94) a. *I expected that everyone [you do] will visit Mary.

b. I expect that everyone will visit Mary [that you do].

The formal means by which the association between the degree head and the degree clause is established in degree constructions in the current proposal are identical to the means by which the relationship between the quantificational DP and the associated relative clause is established in Fox (2002)'s proposal. The one difference is that degree clauses are syntactically and semantically complements of the degree head while relative clauses are syntactically adjuncts and semantically modifiers on the NP inside the QP.

We conclude this paper with a brief discussion of the crosslinguistic implications of our proposal. In our analysis, covert movement of the degree head plays a crucial role. It would therefore be insightful to examine the applicability of our proposal for languages whose 'LF syntax' seems to differ from English. It would be useful, in particular, to examine languages like Chinese, Japanese, and Hungarian. For Chinese and Japanese, an isomorphism principle is widely assumed according to which LF-relations are faithful to surface c-command relations between quantificational elements. Hungarian has been argued to 'wear its LF on its sleeve' i.e. it does in overt syntax several operations that have been proposed to take place in the covert syntax of other languages. The manifestations of the movement of the degree head and its correlation with the surface position of the degree clause in these languages would help us understand the nature of the

movement of the degree head better.

References

Abney, S. (1987) *The English Noun Phrase in its sentential aspect*, Doctoral dissertation, Massachusetts Institute of Technology, Cambridge, Massachusetts. Distributed by MIT Working Papers in Linguistics.

Andrews, A. (1985) *Studies in the Syntax of Relative and Comparative Clauses*, Outstanding dissertations in linguistics, Garland, New York.

Aoun, J., N. Hornstein, and D. Sportiche (1981) "Some aspects of wide scope quantification," *Journal of Linguistic Research* 1:3, 69–95.

Beck, S. (1996) *Wh-constructions and transparent Logical Form*, Doctoral dissertation, Universität Tübingen, Tübingen, Germany. Distributed SfS Report 04-96.

Bobaljik, J. D. (1995) *Morphosyntax: The Syntax of Verbal Inflection*, Doctoral dissertation, Massachusetts Institute of Technology, Cambridge, Massachusetts. Distributed by MIT Working Papers in Linguistics.

Bobaljik, J. D. (2002) "A-Chains at the PF-Interface: Copies and 'Covert' Movement," *Natural Language and Linguistic Theory* 20, 197–267.

Bresnan, J. (1973) "The Syntax of the Comparative Clause Construction in English," *Linguistic Inquiry* 4, 275–343.

Bresnan, J. (1975) "Comparative Deletion and Constraints on Transformations," *Linguistic Analysis* 1:1, 25–74.

Carlson, G. N. (1977) "Amount Relatives," *Language* 53:3, 520–542.

Chomsky, N. (1965) *Aspects of the Theory of Syntax*, MIT Press, Cambridge, MA.

Chomsky, N. (1993) "A minimalist program for linguistic theory," in K. Hale and S. Keyser, eds., *The view from Building 20: essays in linguistics in honor of Sylvain Bromberger*, MIT Press, Cambridge, MA, 1–52.

Chomsky, N. (2001) "Beyond Explanatory Adequacy," unpublished manuscript, MIT.

Corver, N. (1990) *The Syntax of Left Branch Extractions*, Doctoral dissertation, Tilburg University.

Corver, N. (1993) "A Note on Subcomparatives," *Linguistic Inquiry* 24, 773–781.

Cresswell, M. (1976) "The Semantics of Degree," in B. Partee, ed., *Montague Grammar*, Academic Press, New York, 261–292.

Diesing, M. (1992) *Indefinites*, Linguistic Inquiry Monographs 20, MIT Press, Cambridge, Massachusetts.

Embick, D., and R. Noyer (2001) "Movement Operations after Syntax," *Linguistic Inquiry* 32:4, 555–598.

Fox, D. (2000) *Economy and Semantic Interpretation*, Linguistic Inquiry Monographs 35, MIT Press, Cambridge, Massachusetts.

Fox, D. (2001) "The Syntax and Semantics of Traces," handout of Nov. 2001 talk at the University of Connecticut.

Fox, D. (2002) "Antecedent Contained Deletion and The Copy Theory of Movement," *Linguistic Inquiry* 34.

Fox, D., and J. Nissenbaum (1999) "Extraposition and Scope: A case for overt QR," in *Proceedings of the 18th West Coast Conference on Formal Linguistics*, Palo Alto, CA, CSLI.

Guéron, J., and R. May (1984) "Extraposition and Logical Form," *Linguistic Inquiry* 15:1, 1–32.

Hackl, M. (2001) *Comparative Quantifiers*, Doctoral dissertation, Massachusetts Institute of Technology, Cambridge, Massachusetts.

Hankamer, J. (1973) "Why there are two *than*'s in English?" in *Papers from the Ninth Regional Meeting of the Chicago Linguistics Society*, Chicago Linguistics Society, Chicago.

Heim, I. (1985) "Notes on Comparatives and Related Matters," unpublished manuscript, University of Texas at Austin.

Heim, I. (2000) "Degree Operators and Scope," in *Proceedings of SALT X*, Cornell University, Ithaca, NY, Cornell Linguistics Club, 40–64.

Heim, I. (2002) "Notes on Williams 1974, Fox & Nissenbaum 1999, Fox 2002, and Bhatt & Pancheva ms. 2002," handout from Nov. 25, 2002 lecture, Massachusetts Institute of Technology.

Heim, I., and A. Kratzer (1998) *Semantics in Generative Grammar*, Blackwell textbooks in Linguistics 13, Blackwell, Oxford, England.

Hellan, L. (1981) *Towards an integrated theory of comparatives*, Narr, Tübingen.

Izvorski, R. (1995) "A Solution to the Subcomparative Paradox," in J. Camacho, L. Choueiri, and M. Watanabe, eds., *Proceedings of WCCFL 14*, Stanford, CA, CSLI Publications, 203–219.

Kennedy, C. (1997a) "Antecedent Contained Deletion and the Syntax of Quantification," *Linguistic Inquiry* 28:4, 662–688.

Kennedy, C. (1997b) *Projecting the Adjective: The Syntax and Semantics of Gradability and Comparison*, Doctoral dissertation, University of California, Santa Cruz, Santa Cruz, California. published by Garland in the Outstanding Dissertations in Linguistics series.

Kennedy, C. (2002) "Comparative Deletion and Optimality in Syntax," *Natural Language and Linguistic Theory* 20:3, 553–621.

Larson, R. (1988) "Scope and Comparatives," *Linguistics and Philosophy* 11:1, 1–26.

Lebeaux, D. (1990) "Relative clauses, licensing, and the nature of the derivation," in *Proceedings of NELS 20*, Amherst, Massachusetts, GLSA, 318–332.

Lechner, W. (1999) *Comparatives and DP-Structure*, Doctoral dissertation, University of Massachusetts, Amherst.

Moltmann, F. (1992) *Coordination and Comparatives*, Doctoral dissertation, Massachusetts Institute of Technology, Cambridge, Massachusetts.

Napoli, D. J. (1983) "Comparative Ellipsis: A Phrase Structure Analysis," *Linguistic Inquiry* 14:4, 675–694.

Pesetsky, D. (2000) *Phrasal Movement and Its Kin*, MIT Press, Cambridge, MA.

Postal, P. (1974) *On Raising*, MIT Press.

Rullmann, H. (1995) *Maximality in the semantics of Wh-Constructions*, Doctoral dissertation, University of Massachusetts-Amherst, Amherst, Massachusetts. Distributed by GLSA.

Russell, B. (1905) "On Denoting," *Mind* 14, 479–493.

Sag, I. (1976) *Deletion and Logical Form*, Doctoral dissertation, Massachusetts Institute of Technology, Cambridge, Massachusetts. Distributed by MIT Working Papers in Linguistics.

Selkirk, E. (1970) "On the determiner systems of noun phrases and adjective phrases," manuscript, Massachusetts Institute of Technology.

Sportiche, D. (1997) "Reconstruction and Constituent Structure," handout of a talk presented at MIT.

Sportiche, D. (1999) "Reconstruction, Constituency and Morphology," paper presented at GLOW, Berlin.

Stateva, P. (1999) "In Defense of the Movement Theory of Superlatives," in *Proceedings of ESCOL 99*, Cornell University, Ithaca, NY, Cornell Linguistics Club.

Stechow, A. von (1984) "Comparing semantic theories of comparison," *Journal of Semantics* 3:1, 1–77.

Tiedeman, R. (1995) "Some remarks on antecedent contained deletion," in S. Haraguchi and M. Funaki, eds., *Minimalism and linguistic theory*, Hituzi Linguistics Workshop Series 4, Hituzi Syobo, Japan, 67–103.

Wilder, C. (1995) "Antecedent containment and ellipsis," in *FAS Papers in Linguistics*, FAS, Berlin, 132–165.

Williams, E. (1974) *Rule Ordering in Syntax*, Doctoral dissertation, Massachusetts Institute of Technology, Cambridge, Massachusetts. Distributed by MIT Working Papers in Linguistics.

[2] Although we do not discuss result clauses and the degree heads that license them here, as in (i), our general proposal extends to them too.

(i)   a. Julien is *so* crazy <u>that he eats ants</u>.

b. Monica is *too* cool <u>to care</u>.

[3] For ease of exposition, from now on we will refer to the degree head as *-er*, although the proposal is meant to apply to the equative degree head *as* as well.

[4] While Chomsky's recent proposal is that derivations proceed through multiple phases, each phase has a single-component syntax, with the output of the sequence of derivations being sent to PF and LF simultaneously. Thus, the timing of operations does not correlate with whether or not they have an effect at PF.

[5] For nominal comparatives, we follow Bresnan all the way, but we assume that *much* is not present in adjectival comparatives, unless overtly realized. We believe that the presence of *many/much* in nominal comparatives is motivated morphosyntactically (and likely semantically as well, though discussion of this is beyond the scope of our inquiry). It is overtly observed in the case of equatives: *as *(many) people*, *as *(much) coffee*, and also with other degree expressions: *so/too/that *(many) people*, *so/too/that *(much) coffee*. Degree heads such as *-er, as, so, too, that* can combine only with adjectives. Because of this category restriction, in nominal comparatives an adjective *many* or *much* is needed. With adjectives, the category restriction on the presence of degree heads is already satisfied. Correspondingly, *much* is not necessary, morphosyntactically or semantically: e.g., *as (much) happy*, *so (much) happy*. Thus, unlike Bresnan, we do not posit that forms such as *happier* have the structure [[*-er* much] happy]. Yet expressions such as *more interesting* need the presence of *much*, for a different reason. All degree heads in the class of *-er, so, too*, etc. have the requirement that their sister be an adjective. In addition, *-er* has the lexically specific requirement that its sister adjective be mono- or bisyllabic. Since *interesting* does not meet this idiosyncratic property of *-er*, *much* is required. Further questions are raised by cases like *unhappier*.

[6] Selectional restrictions also apply in the case of result clauses, with *so* selecting *that* and *too* selecting infinitival *to*.

(i)   a. Julien is *so* crazy *that/ *than/ *as* he eats ants.

      b. Barbra is *too* cool *to* care / *that/ *than/ *as* she cares.

As indicated in footnote (2), our analysis is meant to extend to such constructions as well.

[7] There are certain problems for handling selectional restrictions between degree heads and degree clauses by postulating a head-complement relationship. The first involves cases of multiple *-er*s with a single degree clause, as in (i). Such multi-head comparatives are rarely discussed in the literature (see Corver (1993), von Stechow (1984), Kennedy (2002)).

i.   a. People do crazi*er* things at high*er* speeds on the McGrath Highway *than they do other places.* (from Andrews (1985))

     b. Marcille gave a long*er* talk at a bett*er* attended session *than did her husband.*

Furthermore, there are comparatives where the opposite holds - there are multiple degree clauses with a single *-er*, as in (ii).

ii.   a. John is (much) tall*er than Mary than Bill is.*

     b. John has (much) *more* CDs *than Mary than Bill does.*

As far as we are aware, cases such as (ii) have not been previously discussed. They are interpreted as comparisons on the differential argument of *-er*, e.g. (ii.a)

compares the difference between John's and Mary's heights to the difference between Bill's and Mary's heights and may involve, as an anonymous reviewer suggests, coalescence of two adjacent *-er*s.

[8] There is some variability in judgements concerning the availability of ACD with bare plurals and weak NPs in general. See Diesing (1992), page 144, fn. 19 for relevant discussion.

[9] Wold's argument comes to us through Heim (2000).

[10] The *wh*-operator in Spec, CP of the degree clause in (6a) reflects the proposal in Chomsky 1977, a.o., of *wh*-movement in degree clauses (although for Chomsky the operator in this case would be a DP, binding an individual variable, and not a DegP binding a degree variable.) We further assume that an operation of maximalization applies to the *wh*-predicate complement of *than* and the degree clause receives the interpretation of a definite description of degrees, cf. von Stechow (1984), Rullmann (1995).

[11] The one exception that we are aware of involves phrases like *more than n* (cf. i).

   i.   a.   Ralph must be [*more than 6 feet*] but [*less than 6 feet, 4 inches*] tall.

          b.   Julia owns [*more than 5*] but [*fewer than 8*] cars.

We will have a few things to say about these in footnote 31.

[12] The individual argument of the gradable predicate, is presumably the internal argument - the complement of A - or, it could be introduced as an external argument by a *little a* head taking the AP as its complement.

[13] There are variants of the classical analysis that posit shell structures (Izvorski (1995), Lechner (1999).

[14] Heim (2000), who assumes the classical analysis, nevertheless notes that the obligatory superficial (not feeding LF) extraposition is problematic.

[15] As discussed earlier, we assume that *more apples* is the spellout of *many-er apples.*

[16] We also assume that the individual argument of the degree predicate *tall* is introduced in a higher functional head, a *little a.*

[17] This is as far as degree clauses are concerned. Outside of the domain of comparatives, we adopt Fox & Nissenbaum's position that 'extraposed' arguments are A′-moved, whereas 'extraposed' adjuncts are base-generated following QR of the source DP.

[18] The adjunct/argument distinction persists when clausal arguments are considered. A minimal pair is given in (i):

(i) a. ?? Which rumor that John$_i$ liked Mary did he$_i$ later deny?

  b. Which rumor that John$_i$ liked did he$_i$ later deny?

[19] If the high attachment yields a new scopal interpretation, the high attachment is available even if it is string vacuous. This point is discussed at length in §6.2.

[20] It is worth noting that the 'exactly as high' correlation falls out from the analysis proposed by Guéron and May (1984). Within their proposal, extraposed clause and the source XP do not form a constituent at LF. Thus their inability to undergo further covert movement together is to be expected.

[21] This is a simplified representation of the meaning of *before* but is sufficient for the present purposes. A more accurate representation will be crucial for (43).

[22] Example (ia) serves as evidence that extraposition of the degree expression *than*

1000 feet is, in principle, possible, as long as the constituent across which extraposition has applied, does not enter into scopal relationship with *-er*. Example (ib) shows that it is not the presence of ellipsis inside the *before* clause that causes the problem in (44).

(i)   a. Mary climbed higher yesterday than 1000 feet.

        b. ??Mary climed higher before Bill arrived than 1000 feet.

[23] Williams' generalization was illustrated earlier with a strong DP (*every book*). Note that examples like (i), with a weak DP, are ambiguous too. (i) is ambiguous in exactly the same way as (40): its two readings reflect the scope of the object DP with respect to the *before*-clause.

(i)   John read several books before you did.

[24] Some speakers find it hard to get the reading indicated in (47b) for (45). This reading seems to require special contextual support - we have to be talking about a particular set of books. This is unsurprising. The example in footnote 23, repeated here, also shows the relevant ambiguities; and, on the wide scope reading of *several books*, the most easily available interpretation is one where a particular set of books is presupposed.

(i)   John read several books before you did.

[25] This example is modeled on ex. 68 from Guéron and May (1984).

[26] We note that a similar argument can be made for extraposed relative clauses also. (i) shows that the subject QP can take scope over the extraposed result clause.

i. (based on examples (33c, d) from Williams (1974), pg. 205)

   a. No one [who was using a shovel] got so tired yesterday [that he had to quit].

   b. *No one got so tired yesterday [that he had to quit] [who was using a shovel].

   c. *No one got so tired yesterday [who was using a shovel] [that he had to quit].

As before, we relate the ungrammaticality of (i.c) to the Heim-Kennedy constraint. These facts suggest that the subject QP and the extraposed relative clause do not move further covertly. If they could, then at LF, (i.c) could have a representation where it is identical in the relevant aspects to (i.a) and we would lose our explanation for the ungrammaticality of (i.c).

However, the facts seem to be more complicated. If we consider instances involving extraposition from objects, it seems that the 'exactly as high' generalization does not hold (cf. ii).

ii. John is required [to publish [no articles] this year [that are about Tree Adjoining Grammar] in a major journal] [to get tenure].

Despite the extraposed relative clause being in the embedded clause, (ii) seems to allow for both scopal orders.

[27] It might seem that von Stechow's sensible (narrow scope) *de-re* reading for (55) is distinct from the sensible (wide scope) *de-re* reading for (discussed above (cf. von Stechow (1993, 1998)). The intuition is that on the narrow scope *de re* reading, John is mistaken as well, but his thought is comparative, i.e., he thinks Mary

is taller than a particular degree - her actual height. However, as Heim (2000), fn. 17 points out, a DegP interpreted *de re* under an intensional predicate does not, despite its form, represent a comparative thought. See also §4.2 for related discussion.

[28] The cases where the comparison is within the scope of *tell* are systematically further ambiguous between *de re* and *de dicto* interpretations. We do not consider the *de re/de dicto* ambiguity to be a matter of syntactic scope and for this reason we abstract away from it here.

[29] This puzzle is potentially related to an observation made by Kennedy (1997a) and reported in Fox (2000) that there is no contrast between (i.a) and (i.b).

    i.   a.   Polly introduced him$_i$ to everyone Erik$_i$ wanted her to.

          b.   Polly introduced him$_i$ to everyone Erik$_i$ wanted to meet.

Unlike (i.a), there is no antecedent containment to be resolved in (i.b). However, we find that QR is still able to bleed Condition C.

[30] The proposal advanced in this section is based on the discussion in Heim (2002).

[31] As noted in footnote 11, there is one exception to the observation that complements of *-er* are obligatorily extraposed and it involves degree phrases of the form *than* n.

    i.   Mary read [more than three] books.

We do not provide an analysis of such cases. Note though that the closely related sentence **Mary read more than that books* is ungrammatical. It is rempting to relate the acceptability of (i) to the fact that if *more than* n is treated as a complex deter-

miner, it has conservative semantics. However, there are reasons to believe that *more than* n can still behave as a degree quantifier, a point made by Hackl (2001) based on the availability of split readings for sentences like (ii).

ii. John is required to read fewer than 5 papers on this topic.
Split Reading: [fewer than 5] > required > papers

The non-obligatoriness of extraposition for *than* n is potentially related to the fact that they do not obey the 'exactly as high' generalization as shown by the availability of a split reading in (ii).

[32] An alternate explanation for the **exactly** part of the generalization follows from a generalization according to which overt A-bar movement blocks further cover A-bar movement (cf. Aoun et al. (1981)). The movement of the degree head/QP even though itself 'covert' counts as overt for this generalization because the 'covert' movement has effects on phonology by providing a site for the overt merge of degree clauses/relative clauses. Hence the constituent formed after the merge of the degree head with the degree clause is unable to move further in covert syntax. One piece of evidence that would help distinguish between the proposal in the main text and this alternative concerns the behavior of relative clauses extraposed from NPs. If relative clauses extraposed from NPs do not display 'exactly as high' effects, it would argue against the alternative proposal. See the discussion in footnote 26.

[33] We exploit the fact that Trace Conversion allows for late merge of the complements of elements that are replaced by Determiner Replacement to explain why the complement of *-er* can be late merged. This raises the following problem noted in footnote 11 of Fox (2002), and attributed by Fox to Uli Sauerland: why

can't NP's be late merged in a position discontinuous from their determiners? Note the ungrammaticality of (i).

i.   *I gave him [$_D$**every**] yesterday [$_{NP}$**book that John wanted to read**].

We do not have a proper answer to this question. The following lines of inquiry seem to hold some promise. The first is that only elements that do not need to be case-licensed like PP/CPs can extraposed. NP's need to be case-licensed and hence they cannot be extraposed. A related idea is that NP's receive θ-roles and hence must be merged locally with the θ-assigning head. Degree clauses do not receive a θ-role from *-er* and hence can be merged late.