Late Merger of Degree Clauses

Rajesh Bhatt and Roumyana Pancheva

In this article, we propose that degree heads and degree clauses form a constituent not at the point where the degree head is merged, but after QR of the degree head and countercyclic merger of the degree clause. We derive a generalization originally outlined in Williams 1974 that the scope of the comparative degree quantifier is exactly as high as the site of attachment of the degree clause. This generalization is shown to follow from the derivational mechanism of countercyclic merger and a semantic property of the comparative degree head, namely, its nonconservativity.

Keywords: comparatives, scope, interpretation of copies, antecedent-contained deletion (ACD), conservativity, extraposition

The focus of this article is the place of degree clauses (i.e., *than-/as*-phrases) in the overall architecture of comparatives.[1] We propose that degree clauses are merged late, after the degree head *-er/as* has moved to a scope position.[2] The position where the degree clause merges is the position in which it 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 post-Spell-Out timing; temporally, it *precedes* the merger 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 Quantifier Raising (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's 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 that ensures that the scope of a source DP is at least as high as the attachment site of the extraposed

[1] 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 *that he eats ants.*
    b. Monica is *too* cool *to care.*

[2] 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.

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 specification ''exactly as high'' 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 (see Fox 2001, 2002). Finally, the question of what can be merged late arises. It has been proposed that adjuncts but not arguments can be merged countercyclically (see 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 merger 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 merger 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 (see Fox and Nissenbaum 1999, Fox 2002) or variants thereof, such as grammar based on multiple Transfer (see Chomsky 2001).[3] Our proposal also provides indirect support for what has come to be known as the phonological theory of QR (see, e.g., Bobaljik 1995, 2002, Pesetsky 2000), 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* and the gradable predicate (e.g., *tall*) and the relationship between *-er* and the degree clause, the overall structure of comparatives has remained elusive. A number of divergent analyses have been proposed regarding the position in which the degree clause merges (a representative list of the various approaches includes those in Bresnan 1973, 1975, Larson 1988, Moltmann 1992, Kennedy 1999, Lechner 1999, Heim 2000). In section 1, we present the arguments that have motivated these analyses.

1 Constituency in Degree Constructions

Comparatives exhibit properties that cannot be easily accommodated in a single structure. In particular, there is convincing syntactic and semantic evidence that the degree clause is the complement of *-er*. However, there is also 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 nonadjacent in the majority of cases, more often than not they may not even appear together as a constituent that excludes the gradable predicate (e.g., *tall*). These

---

[3] While Chomsky's recent proposal is that derivations proceed through multiple phases, each phase has a single-component syntax, 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.

diverging properties suggest that the degree clause is obligatorily extraposed. In section 7, we will consider why this is so.

1.1 The Degree Clause as an Argument of the Degree Head

1.1.1 Selectional Restrictions despite Nonadjacency   One of the clearest syntactic pieces of evidence that the degree clause forms a constituent with the degree head is the selectional restrictions between the two. As shown in (1), differential comparatives (i.e., comparatives with *more, less,* and *fewer*) cooccur with a degree clause introduced by *than,* whereas equatives (i.e., comparatives with *as*) cooccur with 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), among others, 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 as *fewer people* is derived from [*-er few*] *people,* so *more people* has the underlying structure [*-er many*] *people, more coffee* has the structure [*-er much*] *coffee,* and *less coffee* has the structure [*-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.[4] Given this analysis, it is possible to reduce the cooccurrence 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*

---

[4] For nominal comparatives, we follow Bresnan in all respects, 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,* and *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,* and the like, require 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.*

In other words, there are selectional restrictions between *-er* and *than,* and between *as* and *as.*[5] 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, 1975) and Carlson (1977), among others.[6]

1.1.2 Semantic Constituency despite Nonadjacency   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, as in (3), it is also to be expected that they can be quantified over, as in (4).

(3) a. John is 6 feet tall.
    b. John is that (much) tall.

(4) a. John is taller than 6 feet.
    b. John is [$_{AP}$[$_{DegP}$ -er than 6 feet] tall]
    c. [$_{DegP}$ -er than 6 feet]$_1$ 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 (see, e.g., Cresswell 1976, von Stechow 1984, Heim 1985, 2000). Under fairly standard assumptions, quantificational expressions undergo QR; similarly, the degree quantifier [*-er* + 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).

---

[5] 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 (1), our analysis is meant to extend to such constructions as well.

[6] Two problems face the attempt to handle 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 multihead comparatives are rarely discussed in the literature (see Corver 1993, von Stechow 1984, Kennedy 2002).

(i) a. People do craz*ier* 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 (iia–b) have not been previously discussed. They are interpreted as comparisons on the differential argument of *-er;* for example, (iia) compares the difference between John's and Mary's heights with the difference between Bill's and Mary's heights and may involve, as an anonymous reviewer suggests, coalescence of two adjacent *-er*s.

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.* However, there is additional evidence, coming from considerations of ellipsis resolution, that *-er* and the degree clause can take 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)).[7] 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 Ā-gap to serve as a suitable antecedent for the ellipsis in the relative clause. Within the context of Diesing's (1992) 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 this puzzle.[8] He proposes that what undergoes QR in (5b–c) is the degree quantifier formed of *-er* and the degree clause, as shown in (6a).[9] 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 representation for (5b)*
    a. [-er [(than) *wh* Bill was Δ]] [John was climbing *d*-many trees]
    b. [-er [(than) λ*d* Bill was <u>climbing *d*-many trees</u>]] [λ*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] There is some variability in judgments concerning the availability of ACD with bare plurals and weak NPs in general. See Diesing 1992:144, fn. 19, for relevant discussion.

[8] Wold's argument comes to us through Heim 2000.

[9] The *wh*-operator in Spec,CP of the degree clause in (6a) reflects the proposal made by Chomsky (1977), among others, that *wh*-movement takes place 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 (see von Stechow 1984, Rullmann 1995).

(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]$_1$ [the paper be t_1-long]]
        $\forall w \in$ Acc: max {d: long$_w$(p,d)} = 15 pages
        (where *Acc* = the set of accessible worlds)

    b. *-er > required:* [exactly 5 pages -er than that]$_1$ [required [the paper be t_1-long]]
        max {d: $\forall w \in$ Acc: long$_w$(p,d)} = 15 pages

Under one of its readings, (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 worlds 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 is predicted if *-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 *less*-comparatives (Stateva 1999).

(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 that the degree head and the degree clause form 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. In the next section, we present this evidence in some detail.

1.2 Constituency of the Degree Head and the Matrix Gradable XP

1.2.1 Suppletive Forms One reason for positing that *-er* and the gradable predicate form a constituent 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, *-er* and *tall* can still merge, 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 Nonadjacency between -er *and the Degree Clause* Much stronger evidence against the constituency of *-er* and the degree clause seems to come from the fact that in general,[10] it is not possible for the two to appear together.

(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.

---

[10] The one exception that we are aware of involves phrases like *more than n.*

(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.

    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 analyzed as a clausal comparative (like (11)) to which comparative ellipsis has applied, 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 that degree clauses are arguments 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 (see (13)). For the ones that do, extraposition alone cannot be the answer to ellipsis resolution. Extraposition may resolve antecedent containment, but it will not, by itself, create a gap of the requisite 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 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 alternative 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 extraposition of the degree clause is, in effect, obligatory, and in section 7 we will explain why this is so.

2 The Architecture of Degree Constructions

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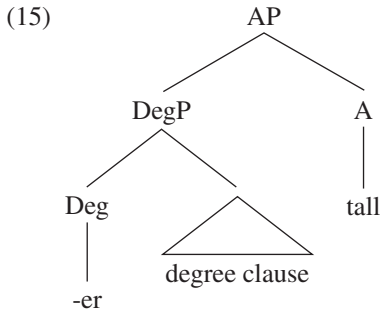
2.1 The Classical View

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 of the gradable predicate, as in (15)[11] (see Chomsky 1965, Selkirk 1970, Bresnan 1973, Heim 2000).[12]

---

[11] The individual argument of the gradable predicate is presumably the internal argument—the complement of A. Alternatively, it could be introduced as an external argument by a ''little a'' head taking the AP as its complement.
    [12] There are variants of the classical analysis that posit shell structures (Izvorski 1995, Lechner 1999).

(15)   AP
                          ╱   ╲
                    DegP   A
                   ╱   ╲   │
               Deg   ╱╲   tall
                │   ╱   ╲
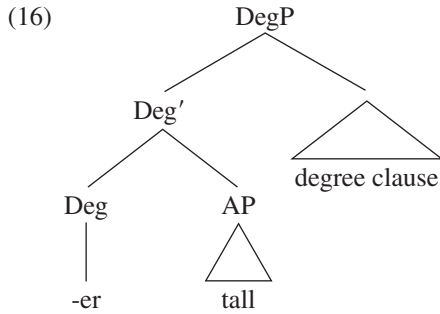   -er degree clause

As one might 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 (see (11)–(14)). This movement is not motivated by anything apart from the need to derive the correct surface word order.[13] Analogous extraposition of relative clauses, for instance, is never obligatory. Furthermore, LF movement of the degree clause together with *-er* (see Heim 2000 and the discussion around examples (4) and (5)) would necessitate reconstruction of the degree clause after the extraposition. This raises several issues. Given that reconstruction is currently taken to involve not literally putting the moved element back in its original position, but simply interpreting its copy, the LF movement of [*-er* + degree clause] would require moving the copy of the extraposed degree clause at LF. But if that were a general option in the grammar, a well-known generalization about Ā-movement would be directly contradicted: overtly Ā-moved elements have frozen scope (i.e., they cannot undergo further movement at LF; see Aoun, Hornstein, and Sportiche 1981). If overt Ā-movement could be ''undone'' by LF movement of the copy, frozen scope would not obtain. An account invoking LF movement of the copy also predicts that the surface position of the degree clause will be irrelevant—in fact, invisible—to semantic interpretation. But as it turns out (see section 5.2), 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 that *-er* and the degree clause do not form a constituent. A representative structure is illustrated in (16) (see Abney 1987, Larson 1988, Corver 1990, 1993, Kennedy 1999, 2002, for variants on this structure).

---

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

(16)

DegP
Deg′
Deg
-er
AP
tall
degree clause

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 that posit a coordination structure for comparatives (see Hankamer 1973, Napoli 1983, Moltmann 1992; for a subset of cases, see Corver 1990, 1993, and parts of Lechner 1999). These do not have the problems of extraposition; nonetheless, they still have difficulties with assigning scope for *-er* and the degree clause, and they run into additional problems related to the posited coordination structure for comparatives.

3 Late Merger of the Degree Clause Resolves the Conflict

We propose that the conflicting evidence regarding constituency in degree constructions can be reconciled by positing 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, during 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;* that is, the selectional restrictions can be stated in a local head-to-head configuration. There is no need to posit rightward Ā-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 and the semantic scope of the degree head.

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, 2002, Pesetsky 2000, Fox and Nissenbaum 1999, Fox 2001). It also appears that a morphological condition on *-er* is responsible for the pronunciation of the copy; that is to say, *-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.[14] This point also applies to *as, so,* and *too,* which are morphologically prefixes rather than suffixes like *-er.*

In its essentials, this proposal follows Fox and Nissenbaum's (1999) analysis of relative clause extraposition. Developing Lebeaux's (1990) proposal that relative clauses can be merged countercyclically, Fox and 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 merger to the domain of comparatives, it also shows that not only adjuncts but complements as well can be merged late.
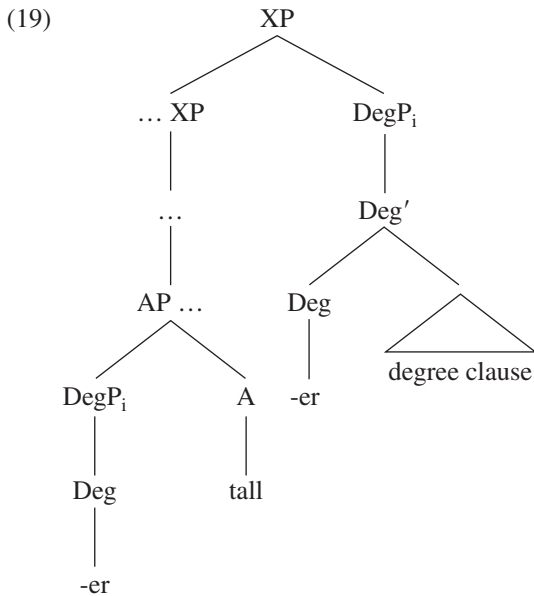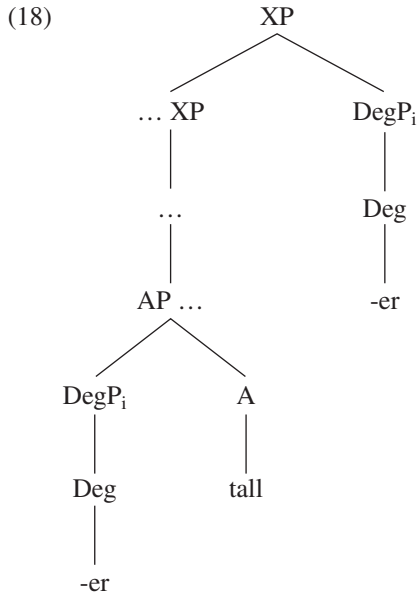
Let us illustrate the proposal in some more detail. As shown in (17), we assume that *-er* is the head of a DegP and (for ease of presentation) that the DegP is an argument of the gradable predicate, as in the classical analysis. We also adopt (for further convenience) the notion that the DegP is a specifier of the gradable predicate.[15]

(17)
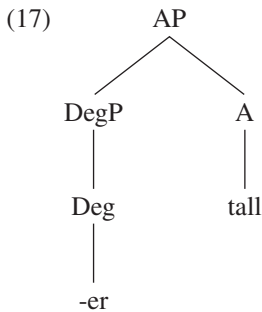
Being a quantificational expression, the DegP headed by *-er* undergoes QR, leaving behind a copy. Following Fox and Nissenbaum (1999), we assume that it right-adjoins in a scope position; that is, it adjoins to a node of type ⟨t⟩ (indicated as XP in the trees in (18) and (19)). The degree clause is then merged as an argument to the QR-ed *-er* (see (18) and (19) for an illustration).

[14] As discussed earlier, we assume that, for example, *more apples* is the spell-out of *many-er apples.*
[15] We also assume that the individual argument of the degree predicate *tall* is introduced in a higher functional head, ''little a.''

(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 merger of the degree clause is countercyclic, it targets a position that is not the root. However, since this is precisely the remerger operation that also happens in the case of

head movement, it should not be seen as problematic (at least no 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 undergo 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 LF representations in (20a–c) reflect the possible positions in which the degree clause can merge.

(20) Bill wants to be taller than John is.
    a. Bill wants PRO$_j$ to be [$_{aP}$[$_{aP}$ t$_j$ [$_{AP}$ t$_i$ tall] [-er$_i$ [than John is tall]]]]
    b. Bill wants [$_{IP}$[$_{IP}$ PRO to be [$_{AP}$ t$_i$ tall]] [-er$_i$ [than John is tall]]]
    c. [$_{IP}$[$_{IP}$ Bill wants [$_{IP}$ PRO to be [$_{AP}$ t$_i$ tall]]] [-er$_i$ [than John 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 Ā-movement of the degree clause. In our analysis, such rightward Ā-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*.

(21) Nicole made more money last year than Tom did.
    a. 'Nicole made more money last year' is generated.
    b. QR moves *-er* to the matrix clause. The lower copy is pronounced, and so this step is ''covert.''
       Overt structure: [[Nicole made more money last year] -er]
       LF structure: [[λ*d* 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 structure: [[λ*d* Nicole made *d*-much money last year] [-er [λ*d* Tom made *d*-much money last year]]]

Our approach is similar in spirit to Guéron and 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. Guéron and 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 Guéron and 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 Guéron and 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 that of Guéron and May—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 were 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 merger 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 why can *-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, 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 merger of the degree clause with *-er* in a scope position is the only way of achieving a semantically well formed structure at 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.

In this section, we have 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 that has not been previously considered and that, importantly, relies on no special assumptions. The two crucial aspects of the proposal—countercyclic merger and comparative operator movement—are operations that have been independently proposed and are, arguably, well justified (the former outside 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 directly relates the surface position of the degree clause with the scope of *-er*. The rest of the article is devoted to presenting detailed evidence in support of our proposal for late merger 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 *-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 (see Heim 2000). Arguments have been made, however, against positing such movement in the LF syntax of comparatives. One influential argument against movement of the DegP is due to Kennedy (1999) and has been developed further by Heim (2000). Kennedy notes that putative movement of the DegP predicts the existence of readings that are not in fact attested. Consider (22).

(22) John is 4 feet tall. Every girl is exactly 1 inch 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)]$\}$ = $4'$ + $1''$
     (Heim 2000:(22))

(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 were available, we could judge (22) as true in the scenario where the shortest girl is exactly 4 feet 1 inch tall but the other girls are taller. Similar points are made by (23).

(23) a. (The frostline is $3\frac{1}{2}$ feet deep.) Mary set every post exactly 2 feet deeper than that.
     (Heim 2000:(25))

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

Both (23a) and (23b) lack the reading whereby the DegP takes scope over the QP. For example, (23) is not judged true in the scenario where Mary set just the least deep post exactly 2 feet 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.
     (Heim 2000:(27))

On the basis of his observation that a DegP can never take scope over a QP and that more generally, putative degree quantifiers always take the narrowest possible scope, Kennedy (1999) 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 it 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) *$\lambda d \ldots$ QP $\ldots d \ldots$

It is not clear why a constraint like (25) exists, 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 (see 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 examples we have used so far to demonstrate the semantic effects of DegP movement have had a very particular form. These examples are repeated in (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* versus *de dicto* interpretation. We take *de re* interpretation to not be a matter of scope; that is, 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 LF structures (28a–c) to (28). The *de dicto* narrow scope interpretation is quite clearly distinct 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 noncomparative desire, while the narrow scope *de re* reading seems to involve the expression of a comparative desire.

  Unfortunately, our intuitions turn out to be unreliable here. As Heim (2000:fn. 17) points out, despite their form LF structures like (28b) do not (necessarily) involve comparative desires. To see this point clearly, consider the semantic representations assigned to the *de re* LF structures 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)} > \$1 billion
   (where *Boul* = the set of bouletically accessible worlds)
  b. *than*-clause *de re, want > -er*
   $\forall w \in$ Boul: [max {d: rich$_w$(j,d)} > \$1 billion]

Given the semantics assumed above for *want,* it turns out that the wide scope *de re* LF structure and the narrow scope *de re* LF structure 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 judgments 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
　　　　(where *Acc =* the set of accessible worlds)
　　　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:* $\exists w \in$ Acc: [max {d: long$_w$(p,d)} > 10]

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

There is also another complication that needs to be kept in mind when trying to detect the ambiguity of a DegP with respect to an intensional operator using truth-conditional judgments. The semantics of certain intensional operators combines with DegP movement in such a way that the wide and narrow 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 semantics of these predicates and the semantic noneffect 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 movement. 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 movement even when truth-conditional judgments do not help to distinguish between wide and narrow scope.

5 Evidence from Extraposition

As a descriptive term, *extraposition* refers to the phenomenon whereby a constituent of a phrase appears separated from the phrase and to its right. This phenomenon has been analyzed in various ways. Perhaps most commonly, it is proposed that the discontinuous constituent is merged with its source phrase and is then rightward-moved to its surface position—the Ā-movement analysis. According to another analysis, the discontinuous constituent moves to the left and then is stranded by the source phrase, which moves leftward even higher—the remnant movement analysis. Finally, it has been proposed that the discontinuous constituent is base-generated in its surface

position, and various mechanisms are employed to ensure that it is interpreted together with its ''source'' phrase—the base generation analysis. Below, we use *extraposition* largely pretheoretically, as a description of the cases where the degree clause appears discontinuous from the degree predicate, separated from it by sentential material such as adjuncts (as in (21)). As we stated above, we take such degree clauses to be merged in their surface positions; that is, we employ a version of the base generation analysis.[16] We consider the interaction of ''extraposition'' with (a) binding theory Condition C effects and (b) scope of the comparison. Both of these interactions provide evidence that degree clauses are merged late, following QR of the DegP headed by *-er*.

5.1 Extraposition and Condition C

It is well known that standard cases of Ā-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 Ā-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 Ā-movement and Condition C. Whereas Ā-moved arguments (as in (31)) behave as if interpreted in the base position, adjuncts do not. In (32b), the R-expression *John* is contained in an adjunct, and Condition C is not violated. 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 that can be interpreted).[17]

(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

---

[16] This is as far as degree clauses are concerned. Outside the domain of comparatives, we adopt Fox and Nissenbaum's (1999) position that ''extraposed'' arguments are Ā-moved, whereas ''extraposed'' adjuncts are base-generated following QR of the source DP.

[17] The adjunct/argument distinction holds for clausal arguments as well. 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?

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 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 article, 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 nonequivalent 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.[18] Hence, (34a) displays Condition C effects. 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 extraposition of 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 the extraposition of complements does not.

(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 one Lebeaux (1990) gives the contrast in (32): the adjunct in (36a) can be merged late in its surface

---

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

position (after QR of the DP *an argument* to the right) and hence does not have to leave behind a copy that would trigger Condition C. (36b) involves a complement that 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) to demonstrate that late merger 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 merger. Likely, the necessity for θ-licensing forces arguments of lexical heads to be merged with their predicates. Adjuncts and degree clauses need not be θ-licensed and are thus free to enter 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 (1997, 1999). According to his proposal, determiners and their NP arguments are merged independently: NPs are merged in the θ-domain of predicates (e.g., verbs), whereas determiners are merged later, in the inflectional domain associated with the predicate.

Another difference between the cases at hand involving late merger of degree clause complements and the cases of late merger of complements disallowed by Lebeaux (1990) and 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 and by Fox and Nissenbaum involve complements of restrictors of determiners. In section 7, we will show that Fox's (2002) mechanism for interpreting traces is sensitive to this distinction. Late merger of complements of restrictors of determiners yields an illegitimate object, while late merger of complements of determiners/*-er* yields a semantically well formed object.

5.2 *Extraposition and Scope of* -er

5.2.1 *The Extraposition-Scope Generalization*   Williams (1974) establishes a correlation between the surface position of an extraposed clause and the scope of its source XP. He states this generalization as follows:

(37) If two scope items *x* and *y* with their determining clauses are represented in Deep Structure as: $[[x \ S_1] \ldots [y \ S_2]]$
and if extraposition yields the structure: $\ldots x \ldots y \ldots S_1 \ S_2 \ldots$ then semantically $(y(x))$;
if it yields $\ldots x \ldots y \ldots S_2 \ S_1 \ldots$ then semantically $(x(y))$.
(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 articulated the extraposition-scope correlation as follows:

(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).
(Fox 2002:(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.[19]

(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 merger site of β.

First, we will illustrate the generalization as articulated in (38); that is, we will show that the scope of the degree head is *at least as high* as the degree clause. Then we will show that the scope of the degree head is *exactly as high* as the degree clause. Finally, in section 7 we will provide an explanation for the generalization.

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 undergo long QR 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 did.

(40) I read every book before you did. (ambiguous)
  a. Reading 1: *before > every book,* 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$_1$, ∀x, book x, I read x, and at t$_2$, ∀x, book x, you read x, and t$_1$ < t$_2$[20]
  b. Reading 2: *every book > before,* 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$_i$]]]]
    ∀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

---

[19] It is worth noting that the ''exactly as high'' correlation follows from the analysis proposed by Guéron and May (1984). Within their proposal, the 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.

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

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. (ambiguous)
  a. Reading 1: *before > every book*
  b. Reading 2: *every book > before*

(42) I read every book before you did that John had recommended. (unambiguous)
  a. *Reading 1: *before > every book*
  b. Reading 2: *every book > before*

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.

The Extraposition-Scope Generalization follows from Fox and Nissenbaum's (1999) and Fox's (2002) proposal in this 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 undergone QR to this position. It follows that only the reading corresponding to the LF structure 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 the DegP headed by *-er* to undergo QR, followed by late merger 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. However, it turns out that *-er* cannot take scope over the *before*-clause. This is a manifestation of a wider phenomenon observed by Kennedy (1999), to which we will return. Importantly for present concerns, the inability of *-er* to take 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 (43) and (44).

(43) Mary climbed higher than 1,000 feet before you did. (unambiguous)
  a. Reading 1: *before > -er*
    i. Mary [$_{VP}$[$_{VP}$ climbed [$_{AP}$ t_i high]] [$_{DegP}$ -er [than 1,000 feet]]] [before you did [$_{VP}$ Δ]]

ii. Mary PAST [$_{VP}$[$_{VP}$ climb [$_{AP}$ t$_i$ high]] [$_{DegP}$ -er [than 1,000 feet]]] [before you did [$_{VP}$[$_{VP}$ climb [$_{AP}$ t$_i$ high]] [$_{DegP}$ -er [than 1,000 feet]]]]

The earliest time t$_1$ [max d [Mary climbed to d at t$_1$] $d > 1,000$ feet] the earliest time t$_2$ [max d [you climbed to d at t$_2$] $d > 1,000$ 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 1,000 feet]]

ii. [$_{IP}$[$_{IP}$ Mary PAST [$_{VP}$ climb [$_{AP}$ t$_i$ high]] [before you did [$_{VP}$ climb [$_{AP}$ t$_i$ high]]]] [$_{DegP}$ -er [than 1,000 feet]]]

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$_2$] $d > 1,000$ feet

(44) *Mary climbed higher before you did than 1,000 feet.

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

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

Example (43) is grammatical, but 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 1,000-foot mark before you do. The reading that results when *-er* has 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 1,000-foot mark, as long as then Mary reached another height (say, 1,100 feet) 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.[21] 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 move by 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.

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 that further supports 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 take 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

---

[21] Example (ia) serves as evidence that extraposition of the degree expression *than 1,000 feet* is in principle possible, as long as the constituent across which extraposition has applied does not enter into a 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 1,000 feet.
　　b. ??Mary climbed higher before Bill arrived than 1,000 feet.

DP, similarly to what we find in (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.*[22] 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 LF structures behind readings 1 and 2 are illustrated in (47).

(47) a. Reading 1 (available in (45), unavailable in (46)): *before > -er . . . d-many books*
    i. John [$_{VP}$ read [[t$_i$ many books] [-er [than Mary published in her life]]$_i$]] [before you did [$_{VP}$ Δ]]
    ii. John PAST [$_{VP}$ read [[t$_i$ many books] [-er [than Mary published in her life]]$_i$]] [before you did [$_{VP}$ read [[t$_i$ many books] [-er [than Mary 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*[23]
    i. John [[$_{VP}$ read t$_i$] [before you did [$_{VP}$ Δ]]] [[t$_j$ many books] [-er [than Mary published in her life]]$_j$]$_i$
    ii. John [[$_{VP}$ read t$_i$] [before you did [$_{VP}$ read t$_i$]]] [[t$_j$ many books] [-er [than Mary 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.

---

[22] Williams's 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.

[23] 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 22, 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.

In contrast to (45), (46) is unambiguous. It only has the reading shown in (47b), where the comparative DP takes scope 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 takes scope 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 has 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).[24]

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

(48b) is ungrammatical because it violates 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 structure of (48a)*
     so [that . . .] λd_1 [[d_1-many people] λx [-er [than . . .] λd_2 [x ate d_2 fast]]]

In a structure like (50) where the two degree abstractions cross, ungrammaticality results 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 9 p.m.].
     b.   *More people ate so fast yesterday [that we were all done by 9 p.m.] [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 to be ruled out by the Heim-Kennedy Constraint.

---

[24] This example is modeled on example (63) from Guéron and May 1984.

(51) *Degree abstractions in the LF structure 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 as in (52) does not help in rescuing the structure.

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

Like (51), (52) 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 exactly marks the scope of the corresponding degree heads. However, these examples involve a confound in that further covert movement of the degree quantifier violates the Heim-Kennedy Constraint. We need to find a case where potential further movement of the degree quantifier would result in a legitimate LF structure and then show that the corresponding reading is unavailable.

      The following examples provide such a case. They involve an intensional predicate and a degree quantifier such that the two LF structures corresponding to the two possible scopes of the intensional predicate and the degree quantifier are semantically distinct (see section 4.2 and Heim 2000 for details).

(53) a. *Degree clause inside the embedded clause*
          John is required [to publish fewer papers this year [than that number] in a major journal] [to get tenure].
          Simplified LF structure: *required* > [*fewer* [*than n*]]
          required [fewer [than n] λd [PRO publish *d*-many papers]]
      b. *Degree clause outside the matrix clause*
          John is required [to publish fewer papers this year in a major journal] [to get tenure] [than that number].
          Simplified LF structure: [*fewer* [*than n*]] > *required*
          fewer [than n] λd [required [PRO publish *d*-many papers]]

(54) a. *Degree clause inside the embedded clause*
          John is required [to publish exactly 5 more papers this year [than that number] in a major journal] [to get tenure].
          Simplified LF structure: *required* > [*exactly 5 more* [*than n*]]
          required [[exactly 5 -er] [than n] λd [PRO publish *d*-many papers]]
      b. *Degree clause outside the embedded clause*
          John is required [to publish exactly 5 more papers this year in a major journal] [to get tenure] [than that number].
          Simplified LF structure: [*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 inside the embedded clause have only the reading in which *required* takes scope over the comparison (see (53a), (54a)). This is why (53a) has the pragmatically odd reading that if John publishes more than a certain number of papers, he will

not get tenure. Likewise, when the degree clause is outside the embedded clause, only the reading in which the comparison takes scope over *required* is available. This is the pragmatically reasonable reading that limits from above the minimum number of papers John needs to publish to get tenure. It does not penalize John for being productive.

The availability of the *-er > required* reading in (53b) and (54b) shows that the structure involving a degree abstraction that crosses *required* is semantically well formed. The absence of this reading in (53a) and (54a) shows that the scope of *-er* is marked exactly by the surface position of the degree clause; in other words, the degree quantifier in (53a) and (54a) cannot move further.[25]

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 (see Russell 1905) is exhibited by sentences such as (55). On one reading, (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 feet 8 inches tall). 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, for example, *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 inside 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 taking

---

[25] A similar argument can be made for extraposed relative clauses. (i) shows that the subject QP can take scope over the extraposed result clause.

(i) 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].
         (based on (33c–d) in Williams 1974:205)

As before, we relate the ungrammaticality of (ic) 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 (ic) could have a representation where it is identical in the relevant aspects to (ia), and we would lose our explanation for the ungrammaticality of (ic).

However, the facts seem to be more complicated. In cases involving extraposition from objects, it seems that the "exactly as high" generalization does not hold.

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

Even though the extraposed relative clause is in the embedded clause, (ii) seems to allow both scopal orders.

scope over *think,* and *think* taking scope over *-er*). Such a nonscopal approach to Russell's ambiguity is required in cases like (56).

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

Assigning the DegP scope outside the *if*-clause is problematic because *if*-clauses are strong islands. In addition, such a movement yields incorrect truth-conditions. Von Stechow's (1984) 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.[26]

Let us now consider the interaction of the Russell's ambiguity cases with extraposition. (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 more closely 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. We will consider 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 at the surface it appears extraposed higher than *tomorrow,* it must be the case that *-er* has undergone QR to the matrix clause. Predictably, only the high scope *de re* reading is available. The derivation is shown in (58).

[26] 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 (55) (see 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; that is, 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 section 4.2 for related discussion.

(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 moves *-er* 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 structure: [[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 structure: [[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 be the case with DegPs as well. 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 alternative derivations that do not involve QR out of a finite clause. In contrast, the only derivation available for (57c) involves 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 presented 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 the c-command domain of the pronoun (see section 5.1). This, together with the fact that Ā-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 Ā-movement; rather, 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 (section 6.1), as does a related correlation between Condition C and scope (section 6.2). Finally, we will discuss cases of trapped ACD that are problematic for alternative analyses and show that our proposal makes the right predictions for these (section 6.3).

6.1 The Ellipsis-Scope Generalization

Sag (1976) and Williams (1974) note that there is a correlation between ellipsis and scope. This correlation has 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*).[27]

(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 (see (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 missing reading (= (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 because of a failure of ellipsis resolution.

Our analysis of the Ellipsis Scope Generalization essentially follows Heim's (2000)—the difference being that in her analysis, the degree clause and the degree head are merged in the

---

[27] 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.

base position of *-er* and move as a unit at LF. (62) demonstrates how the Ellipsis-Scope Generalization follows from our analysis. Following Fox and 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 by QR of *-er* to the right, then by merger of the degree clause as the complement of *-er.* The resulting syntactic object is then combined with the remaining lexical items to yield the structures in (62a–b). Choosing the embedded VP to resolve the VP-ellipsis in the degree clause yields the LF structure in (62a). The reading indicated by the LF structure in (62b) is unavailable because there is no VP in the structure that can license the required ellipsis resolution. 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 the reading (62c) (embedded VP) or (62d) (matrix VP).

(62) Mary's father tells her to work harder than her boss does.
    a.   *tell > -er,* elided material = *work d-hard*
        The degree clause is merged low, at the level of the embedded clause.
        LF structure: Mary's father tells her [[λ*d* PRO to work *d*-hard] [-er [λ*d* her boss works *d*-hard]]]
    b.   **tell > -er,* elided material = *tell her to work d-hard*
        The degree clause is merged low, at the level of the embedded clause.
        LF structure: Mary's father tells her [[λ*d* PRO to work *d*-hard] [-er [λ*d* her boss tells her to work *d*-hard]]]
    c.   *-er > tell,* elided material = *work d-hard*
        The degree clause is merged high, at the level of the matrix clause.
        LF structure: [[λ*d* Mary's father tells her to work *d*-hard]] [-er [λ*d* her boss works *d*-hard]]
    d.   *-er > tell,* elided material = *tell her to work d-hard*
        The degree clause is merged high, at the level of the matrix clause.
        LF structure: [[λ*d* Mary's father tells her to work *d*-hard]] [-er [λ*d* her boss tells her to work *d*-hard]]

6.2 Correlations with Condition C

Further examination of the examples used to motivate the Sag-Williams Ellipsis-Scope Generalization reveals that there is also a correlation between coreference judgments 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 in (63).

(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 = *tell 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 = *tell her to work d-hard,* coreference possible

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

The only reading available for (64a) is the reading where the comparison takes scope over the matrix predicate. In contrast, both readings are available in (64b).

    In Heim's (2000) analysis, the impossibility of coreference in (63a) and the missing reading in (64ai) 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 in the degree clause is possible only if the degree head and the degree clause move by QR 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 structure: her father tells her [[-er [λ*d* Mary's boss works *d*-hard]] [λ*d* PRO to work *d*-hard]]
    b. **tell > -er,* elided material = *tell her to work d-hard,* ruled out by conditions on ellipsis resolution
       LF structure: her father tells her [[-er [λ*d* Mary's boss tells her to work *d*-hard]] [λ*d* PRO to work *d*-hard]]
    c. *-er > tell,* elided material = *work d-hard,* coreference possible, no Condition C violation
       LF structure: [[-er [λ*d* Mary's boss works *d*-hard]] [λ*d* her father tells her to work *d*-hard]]
    d. *-er > tell,* elided material = *tell her to work d-hard,* coreference possible, no Condition C violation
       LF structure: [[-er [λ*d* Mary's boss tells her to work *d*-hard]] [λ*d* 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) (from Fox 2000:175).

(66) a. ??/*You bought him$_i$ every picture that John$_i$ liked.
    b.   *He$_i$ bought you every picture that John$_i$ liked.

Fiengo and May (1994) and Fox (2000) note that there is an environment where QR does bleed Condition C: namely, when QR resolves antecedent containment. This is shown in (67) (from Fox 2000:180).

(67) 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, it is possible for QR to bleed Condition C even when there is no antecedent containment to be resolved, as in (68) (= (64a)). Given what we know about QR, this is a puzzle.[28]

(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 structure: 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 structure: [[-er [λd Mary's boss tells her to work d-hard]] [λd her father tells her to work d-hard]]

    The Condition C–Scope Generalization follows from our proposal. The merger 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 undergoes QR 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
        The degree clause is merged low, at the level of the embedded clause.
        LF structure: her father tells her [[λd PRO to work d-hard] [-er [λd Mary's boss works d-hard]]]
    b. *tell > -er, elided material = tell her to work d-hard, ruled out by conditions on ellipsis resolution
        The degree clause is merged low, at the level of the embedded clause.
        LF structure: her father tells her [[λd PRO to work d-hard] [-er [λd Mary's boss tells her to work d-hard]]]

---

[28] This puzzle is potentially related to Kennedy's (1997) observation, reported by Fox (2000:185), that there is no contrast between (ia) and (ib).

(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 in (ia), there is no antecedent containment to be resolved in (ib). However, QR is still able to bleed Condition C.

    c.   *-er > tell,* elided material = *work d-hard,* coreference possible, no Condition C violation
       The degree clause is merged high, at the level of the matrix clause.
       LF structure: [[λ*d* her father tells her to work *d*-hard] [-er [λ*d* Mary's boss works *d*-hard]]]
    d.   *-er > tell,* elided material = *tell her to work d-hard,* coreference possible, no Condition C violation
       The degree clause is merged high, at the level of the matrix clause.
       LF structure: [[λ*d* her father tells her to work *d*-hard] [-er [λ*d* 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
       The degree clause is merged low, at the level of the embedded clause.
       LF structure: her father tells her [[λ*d* PRO to work *d*-hard] [-er [λ*d* Mary's boss tells her to work *d*-hard]]]
    b.   *-er > tell,* coreference possible, no Condition C violation
       The degree clause is merged high, at the level of the matrix clause.
       LF structure: [[λ*d* her father tells her to work *d*-hard] [-er [λ*d* Mary's boss tells her to work *d*-hard]]]

6.3 Trapped Ellipsis

In all the cases 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, 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.

(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 (72) 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 latter undergoes covert movement. In other words, the degree clause has always been merged in a scope position. However, our analysis does not require this. If we assume, as is plausible, that there are DP-internal scope positions (see Heim and Kratzer 1998:sec. 8.5), 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's (2000) analysis of degree constructions and use QR as she does to resolve antecedent containment. However, for reasons that will become clear, we believe this is the wrong way to go. Instead, the derivation of the examples in (72) adopts a different strategy, involving 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) (from Fox 2002:88).

(74) a. I gave a book yesterday to Mary.
     b. I gave everyone yesterday a book on 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 an additional rightward movement of syntactic material over the degree clause has taken place. The derivations for (72) are indicated in (75). (Words in italics are unpronounced copies.)

(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. *Merger 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. *Merger 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. *Merger 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]]

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 move together by QR to resolve antecedent containment.

Since the examples in (71) do not involve even apparent antecedent containment, derivations like the ones in (75) are not obligatory 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 undergo object shift to a position above the position where the ECM assigner is interpreted, then an alternative 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 degree head adjoined to a DP and not, as has been the case elsewhere, a degree head adjoined to a clausal projection.

Our analysis of trapped ellipsis, as in (72), crucially depends on the availability 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 ungrammaticality should result.

Fox (2001) notes that one environment where rightward movement is not permitted involves finite VPs. An embedded finite VP is 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, as shown in (76) (from Fox 2001).

(76) 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 structure, shown in (78), would be illegitimate as it would involve vacuous quantification.

(78) LF structure: *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. *Merger of degree clause*
        [[Mary desires that more people] [more [than Bill desires people take syntax]]]
    d. LF structure: [[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), a grammatical sentence results.

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's (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 Merger and the Nonconservativity 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. However, 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? Why can the degree

clause, a complement, 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 shown that the ''at least as high'' generalization follows from the mechanism of countercyclic merger developed by Fox and Nissenbaum (1999) and Fox (2002), but so far we have provided no comparable explanation for the ''exactly as high'' generalization.

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 nonconservativity.

7.1 Background: The Interpretation of Copies

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

(80) *Trace Conversion*
    a. *Variable Insertion*
       (Det) Pred $\rightarrow$ (Det) [Pred $\lambda y\ (y = x)$]
    b. *Determiner Replacement*
       (Det) [Pred $\lambda y\ (y = x)$] $\rightarrow$ the [Pred $\lambda y\ (y = x)$]

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 containing a free variable. This free variable is bound by a higher copy. 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 LF structures 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)
    $= Q(A)(B(\text{the}[A\ x]))$

He points out that given Trace Conversion, nonconservative quantifiers could only have trivial meanings.

(83) a. Only Norwegians are tall. (→, QR)
      b. [only Norwegians] [[only Norwegians] are tall] (→, Trace Conversion)
      c. [only Norwegians] λx [the Norwegians x are tall]
      d. Only Norwegians are such that those Norwegians are tall. (tautology)

7.2 The Semantics of -er

While Trace Conversion yields the expected meaning when applied to conservative quantifiers, applying it to nonconservative quantifiers yields a trivial meaning.[29] Note that -er is a nonconservative 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 nonconservativity 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 such that John is d-tall and Bill is not d-tall.)

(86) $[\![\text{-er}]\!](A)(B) \not\Leftrightarrow [\![\text{-er}]\!](A)(A \cap B)$
      a. $[\![\text{-er}]\!](A)(B) = 1$ iff $A \subset B$ (contingent)
      b. $[\![\text{-er}]\!](A)(A \cap B) = 1$ iff $A \subset A \cap B$ (contradiction)

In (87), we consider the case where the first argument of -er, the degree clause, is merged with -er in its base position. The two then undergo QR together, leaving behind a copy. The LF structure after Trace Conversion is shown 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 [λd₁ Bill is d_1 -tall] d] tall]

---

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

The LF structure in (87c) yields a contradiction. The second argument of *-er* in (87c) can only be a subset of the first argument. Yet given the semantics of *-er,* the first argument must be a proper subset of the second.[30]

7.3 Late Merger Avoids the Contradiction

We now consider the option of merging the complement of *-er* late—that is, 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 merger 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 structure 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.* Such merger is ruled out because it leads to a contradictory LF.[31] The explanation based on conservativity also explains why *-er* and the degree

---

[30] It is not clear that equatives face a similar problem with conservativity (Daniel Büring, personal communication). The received wisdom concerning equatives is that their semantics involves subsethood and not equality. Subsethood does not lead to nonconservativity, while equality does. An alternative is to explore a semantics for the equative that involves equality along with a covert multiplier that is always greater than or equal to one. With this semantics, early merger of the degree clause when the multiplier is greater than one leads to a structure that is trivially false. When the multiplier is equal to one, the facts are more complex. We have a nonconservative quantifier, but not one where early merger leads to a structure that is trivially true or false. Instead, the early merger structure yields an 'at least as' meaning, while the late merger structure yields an 'exactly as' meaning. We have not examined the empirical facts concerning equatives, and so we leave this topic for further research.

[31] As noted in footnote 10, there is one exception to the observation that complements of *-er* are obligatorily extraposed. It involves degree phrases of the form *than n.*

   (i) Mary read [more than 3] 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 tempting to relate the acceptability of (i) to the fact that if *more than n* is treated as a complex determiner, it has conservative semantics. However, there are reasons to believe that *more than n* can still behave like 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 nonobligatoriness of extraposition for *than n* is potentially related to the fact that it does not obey the ''exactly as high'' generalization, as shown by the availability of a split reading in (ii).

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 merger of degree clause in the base position of *-er* followed by QR of [-er + degree clause]
    b. Late merger of degree clause after QR of *-er* to a scope position, followed by further QR of [-er + degree clause]

We have now derived the following generalization:

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

7.4 Why Is Late Merger of Certain Complements Possible?

We now consider why the degree clause, a complement, can be merged late. Lebeaux (1990) and Fox and Nissenbaum (1999) show 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 can be merged late, complements must be merged with their selecting head.

To investigate 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 that is affected by the Determiner Replacement part of Trace Conversion. In contrast, the complement in (91b) is the complement of *-er,* which is the element replaced by Determiner Replacement. Trace Conversion replaces *-er* and *which* with *the,* but leaves *rumor* untouched. Now, if the complement of *rumor* undergoes late merger, an illegitimate LF structure results.

(92) LF structure with late merger: [which rumor that John liked Mary] λx [he denied [the rumor x]]
    Higher *rumor* is of type $\langle e, e, t \rangle$, lower *rumor* is of type $\langle e, t \rangle$.

---

[32] An alternative explanation for the ''exactly as high'' part of the generalization follows from a generalization according to which overt Ā-movement blocks further covert Ā-movement (see Aoun, Hornstein, and Sportiche 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 merger of degree clauses/relative clauses. Hence, the constituent formed after the degree head merges 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, this failure would argue against the alternative proposal. See the discussion in footnote 25.

For (92) to be well formed type-theoretically, the higher copy of *rumor* would have to be of type $\langle e, e, t \rangle$ and the lower copy of type $\langle e, t \rangle$. Since two copies of an element cannot be distinct in type, (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

In this article, we have 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 (see, e.g., Bobaljik 2002, Pesetsky 2000, Chomsky 2001, Fox and Nissenbaum 1999, Fox 2002). This model of grammar makes it possible to interleave overt and covert movement. Thus, covert movement of the degree head can precede overt merger of the degree clause with the degree head.

In addition to accounting for the role played by the degree clause in determining the interpretation of a degree construction, our proposal accounts for 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 nonresolution of an antecedent-contained ellipsis. We find parallel examples with canonical cases of ACD (one of these is (94), from Tiedeman 1995 and Wilder 1995, among others, via Fox 2002).

(94) a. *I expect 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's (2002) proposal. The one difference is that degree clauses are syntactically and semantically

---

[33] We exploit the fact that Trace Conversion allows for late merger of the complements of elements that are replaced by Determiner Replacement to explain why the complement of *-er* can be merged late. This raises the following problem, noted by Fox (2002:fn. 11) and attributed by Fox to Uli Sauerland: why can't NPs be merged late 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, such as PP and CPs, can be extraposed. NPs need to be Case-licensed and hence cannot be extraposed. A related idea is that NPs 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.

complements of the degree head, while relative clauses are syntactically adjuncts and semantically modifiers of the NP inside the QP.

We conclude 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 that of English. In particular, it would be useful 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''; that is, 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 better understand the nature of the movement of the degree head.

References

Abney, Steven. 1987. The English noun phrase in its sentential aspect. Doctoral dissertation, MIT, Cambridge, Mass.

Andrews, Avery. 1985. *Studies in the syntax of relative and comparative clauses.* New York: Garland.

Aoun, Joseph, Norbert Hornstein, and Dominique Sportiche. 1981. Some aspects of wide scope quantification. *Journal of Linguistic Research* 1:69–95.

Beck, Sigrid. 1996. *Wh*-constructions and transparent Logical Form. Doctoral dissertation, Universität Tübingen.

Bobaljik, Jonathan David. 1995. Morphosyntax: The syntax of verbal inflection. Doctoral dissertation, MIT, Cambridge, Mass.

Bobaljik, Jonathan David. 2002. A-chains at the PF-interface: Copies and ''covert'' movement. *Natural Language & Linguistic Theory* 20:197–267.

Bresnan, Joan. 1973. The syntax of the comparative clause construction in English. *Linguistic Inquiry* 4:275–343.

Bresnan, Joan. 1975. Comparative deletion and constraints on transformations. *Linguistic Analysis* 1:25–74.

Carlson, Gregory N. 1977. Amount relatives. *Language* 53:520–542.

Chomsky, Noam. 1965. *Aspects of the theory of syntax.* Cambridge, Mass.: MIT Press.

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

Chomsky, Noam. 1993. A minimalist program for linguistic theory. In *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger,* ed. by Kenneth Hale and Samuel Jay Keyser, 1–52. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 2001. Beyond explanatory adequacy. MIT Occasional Papers in Linguistics 20. Cambridge, Mass.: MIT, Department of Linguistics and Philosophy, MITWPL.

Corver, Norbert. 1990. The syntax of left branch extractions. Doctoral dissertation, Tilburg University.

Corver, Norbert. 1993. A note on subcomparatives. *Linguistic Inquiry* 24:773–781.

Cresswell, Maxwell. 1976. The semantics of degree. In *Montague Grammar,* ed. by Barbara Partee, 261–292. New York: Academic Press.

Diesing, Molly. 1992. *Indefinites.* Cambridge, Mass.: MIT Press.

Embick, David, and Rolf Noyer. 2001. Movement operations after syntax. *Linguistic Inquiry* 32:555–598.

Fiengo, Robert, and Robert May. 1994. *Indices and identity.* Cambridge, Mass.: MIT Press.

Fox, Danny. 2000. *Economy and semantic interpretation.* Cambridge, Mass.: MIT Press.

Fox, Danny. 2001. The syntax and semantics of traces. Handout of talk, University of Connecticut, Storrs, November.

Fox, Danny. 2002. Antecedent-contained deletion and the copy theory of movement. *Linguistic Inquiry* 33: 63–96.

Fox, Danny, and Jon Nissenbaum. 1999. Extraposition and scope: A case for overt QR. In *WCCFL 18,* ed. by Sonya Bird, Andrew Carnie, Jason D. Haugen, and Peter Norquest, 132–144. Somerville, Mass.: Cascadilla Press.

Guéron, Jacqueline, and Robert May. 1984. Extraposition and Logical Form. *Linguistic Inquiry* 15:1–32.

Hackl, Martin. 2001. Comparative quantifiers. Doctoral dissertation, MIT, Cambridge, Mass.

Hankamer, Jorge. 1973. Why there are two *than*'s in English. In *Papers from the Ninth Regional Meeting of the Chicago Linguistic Society,* ed. by Claudia Corum, T. Cedric Smith-Stark, and Ann Weiser, 179–191. Chicago: University of Chicago, Chicago Linguistic Society.

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

Heim, Irene. 2000. Degree operators and scope. In *SALT X,* ed. by Brendan Jackson and Tanya Matthews, 40–64. Ithaca, N.Y.: Cornell University, CLC Publications.

Heim, Irene. 2002. Notes on Williams 1974, Fox & Nissenbaum 1999, Fox 2002, and Bhatt & Pancheva ms. 2002. Handout of lecture, MIT, Cambridge, Mass., November 25.

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

Hellan, Lars. 1981. *Towards an integrated theory of comparatives.* Tübingen: Gunter Narr.

Izvorski, Roumyana. 1995. A DP-shell for comparatives. In *Proceedings of ConSole III,* ed. by Alejandro Bisetti, Laura Bruge, João Costa, Rob W. N. Goedemans, Nicola Munaro, and Ruben van de Vijver, 99–121. The Hague: Holland Academic Graphics.

Kennedy, Christopher. 1997. Antecedent-contained deletion and the syntax of quantification. *Linguistic Inquiry* 28:662–688.

Kennedy, Christopher. 1999. *Projecting the adjective: The syntax and semantics of gradability and comparison.* New York: Garland.

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

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

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

Lechner, Winfried. 1999. Comparatives and DP-structure. Doctoral dissertation, University of Massachusetts, Amherst.

Moltmann, Friederike. 1992. Coordination and comparatives. Doctoral dissertation, MIT, Cambridge, Mass.

Napoli, Donna Jo. 1983. Comparative ellipsis: A phrase structure analysis. *Linguistic Inquiry* 14:675–694.

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

Postal, Paul. 1974. *On raising.* Cambridge, Mass.: MIT Press.

Rullmann, Hotze. 1995. Maximality in the semantics of *wh*-constructions. Doctoral dissertation, University of Massachusetts, Amherst.

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

Sag, Ivan. 1976. Deletion and Logical Form. Doctoral dissertation, MIT, Cambridge, Mass.

Selkirk, Elisabeth. 1970. On the determiner systems of noun phrases and adjective phrases. Ms., MIT, Cambridge, Mass.

Sportiche, Dominique. 1977. Reconstruction and constituent structure. Handout of talk, MIT, Cambridge, Mass.

Sportiche, Dominique. 1999. Reconstruction, constituency and morphology. Paper presented at the GLOW Colloquium, Berlin.

Stateva, Penka. 1999. In defense of the movement theory of superlatives. In *Proceedings of the Eastern States Conference on Linguistics 1999,* ed. by Rebecca Daly and Anastasia Riehl, 215–226. Ithaca, N.Y.: Cornell University, CLC Publications.

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

Stechow, Arnim von. 1993. Die Aufgaben der Syntax. In *Syntax: An international handbook of contemporary research,* ed. by Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld, and Theo Vennemann, 1–88. Berlin: Walter de Gruyter.

Stechow, Arnim von. 1998. Some remarks on Ch. Kennedy's ''Syntax and semantics of comparison.'' Ms., Universität Tübingen. Presented at the workshop ''Syntax and Semantics of Comparison,'' Berlin.

Taraldsen, Tarald. 1981. The theoretical interpretation of a class of marked extractions. In *Theory of markedness in generative grammar,* ed. by Adriana Belletti, Luciana Brandi, and Luigi Rizzi, 475–516. Pisa: Scuola Normale Superiore.

Tiedeman, Robyne. 1995. Some remarks on antecedent contained deletion. In *Minimalism and linguistic theory,* ed. by Shosuke Haraguchi and Michio Funaki, 67–103. Tokyo: Hituzi Syobo.

Wilder, Chris. 1995. Antecedent containment and ellipsis. In *FAS papers in linguistics,* 132–165. Berlin: FAS (Forschungsschwerpunkt allgemeine Sprachwissenschaft, Typologie, und Universalienforschung).

Williams, Edwin. 1974. Rule ordering in syntax. Doctoral dissertation, MIT, Cambridge, Mass.

Wold, Dag. 1995. Antecedent-contained deletion in comparative constructions, Ms., MIT, Cambridge, Mass.

(Bhatt)
Department of Linguistics
Calhoun 501
University of Texas at Austin
1 University Station, B5100
Austin, Texas 78712-1196

rbhatt@mail.utexas.edu

(Pancheva)
3601 Watt Way
Grace Ford Salvatori Hall 301
University of Southern California
Los Angeles, California 90089-1693

pancheva@usc.edu