

## Comparative (Sub)deletion and Ranked, Violable Constraints in Syntax

Christopher Kennedy

Northwestern University

### 0. Introduction

This paper investigates the syntax of comparative deletion and comparative subdeletion and argues that the apparently paradoxical behavior of these two English comparative constructions can be explained by a simple distinction between overt and covert movement: comparative deletion involves overt movement of a compared phrase, while comparative subdeletion involves covert movement of the same element. This derivational difference is argued to follow from general constraints on the relation between movement and deletion in English in a model of syntax in which such constraints are ranked and violable. Finally, this approach is demonstrated to be superior to alternative analyses that do not make reference to ranked and violable constraints, and to receive independent support from the interaction of comparatives and VP-deletion.

### 1. Perspectives on the Syntax of Comparison in English

#### 1.1 Towards a Uniform Analysis of Comparatives

*Comparative deletion* (CD) is the term introduced by Bresnan (1973, 1975) to describe expressions of comparison such as those in (1), which compare two quantities of the same sort of stuff (number of games, degrees of height, degrees of carefulness).

- (1) a. The galaxy contains more stars than the eye can see.  
b. At that time, sea level was not as high as it later became.  
c. My sister drives as carefully as I drive.

Somewhat more exotic, but equally well-attested, are examples of *comparative subdeletion* (CSD) such as the naturally occurring sentences in (2), which compare the relative quantities of different sorts of stuff (number of scoring titles vs. number of tattoos, degrees of length vs. degrees of thickness, degrees of carefulness vs. degrees of carelessness).

- (2) a. [Michael Jordan] has more scoring titles than Dennis Rodman has tattoos. (*Chicago Tribune* 7.17.98)  
 b. The shapes seem to be longer than they are thick.... (Bear, Greg, 1997, *Slam*, New York, Tor, p. 262)  
 c. My sister drives as carefully as I drive carelessly. (from Bresnan 1975:47)

As the traditional names for these constructions suggest, both require some element to be omitted from the clausal complement of 'than' or 'as' (henceforth the *comparative clause*). In the case of CD, this element corresponds to the constituent marked by the comparative morphology in the main clause, as illustrated by the examples in (3); in the case of CSD, an amount or degree term must be left out, as shown in (4).<sup>1</sup>

- (3) a. \*The galaxy contains more stars than the eye can see stars.  
 b. \*At that time, sea level was not as high as it later became high.  
 c. \*My sister drives as carefully as I drive carelessly.

- (4) a. \*Michael has more scoring titles than Dennis has two/many tattoos.  
 b. \*The shapes seem to be longer than they are 2 inches/thick.  
 c. \*My sister drives as carefully as I drive so/very carelessly.

Building on the intuition that the similar "deletion" requirements of these two forms of comparison reflect a deeper syntactic connection between the two constructions (in particular, that CD simply involves more deletion than CSD), researchers dating back at least to Lees 1961 have postulated that the basic principles of comparative formation in English generate representations that are structurally CSD configurations, and that the omission of additional material in CD can be derived from general principles of redundancy reduction. On this view, the sentences in (1) are derived from representations like (3a-c).<sup>2</sup>

<sup>1</sup>One exception to the obligatory deletion requirement in CD involves examples in which the compared constituent bears contrastive focus, as in (i), where capitalization indicates focal stress (Chomsky 1977, ex. (247)). This fact follows from the analysis I will develop in section 2 (see note 12).

- (i) A: This desk is higher than that one is wide.  
 B: What is more, this desk is higher than that one is HIGH.

<sup>2</sup>A variety of other constituents can also be omitted from the comparative clause, resulting in so-called *comparative ellipsis* structures. Following Napoli 1983, I will assume that examples of comparative ellipsis like (1a-d) involve either CD or CSD plus the various independently-motivated ellipsis operations listed below (see also Lechner 1999), and I will focus here on the syntactic analysis of the more basic CD/CSD configurations.

- (i) a. The galaxy contains more stars than the solar system does. *VP deletion*  
 b. The galaxy contains more stars than the solar system. *Stripping*  
 c. The galaxy contains more stars than anyone thought. *Null complement anaphora*  
 d. The galaxy contains more stars than it does — planets. *Pseudogapping*

In addition, I will not consider comparatives with *of* in this paper, such as (ii). Following Grimshaw 1987, I will assume for now that these involve CD plus extraposition of the *of*-PP, though this assumption deserves closer scrutiny in future work.

- (ii) There were more of the planets in the photo than there were of the stars.

The hypothesis that CD and CSD reflect a single rule of comparative formation is most fully developed in Bresnan 1975. Bresnan argues that the rule of comparative formation in English is an unbounded deletion operation that obligatorily eliminates a degree term from the phrase in the comparative clause that is compared with the morphologically comparative constituent in the main clause. (I will henceforth refer to the former as the *compared constituent* and the latter as the *head* of the comparative.) Like other transformations, this operation is subject to the relativized A-over-A condition, which requires the additional removal of as much redundant material as possible, up to recoverability (see Bresnan 1975:68). The result is that in CD, where the compared constituent is fully identical with the head, the entire phrase must be deleted. In CSD, however, the lexical component of the compared constituent is distinct from the head, so only the degree term may be deleted. This is illustrated in (5) and (6), where brackets indicate the constituents that count as identical in Bresnan's analysis.

- (5) a. The galaxy contains [more stars] than the eye can see [~~x many stars~~]  
 b. Sea level was not [as high] as it later became [~~x much high~~]
- (6) a. Michael has [more] scoring titles than Dennis has [~~x many~~] tattoos  
 b. The shapes are long[er] than they are [~~x much~~] thick

Subsequent analyses, building on the observation that CD and CSD have properties similar to A-bar movement constructions (Ross 1967, Chomsky 1977), recast Bresnan's analysis in terms of movement or binding of a degree term.<sup>3</sup> While there are important differences between these various approaches, they all share the assumption that CD and CSD are derived in fundamentally the same way, and so predict that both types of comparatives should show a clear set of similar properties, particularly properties relating to movement or unbounded deletion.

Initial confirmation of this prediction comes from island effects. First, both CD and CSD are ill-formed when the deletion site is embedded in an extraction island (see Ross 1967, Chomsky 1977, and Postal 1998), as shown by the examples in (7)-(9), which illustrate complex-NP, *wh*-, and adjunct islands, respectively.

- (7) a. \*Michael has more scoring titles than Dennis is a guy who has.  
 b. \*Michael has more scoring titles than Dennis is a guy who has tattoos.
- (8) a. \*The shapes were longer than I wondered whether they would be.  
 b. \*The shapes were longer than I wondered whether they would be thick.
- (9) a. \*My sister drives as carefully as I avoid accidents when I drive.  
 b. \*My sister drives as carefully as I get into accidents when I drive carelessly.
- A second piece of evidence for a uniform analysis of CD and CSD comes from crossover effects. (10) demonstrates that both CD and CSD show strong crossover effects, and (11) makes the same point for weak crossover.

---

<sup>3</sup>For example, Pinkham 1982 postulates direct binding of a degree term for both CD and CSD and insertion of null pronominal elements in CD, while Heim 1985, Izvorski 1995, and others analyze CD and CSD as *wh*-movement of a degree term and CD as additional deletion of redundant material (though Izvorski makes a novel proposal regarding the nature of the moved degree term; see note 6).

- (10) a. More Democrats<sub>i</sub> voted than anyone/\*they<sub>j</sub>! expected to vote.  
 b. More Democrats<sub>i</sub> voted than anyone/\*they<sub>j</sub>! expected Republicans<sub>j</sub> to vote.
- (11) a. More Democrats<sub>i</sub> voted than anyone/\*their<sub>j</sub>! friends expected to vote.  
 b. More Democrats<sub>i</sub> voted than anyone/\*their<sub>j</sub>! friends expected Republicans<sub>j</sub> to vote.

If sensitivity to islands and crossover effects are indicative of a similar derivational history, a point on which both movement and unbounded deletion analyses mainly agree, then facts like the ones illustrated here provide strong support for the view that CD and CSD should be analyzed in the same way.

A third, somewhat weaker, argument in favor of a uniform analysis comes from the interpretation of CD and CSD. In most contexts, CD and CSD constructions have exactly the same type of truth conditions: both involve comparison of two amounts, differing only in that CD compares amounts of the same sort of stuff, while CSD compares amounts of different sorts of stuff.<sup>4</sup> Both the CD and CSD options in (12a) and (13a), for example, can be assigned the interpretations paraphrased in (12b) and (13b) (see in particular Heim 1985, who explicitly discusses the semantic transparency of CSD and uses it as a basis for handling the interpretation of CD).

- (12) a. Michael has more scoring titles than Dennis has (tattoos).  
 b. *the number of MJ's scoring titles > the number of DR's scoring titles/tattoos*
- (13) a. Michael's hands are as wide as my feet are (long)  
 b. *the width of MJ's hands = the width/length of my feet*

While truth-conditional equivalence (in the relevant sense) is not a sufficient condition for a uniform analysis, it is a necessary one: if CD and CSD have essentially the same syntactic derivations (and involve the same functional vocabulary), then the observed semantic similarity would follow.

## 1.2 Against a Uniform Analysis

The hypotheses that CD and CSD can be given a uniform analysis has received a strong challenge in recent years with the identification of a large set of empirical phenomena that clearly differentiate between the two types of constructions and suggest that they are syntactically distinct in ways that go beyond the superficial difference in the amount of omitted material. These facts have led a number of researchers to develop non-uniform analyses of comparatives, which differ in their implementations, but typically share the

---

<sup>4</sup>Examples of CSD involving adjectives of opposite polarity, such as (i), are an exception.

- (i) Michael's hands are as wide as my feet are short.

(i) does not assert that Michael's hands and my feet are equal in measurement (cf. (13)). Instead, it asserts that Michael's hands exceed a norm of width to the same degree that my feet exceed a norm of shortness (e.g., "Michael's hands are very wide and my feet are very short"). As shown in Kennedy (to appear), this interpretation stems from the interaction of the semantics of comparison and adjectival polarity.

assumption that CD is derived in one of the ways specified above (unbounded deletion or some kind of A'-movement/ellipsis operation), while CSD is different.<sup>5</sup>

The crucial empirical differences between CD and CSD can be summarized as follows. First, in languages that prohibit preposition stranding, such as Czech, CD behaves like movement constructions and obeys this constraint, but CSD does not. (Thanks to Hana Filip for supplying the Czech data; this point was originally made in Corver 1990 for Dutch.) This is illustrated by the contrast between (15a) and (15b); (14) demonstrates the unacceptability of preposition stranding in a comparable Czech question.

(14) \*Kteryich me&stech Vaclav bydlel ve?  
 which city.PL.LOC Vaclav live.PAST.3SG in  
 'Which city does Vaclav live in?'

(15) a. \*Bydlel jsem ve vice me&stech nez& ty jsi bydlel v  
 live.PAST.1SG aux in more city.PL.GEN than you aux live.PAST.2SG in  
 'I have lived in more cities than you have lived in.'  
 b. Chci bydlel ve vice americkyich me&stech nez& jsem  
 want.1SG.PRES live.INF in more American city.PL.GEN than aux  
 bydlel v europskyich me&stech  
 lived.PAST.1SG in European city.PL.LOC  
 'I want to have lived in more American cities than I have lived in European cities.'

Second, CD shows COMP-trace effects in English, but CSD does not (Grimshaw 1987), as shown by the contrast between the (a) and (b) sentences in (16) and (17).

(16) a. More books were published than the editor said (\*that) would be.  
 b. More boys flunked than I predicted (\*that) would pass.

(17) a. More books were published than the editor said (that) articles would be.  
 b. More boys flunked than I predicted (that) girls would pass.

Third, like other types of movement and deletion operations, CD blocks contraction of an immediately preceding auxiliary; contraction before a CSD site is perfectly acceptable, however (Grimshaw 1987):

(18) a. I thought there was more meat than there is/\*'s.  
 b. John was more upset than he is/\*'s now.  
 c. She was as happy about it then as she is/\*'s now.

(19) a. There's more meat than there's rice.  
 b. John was more upset than he's angry now.  
 c. She was as happy about it then as she's sad now.

<sup>5</sup>For example, Grimshaw (1987) claims that CSD structures are base-generated in their surface form; Corver (1993) and Hendriks (1995) analyze CSD in terms of across-the-board movement; Kennedy (1998, 1999) claims that CD and CSD constructions are headed by different *more*'s, which select for syntactically distinct comparative clauses; Chomsky (1977) and Knowles (1984) claim that constituent targeted by *wh*-movement in CSD is different from that targeted by *wh*-movement in CD, while Rivero (1981) argues that it is the landing site of the moved constituent in CSD that differentiates it from CD.

Fourth, CD licenses parasitic gaps, a fact that has been taken as further evidence for its status as a type of *wh*-movement construction (see Postal 1998 for recent discussion), but CSD does not (Grimshaw 1987):

- (20) a. I threw away more books than I kept without reading.  
 Jerome followed more suspects than Arthur interrogated without arresting.
- (21) a. \*I threw away more books than I kept magazines without reading.  
 \*Jerome followed more leads than Arthur interrogated suspects without arresting.

A fifth contrast comes from so-called *multiply-headed comparatives*, such as the CSD constructions in (22).

- (22) a. Christmas makes as many children as happy as it makes adults unhappy.  
 Max persuaded more men to buy more cars than you persuaded women to buy trucks.

The semantic analysis of multiply-headed comparatives is exceedingly complex (see von Stechow 1984 and Hendriks 1995 for discussion), involving multiple instances of comparison. (22a), for example, has the interpretation in (23).

- (23) The number of children that Christmas makes happy equals the number of adults that it makes unhappy and the degree to which Christmas makes children happy equals the degree to which it makes adults unhappy.

The syntactic properties of these constructions are relatively clear, however: multiply-headed CD constructions that are structurally parallel to the examples of CSD in (22) are unacceptable (Corver 1990, 1993, Hendriks 1995):

- (24) a. \*Christmas makes as many children as happy as birthdays make.  
 b. \*Max persuaded more people to buy more cars than you persuaded to buy.

A final argument for treating CD and CSD differently is relevant only to accounts that seek to analyze both constructions in terms of movement. According to such analyses, CSD involves A'-movement of a degree term from inside the nominal or adjectival projection to SpecCP, as in (25). (I assume adjectival projections to be Degree Phrases (DegPs); see Abney 1987, Corver 1990, Grimshaw 1991, and Kennedy 1999.)

- (25) a. Michael has more scoring titles than [CP *Op*! Dennis has [DP *t* tattoos]]  
 b. The shapes are longer than [CP *Op*! they are [DegP *t* thick]]

The problem with this proposal is that such movement is impossible when the displaced phrase has phonological content, as shown by the examples in (26), which violate the "Left Branch Constraint" (see Ross 1967).

- (26) a. \* [CP How many<sub>i</sub> does Dennis have [DP *t* tattoos]]?  
 b. \* [CP How (much)<sub>i</sub> were the shapes [DegP *t* thick]]?

Chomsky (1977:123) suggests that the examples in (25) may involve feature movement only, and so may somehow avoid the Left Branch Constraint (a proposal recently revived

by Donati (1998)). Given Corver's (1990) arguments that the moved elements in (26) are heads, however, and that the actual reason that (26a-b) are unacceptable is that they violate the Head Movement Constraint, Chomsky's proposal is untenable: movement of formal features alone in (25a-b) would violate Generalized Pied Piping (see Chomsky 1995:262-264). Thus the apparent absence of left branch in CSD effects remains a problem for most movement-based approaches to comparatives.<sup>6</sup>

### 1.3 Summary

The facts discussed in this section indicate that while there are compelling arguments in favor of assigning similar syntactic analyses to CD and CSD, there are also clear empirical differences between the two types of constructions that must be resolved in order to maintain a minimum level of descriptive adequacy. One conclusion that could be drawn from this is that CD and CSD have distinct syntactic representations; indeed, this is the position adopted by the researchers who have addressed the data discussed in section 1.2 (see note 5). An alternative conclusion is that CD and CSD are the same in their basic syntactic properties – both types of comparative involve the same functional vocabulary and are subject to the same syntactic operations – but that they differ with respect to how exactly these operations apply. On this view, the facts observed in the previous section would be explained not in terms of differences in representation, but rather in terms of differences in derivation. This is the hypothesis that I will pursue in the next section.

## 2. A New Look at an Old Idea

### 2.1 Overt vs. Covert Movement in Comparatives

The goal of this section is to present and motivate a uniform analysis of comparatives that accounts for the apparently paradoxical properties of CD and CSD in a principled way. I begin by outlining the proposal and showing how it explains the facts, and then in section

<sup>6</sup>Precisely this problem constitutes Bresnan's (1975) central argument against a movement analysis and in favor of the unbounded deletion approach, which she claims is not subject to the Left Branch Constraint. Pinkham (1982) provides compelling arguments against an unbounded deletion analysis, however, demonstrating that it overgenerates in comparatives involving attributive adjectives (see also Kennedy and Merchant 2000). Izvorski (1995) develops a movement account that does not suffer from this problem, because it claims that the term targeted by movement in comparatives (both CD and CSD) is not syntactically parallel to *how (many/much)* in (26), but rather to the adjuncts in *what quantity* and *to what degree* in (i).

- (i) a. In what quantity does Dennis have tattoos?
- b. To what degree were the shapes thick?

While this analysis also succeeds in explaining some of the differences between CD and CSD, it runs into problems with parasitic gaps. Parasitic gaps are licensed by movement; therefore, if CD and CSD involve the same sort of movement, we would expect both to license parasitic gaps, contrary to fact. More importantly, the type of movement that Izvorski assumes to be operating in CD does not license parasitic gaps, as shown by (ii), while the type of movement she rejects (targeting *how many*) does:

- (ii) a. \*In what quantity did you throw away books without reading?
- b. How many books did you throw away without reading?

2.2 I argue that my specific claims about comparatives follow from more general principles governing movement and deletion in a model in which syntactic constraints are ranked and violable.

I propose that comparatives in English – both CD and CSD – involve movement of the compared constituent to a clause-initial position (which we can assume for now to be SpecCP, though this assumption will be revised below), but that the two constructions differ in when this movement applies. The central claims are summarized in (27).

- (27) *English Comparative Formation* (version 1)
- i. CD involves overt movement of the compared constituent plus deletion under identity with the head of the comparative (cf. Chomsky 1977).
  - ii. CSD involves covert movement of the compared constituent.

Focusing on nominal and adjectival comparatives (I assume adverbial phrases to also be DegPs; see Larson 1987), this analysis assigns the structures in (28a-b) to (1a-b), where struck-through text indicates lexical material that is deleted from the PF representation.

- (28)
- a. The galaxy contains more stars than [CP [DP ~~stars~~] the eye can see *t<sub>i</sub>*]
  - b. Sea level was not as high as [CP [DegP ~~high~~] it later became *t<sub>i</sub>*]

Since movement is overt, the LF and PF representations of CD are the same (modulo deletion). Examples of CSD, on the other hand, differ at PF and LF, as shown by (29) and (30), where the (a) and (b) examples are PF and LF representations, respectively.

- (29)
- a. Michael has more scoring titles than [CP [DP Dennis has [DP tattoos]]
  - b. Michael has more scoring titles than [CP [DP tattoos]<sub>i</sub> Dennis has *t<sub>i</sub>*]
- (30)
- a. The shapes are longer than [CP they are [DegP thick]]
  - b. The shapes are longer than [CP [DegP thick]<sub>i</sub> they are *t<sub>i</sub>*]

More generally, the analysis claims that the two types of comparatives are structurally identical at LF, but they differ at PF. This result leads to two clear predictions: 1) CD and CSD should behave the same with respect to constraints on LF representations, and 2) all syntactic differences between the two types of comparatives should be localized to PF. As it turns out, the similarities and differences between the two constructions break down in exactly this way.

Let us consider first the similarities between CD and CSD. The strongest syntactic argument for a uniform analysis comes from extraction islands and crossover effects, two sets of phenomena that, in models that assume multiple levels of representation, have traditionally been analyzed as involving (at least) constraints on LF representations. If CD and CSD are the same at LF, it follows that they should show the same sort of behavior in these contexts.

The analysis also supports a single compositional semantic analysis for both constructions, allowing us to account for their similarity in meaning without positing multiple lexical entries for the comparative morphology (as in Hendriks 1995 and Kennedy 1998, 1999). Consider first the case of adjectival comparatives like the examples in (31), which have the LFs in (32a-b). (I take the impossibility of overt morphology (see (4a-c)) as an indication that the compared constituent is headed by a null morpheme, notated Deg<sup>0</sup><sub>[+C]</sub>.)

*Comparative (Sub)deletion*

- (31) a. Michael's hands are wider than my feet are.  
 b. Michael's hands are wider than my feet are long.
- (32) a. Michael's hands are wider than [CP [DegP Deg<sup>0</sup>[+C] wide]<sub>i</sub> my feet are *t*<sub>i</sub>  
 b. Michael's hands are wider than [CP [DegP Deg<sup>0</sup>[+C] long]<sub>i</sub> my feet are *t*<sub>i</sub>

Following Kennedy 1999, I assume that gradable adjectives denote functions from objects to degrees and combine with degree morphology to generate properties of individuals. The degree morphemes that head the compared constituent and the head of the comparative can then be assigned the interpretations in (33), respectively (*less* and *as* differ from *more* only in the nature of the ordering relation), where  $G$  is a function from objects to degrees,  $\bar{Q}$  is a function from properties to truth values (the semantic value of a clausal constituent with an extracted DegP), and  $t$  is an operator that returns the maximal element of an ordered set of degrees or amounts (see von Stechow 1984, Rullmann 1995, Kennedy 1997b).

- (33) a.  $\text{Deg}^{0[+C]} = \lambda G \lambda \bar{Q} [\lambda x [G(x) = d]]$   
 b.  $\text{er/more} = \lambda G \lambda d \lambda x [G(x) > d]$

Without going through the details of semantic composition (see Kennedy 1999), the interpretations that we end up assigning to (32a-b) are shown in (34), which are equivalent to the informal representations given above in (12b) and (13b).

- (34) a.  $\text{wide}(\text{michael's-hands}) < \text{id}[\text{wide}(\text{my-feet}) = d]$   
 b.  $\text{wide}(\text{michael's-hands}) < \text{id}[\text{long}(\text{my-feet}) = d]$

Nominal comparatives can be analyzed in essentially the same way. The examples in (35) have the LFs in (36).

- (35) a. MJ has more scoring titles than Dennis has.  
 b. MJ has more scoring titles than Dennis has tattoos.
- (36) a. MJ has more scoring titles than [CP [DP D<sup>0</sup>[+C] scoring titles]<sub>i</sub> Dennis has *t*<sub>i</sub>  
 b. MJ has more scoring titles than [CP [DP D<sup>0</sup>[+C] tattoos]<sub>i</sub> Dennis has *t*<sub>i</sub>

Building on the semantic similarity between the vague determiners *many* and *much* and gradable adjectives (see Klein 1980 for discussion), I will assume that part of the meaning of nominal degree morphology is a function MANY from plural objects to amounts. We can then assign the meanings in (37) to the compared constituent and *more*, respectively, where  $P$  is a (plural) NP meaning and  $\bar{Q}$  a function from plural objects to truth values (as above, the semantic value of a clausal constituent with an extracted DP), and  $t$  as above.

- (37) a.  $\text{D}^{0[+C]} = \lambda P \lambda \bar{Q} [\text{in}[\text{EX}[P(X) \ \& \ \text{MANY}(X) = n \ \& \ \bar{Q}(X)]]]$   
 b.  $\text{more} = \lambda P \lambda n \lambda \bar{Q} [\text{EX}[P(Y) \ \& \ \text{MANY}(Y) > n \ \& \ \bar{Q}(Y)]]]$

The interpretations that this analysis assigns to (35a-b) are given in (38), which accurately characterize the truth conditions of these sentences.

- (38) a.  $\text{EX}[\text{scoring-titles}(Y) \ \& \ \text{MANY}(Y) < \text{in}[\text{EX}[\text{scoring-titles}(X) \ \& \ \text{MANY}(X) = n \ \& \ \text{have}(\text{dennis}, X)]]]$   
 b.  $\text{EX}[\text{scoring-titles}(Y) \ \& \ \text{MANY}(Y) < \text{in}[\text{EX}[\text{tattoos}(X) \ \& \ \text{MANY}(X) = n \ \& \ \text{have}(\text{dennis}, X)]]]$

Moving to the properties that differentiate CD and CSD, the first three – the COMP-trace filter, P-stranding, and contraction – clearly involve conditions on overt but not covert movement, as shown by the fact that situ *wh*-phrases do not violate the relevant constraints:

- (39) a. Which editor said (that) how many books would be published this year?  
 b. Who predicted (that) how many boys would flunk?

- (40) Kdo bydlel ve kterých městech? who live.PAST.3SG in which city.PL.LOC  
 'Who lives in which city?'  
 Czech

- (41) a. Who said there's how much rice?  
 b. Who's how angry now?

While the details of these constraints need to be worked out (see e.g. Chomsky and Lasnik 1977 and Honegger 1996 for PF analyses of COMP-trace effects), at this point it is enough to observe that if CSD involves covert movement, then the fact that the examples in (39)-(41) are acceptable means that CSD should also be acceptable in the same environments. This is exactly what we saw in section 1.2.

Regarding multiply-headed comparatives, aside from the unique nature of the construction itself, the acceptability of CSD in this context is unsurprising, since the proposed analysis predicts that such constructions should be structurally analogous to multiple-*wh* questions. Since English allows multiple instances of covert A'-movement (assuming e.g. Higginbotham & May's (1981) analysis of multiple questions or May's (1985) proposals about Quantifier Raising), the LF postulated for a sentence like (42a), shown in (42b), should be well-formed.

- (42) a. Christmas makes as many children as happy as it makes adults unhappy.  
 b. Christmas makes as many children as happy as [CP [DP adults]<sub>i</sub>] [<sub>DegP</sub> unhappy]<sub>j</sub>; it makes *t<sub>i</sub>* *t<sub>j</sub>*

English typically does not allow multiple instances of *overt* A'-movement, however, as shown by examples like those in (43).

- (43) a. \*The children<sub>i</sub> quite happy<sub>j</sub> Christmas will make *t<sub>i</sub>* *t<sub>j</sub>*  
 b. \*Which children<sub>i</sub> how happy<sub>j</sub> will Christmas make *t<sub>i</sub>* *t<sub>j</sub>*

It follows that multiply-headed CD should also be impossible, since it would involve multiple A'-movement. This is illustrated by (44b), the PF assigned to (44a).

- (44) a. \*Christmas makes as many children as happy as birthdays make.  
 b. \*Christmas makes as many children as [CP [DP children]<sub>i</sub>] [<sub>DP</sub> happy]<sub>j</sub>; birthdays make *t<sub>i</sub>* *t<sub>j</sub>*

Although multiply-headed CD is ruled out, the analysis correctly predicts that "mixed" multiply-headed comparatives involving both CD and CSD, such as (45a-b), should be acceptable. Such examples involve only one instance of overt movement, namely the one that targets the deleted constituent, as shown in (46a-b).

Comparative (Sub)deletion

- (45) a. Christmas makes as many people as happy as it makes unhappy.  
 Max persuaded more people to buy more cars than you persuaded to buy trucks.
- (46) a. Christmas makes as many people as happy as [CP DP ~~people~~] it makes  $t_i$  [DegP unhappy] $_j$   
 b. Max persuaded more people to buy more cars than [CP DP ~~people~~] you persuaded  $t_i$  to buy [DP trucks] $_j$ .

The differing acceptability of parasitic gaps in CD and CSD also follows from the over/cover distinction. The fact that CD allows parasitic gaps is not surprising: the movement postulated to occur in (47a) is exactly the same as the movement in a *wh*-question like (47b); the only difference is that in the former, the moved constituent is deleted. (The analysis developed here thus avoids the problems for Izvorski's (1995) proposal discussed in note 6.)

- (47) a. I threw away more books than [DP ~~books~~] I kept  $t_i$  without reading  $e_i$   
 b. [DP How many books] $_i$  did you keep  $t_i$  without reading  $e_i$

A well-known property of parasitic gaps is that they are dependent on *overt* A'-movement (see Nissenbaum 1998 for a new analysis of this requirement). It follows that the compared constituent in (48a), like the *wh*-phrase in (48b), should fail to license a corresponding parasitic gap.

- (48) a. \*I threw away more books than [CP I kept [DP papers] $_i$  without reading  $e_i$ ]  
 b. \*Who kept [DP how many papers] $_i$  without reading  $e_i$

There is one context in which a parasitic gap may be associated with a phrase that is moved covertly, however. As shown by Nissenbaum (this volume), covert A'-movement can license a parasitic gap only if a first parasitic gap is already licensed by overt movement. This is illustrated by the contrasts in (49).

- (49) a. [DP Which senator] $_i$  did you convince  $t_i$  to buy [DP which car] $_j$  after getting an opponent of  $e_i$  to put a bomb in  $e_j$   
 b. \* [DP Which senator] $_i$  did you persuade  $t_i$  to buy [DP which car] $_j$  after putting a bomb in  $e_j$

CSD also license parasitic gaps in such contexts, providing further support for the claim that it involves covert A'-movement (many thanks to Jon Nissenbaum for bringing this fact to my attention):

- (50) a. I persuaded as many senators to buy as many cars as [CP DP ~~senators~~] you persuaded  $t_i$  to buy [DP trucks] $_j$  after getting opponents of  $e_i$  to put bombs in  $e_j$   
 b. \*I persuaded as many senators to buy as many cars as [CP DP ~~senators~~] you persuaded  $t_i$  to buy [DP trucks] $_j$  after putting bombs in  $e_j$

Both examples in (50) are "mixed" multiply-headed comparatives, in which one of the compared constituents undergoes overt movement (CD) and the other remains in situ (CSD). As shown by the contrast between (50a-b), the CSD phrase licenses a parasitic gap only if the CD phrase also licenses a parasitic gap, just as with *wh*-in situ in (49a-b).

Finally, the proposed analysis avoids the problems that arise with respect to the Left Branch Constraint, even though it postulates A'-movement for both CD and CSD. The crucial difference between the analysis proposed here and earlier movement analyses is that neither CD nor (crucially) CSD involve movement of a left branch degree term out of DP/DegP; instead, movement targets the entire compared constituent (cf. Rivero 1981). In other words, the comparatives in (51) are structurally analogous not to the questions in (52), as on standard movement analyses, but to those in (53) (or more accurately, questions in which these phrases remain in situ), which are perfectly well-formed.

- (51) a. Michael has more scoring titles than Dennis has tattoos.  
 b. The shapes are longer than they are thick.

- (52) a. \*How many<sub>t</sub> does Dennis have <sub>t</sub> tattoos?  
 b. \*How (much)<sub>t</sub> were the shapes <sub>t</sub> thick?

- (53) a. How many tattoos<sub>t</sub> does Dennis have <sub>t</sub>?  
 b. How thick<sub>t</sub> were the shapes <sub>t</sub>?

The end result is that the Left Branch Constraint, however it is formalized (see Kennedy and Merchant 2000 for a recent proposal), does not come into play.

## 2.2 Comparative (Sub)deletion and Optimality in Syntax

The previous section demonstrated that the analysis of comparatives summarized in (27) achieves the goal of explaining both the similarities and the differences between CD and CSD, achieving a level of descriptive adequacy not matched by earlier uniform approaches. However, the assumption that is crucial to achieving this result – the claim that CD involves overt movement plus deletion, while CSD involves covert movement – is, at this point, a stipulation. The question that must now be addressed is whether there is a principled explanation for this derivational difference between the two types of comparatives.

In fact, an analysis very similar to the one I have proposed here is considered by Bresnan (1975), who rejects it precisely because of its stipulative and apparently ad hoc nature (Borsley (1984:281) makes a similar objection). In particular, Bresnan (1975:63) points to the unacceptability of pairs like those in (54), and objects that “[t]o guarantee that only the maximally recoverable constituent is moved, one would have to place a special identity condition in the rule itself.... [O]n this analysis, it becomes accidental that the moved constituents undergo deletion and that the elements moved just happen to be those which would be maximally recoverable if deleted.”

- (54) a. \*The galaxy contains more stars than the eye can see stars.  
 b. \*The galaxy contains more stars than [DP planets]<sub>t</sub> the eye can see <sub>t</sub>.

Put another way, Bresnan’s criticism is that the fact that we get deletion whenever we have movement and identity, but neither movement nor deletion when we don’t have identity, remains unexplained under standard assumptions about constraints on movement.<sup>7</sup> What I will argue in the rest of this section is that Bresnan’s objection can be overcome by mod-

<sup>7</sup>In Bresnan’s original analysis, these facts are explained by the assumption that comparatives involve deletion, not movement (which rules out (54b)), and the relativized A-over-A condition (which rules out incomplete deletion in (54a)).


Comparative (Sub)deletion

Trying the standard assumptions about these constraints. In particular, I will show that by adopting a model in which the constraints governing movement and deletion are ranked and violable, and syntactic derivations are evaluated according to how well they satisfy such constraints, the derivational difference between CD and CSD follows. In short: deletion is good and overt movement is bad, but it's better to delete than to avoid overt movement.

Using the tools of Optimality Theory (Prince and Smolensky 1993), the basic idea can be implemented as follows.<sup>8</sup> First, assume two well-motivated constraints: DELETE, which requires minimization of the PF representation (cf. Economy of Representation); and STAY, which forbids overt movement (Grimshaw 1997; cf. Economy of Derivation, Procrastinate, etc.). The derivational difference between CD and CSD can then be derived by ranking these two constraints in English as in (55).

(55) DELETE >> STAY

Consider first the case of CD. As illustrated in Tableau 1, the ranking of DELETE over STAY has the result that when the head and compared constituent are identical, a candidate in which the compared constituent moves and deletes is preferred to options in which no deletion occurs.

	DELETE	STAY
a.  The galaxy contains more stars than [CP [DP <del>stars</del> ] <sub>i</sub> ] the eye can see <i>t</i> ]		*
b. The galaxy contains more stars than [CP [DP stars] <sub>i</sub> ] the eye can see <i>t</i> ]	*i	*
c. The galaxy contains more stars than [CP the eye can see [DP stars] <sub>i</sub> ]	*i	

This tableau is incomplete, however. A fourth candidate in which the compared constituent is deleted in situ, satisfying both DELETE and STAY, must also be considered:

d. The galaxy contains more stars than [CP the eye can see [DP <del>stars</del> ] <sub>i</sub> ]		
--	--	--

That (d) is not the optimal candidate is clear: the differences between CD and CSD indicate that CD involves overt movement. This suggests that some higher-ranked principle governing deletion in English rules out this candidate; the question is what?

The answer to this question requires a closer look at the nature of deletion. It is a fact that languages differ in the array of constituents that may be targeted by deletion

---

<sup>8</sup>I maintain here most of the standard Principles and Parameters assumptions about the architecture of the grammar, in particular, the hypothesis that the only representations that are subject to syntactic constraints are the "interface" levels LF and PF. Given the nature of the two crucial constraints that I claim are relevant to capturing the CD/CSD distinction, the constraint rankings and candidates that I will consider below must involve the PF interface. Presumably different constraints apply to the LF interface, though this is not a question that I will address here.

operations (VP-deletion, for example, is possible in a relatively small number of languages, while “slicing” is possible in quite a large number of languages; see Merchant 1999). In English, deletion of nominals is generally impossible, even when the information they convey is recoverable, except in certain contexts. One such context is when the nominal material is the complement of certain overt determiners (‘N?-deletion’), as in (56a).

- (56) a. Kim could see stars, but I couldn't see [NP ~~stars~~].  
 b. \*Kim could see stars, but I couldn't see [DP ~~stars~~].

Examples like (56b) indicate that null determiners cannot license deletion of nominal material, however. Assuming that this type of deletion is forbidden by a highly-ranked constraint, candidate (d) in Tableau 1 is correctly ruled out.

A second context in which a nominal can be deleted is when it is part of a movement chain (assuming the “copy and delete” theory of movement). Typically, only the lower member(s) of a movement chain are deleted (in English), leaving the head intact. If deletion is a general operation on copies, however (arguably the null hypothesis), then deletion of the highest copy in a chain ought to be possible as well, as long as the content is recoverable. Comparatives in which the head and the compared constituent are identical – i.e., CD constructions – are just such a context (relative clauses are arguably another; see Cresti (this volume)). On this view, the actual structure of the (a) candidate in Tableau 1 is (57), where the compared constituent has been moved (copied) to a clause-initial position, and both of the resulting copies are deleted by the copy and delete mechanism.

- (57) The galaxy contains more stars than [CP [DP ~~stars~~]<sub>i</sub>] the eye can see [DP ~~stars~~]<sub>i</sub>].

In essence, I am claiming that CD can be analyzed as ordinary A'-movement in which both copies of a moved phrase – the compared constituent – are deleted. More precisely, if DELETE requires redundant material to be eliminated from the PF representation, then in comparatives both copies *must* be deleted. The analysis thus explicitly makes deletion in CD obligatory (explaining e.g. (3a-c)), a property that sets it apart from approaches that rely on ellipsis to do the same job, since ellipsis operations are always optional.

The obligatoriness of deletion in CD, and in movement more generally (all copies in a movement chain must be deleted, up to recoverability), represents an important difference between “movement deletion” and “ellipsis deletion”. A second crucial difference between these two types of deletion is that the former is local (Kennedy 1998, 1999; cf. Williams 1977). This is illustrated for comparatives by examples like (58), which is contradictory.<sup>9</sup>

(58) Now that the remodeling has been completed, the space station is longer than it is used to be, and it's even wider than it is.

- (59) a. \*...the space station is longer than [CP [DegP ~~longer~~]<sub>i</sub>] it used to be [t]<sub>j</sub>, and it's even wider [CP [DegP ~~longer~~]<sub>i</sub>] than it is [t]<sub>j</sub>.  
 b. ...the space station is longer than [CP [DegP ~~longer~~]<sub>i</sub>] it used to be [t]<sub>j</sub>, and it's even wider [CP [DegP ~~wider~~]<sub>i</sub>] than it is [t]<sub>j</sub>.

(58) has only the contradictory interpretation corresponding to (59b), not the one in (59a),

<sup>9</sup>For simplicity, I will continue to represent the base position of the moved comparative constituent as an empty category, unless the properties of the lowest copy are relevant.

even though the reading represented by (59a) is the more sensible of the two. If deletion of the moved phrase could be licensed nonlocally, then (59a) ought to be possible, where *long* in the second clause is deleted under identity with the occurrence of *long* in the first clause. The fact that this reading is unavailable indicates that deletion of the compared constituent is licensed only by the head that it is in construction with.

In contrast, deletion in ellipsis may be licensed nonlocally. This is most clearly illustrated by an example like (60).

(60) Jones didn't meet all the candidates she wanted to, but I bet she liked the ones she did [vp meet]

The surface string corresponding to (60) is actually ambiguous, but only the reading shown here, in which deletion is licensed nonlocally by *meet* in the first clause, is informative (the reading in which the deleted VP is headed by *like* is a tautology), hence it is preferred.

These facts indicate that the simple notion of "minimizing the PF representation" used to define DELETE above needs to be refined. On the one hand, we must allow for the optionality of ellipsis; on the other hand, we need to forbid nonlocal deletion in comparison (and movement in general). These results could be accomplished in a number of ways; for now, I will adopt the following two assumptions. First, I assume that DELETE specifically refers to (possibly trivial) chains, requiring all elements of chain to be eliminated from the PF representation.<sup>10</sup> Since ellipsis is distinct (functionally and grammatically) from chain formation, it does not interact with this constraint.<sup>11</sup> Second, in order to explain the locality effects in CD, I will further assume that deletion in movement is more restricted than deletion in ellipsis, obeying an appropriately formulated locality requirement.

Note that these assumptions do not entail that DELETE can only be satisfied through movement. The constraint is satisfied as long as all elements of a chain are deleted; it does not matter how deletion is brought about. This leads to a clear prediction: if a compared constituent identical to the head is part of a constituent that is targeted by ellipsis, then it should remain in its base position (as in (d) in Tableau 1), since by doing so it would simultaneously satisfy DELETE and STAY. This prediction will be evaluated in section 3.2.

We are now in a position to examine CSD, the crucial property of which is that the compared constituent is not identical with the head. Because the two phrases are not identical, deletion of (copies of) the compared constituent (in either a movement chain or a trivial chain) would violate recoverability, which I assume to be inviolable (see Fienngo and Lasnik 1972). The result is that STAY emerges as the crucial constraint, and the optimal derivation is one in which the compared constituent remains in place. (This explains why a compared constituent that is lexically identical to the head but contrastively focused is not deleted (see note 1): deletion of contrastive focus information would violate recoverability.) This is illustrated by Tableau 2, which shows the two relevant candidates.

<sup>10</sup>A result of this assumption is that DELETE is violated by any overt phrase. This is not a problem for a model that allows constraint violation; it simply means that some higher-ranking constraint must be satisfied instead, in this case, a constraint requiring recoverability of deletion (see below).

<sup>11</sup>In essence, the separation between movement and ellipsis that I am getting at here corresponds directly to William's (1977) distinction between "sentence grammar" and "discourse grammar". In these terms, my proposal is that DELETE governs only aspects sentence grammar (which includes movement), not principles of discourse grammar (which includes ellipsis).

<i>Tableau 2</i>		
a.	Michael has more scoring titles than [CP [DP tattoos] <sub>i</sub> ]	*
b.	Michael has more scoring titles than [CP Dennis has [DP tattoos]]	*
		DELETE
		STAY

Viewed from the perspective of a model that assumes ranked and violable constraints, CSD thus represents an expected outcome: a construction in which a lower-ranked (and typically violated) constraint plays a crucial role in determining well-formedness because of violations of higher-ranked constraints.

An immediate challenge to this proposal is that it appears to (incorrectly) predict that English should be a *wh*-in situ language, since STAY is violated by any instance of overt A'-movement. Not only would this be the wrong result in e.g. questions and relative clauses, it would be wrong for comparatives in dialects of English that contain overt *wh*-words:

- (61) a. The flooding was less than what we had thought it would be.  
 b. \*The flooding was less than we had thought it would be what.

The response to this challenge that I will advocate here is that A'-movement in comparatives that do not involve lexical *wh*-XPs is actually a form of Quantifier Raising (QR), not feature-driven movement to SpecCP (cf. Larson 1988). This assumption can be semantically motivated by adopting the semantic analysis of comparative morphology proposed in section 2.1, which requires the compared constituent to raise in order to generate an appropriate interpretation for the comparative clause (as a definite description of a degree; see also the discussion in section 2.3 below). Syntactically, it puts the compared constituent in the same class as quantificational DPs, which also do not move overtly. The fact that lexical *wh*-phrases must move follows if STAY is ranked below a constraint OPSPEC, which requires syntactic operators to occupy a specifier position in overt syntax, as argued in Grimshaw 1997. Assuming that only compared constituents that are headed by a lexical *wh*-expression are "syntactic operators", the contrast between the examples in (61) follows, as shown by Tableau 3.

<i>Tableau 3</i>		
a.	The flooding was less than [CP what <sub>i</sub> we had thought it would be <i>t</i> <sub>i</sub> ]	*
b.	The flooding was less than [CP we had thought it would be what]	*i
		DELETE
		STAY

### 2.3 Summary: Movement and Deletion in English Comparatives

The basic "rule" of English comparative formation that emerges from the proposals outlined here can be succinctly stated as in (62).

- (62) English Comparative Formation (final)  
 Adjoin the compared constituent to the complement of *than*.

(62) is of course not a grammatical rule in the standard sense, since it describes a movement operation that is driven by independent properties of the compared constituent, which I assumed above to be semantic in nature.<sup>12</sup> The crucial point is that within the analytical framework developed in the previous sections, (62) is all that needs to be said about the syntax of comparatives: the derivational distinction that crucially differentiates CD and CSD (i.e., overt vs. covert movement) follows from the interaction of two general constraints on movement and deletion, DELETE and STAY. To the extent that this approach provides a more comprehensive account of the range of data associated with comparatives in English than earlier proposals, it makes a very strong case for adopting a model in which syntactic constraints are ranked and violable. In the next section, I will introduce two other types of data that provide additional support for the analysis.

### 3. Additional Evidence for Ranked Constraints in Comparatives

#### 3.1 Cross-linguistic Variation in Comparatives

An expectation of any optimality-theoretic analysis is that constraint re-ranking should correspond to typological variation. In the case of expressions of comparison, this is a difficult expectation to evaluate, because of the sheer level and variety of cross-linguistic variation they display (see Stassen 1985 for a survey). In many languages, strategies of movement and deletion are arguably not operative, making the possibility of capturing typological variation in terms of the relative ranking of constraints governing these syntactic options remote at best. However, it is still worth considering what we might expect to find in languages where these principles do play a role in the syntax of comparison.

Clearly, the most obvious constraint re-ranking to check is one in which STAY is superior to DELETE. This type of ordering would result in a language in which the compared constituent is either not moved and not deleted, or else not moved but deleted by some other deletion operation available in the language. So-called “conjoined comparative” languages such as Banda and Ilocano may fit this description, as illustrated by the examples in (63) and (64) (data from Stassen 1985).

- (63) *Banda* Anda ne mo gere nini, e ne ze gere.  
house of me big not, it of you big  
'Your house is bigger than my house.'
- (64) *Ilocano* Nasayaat ni Dolores, sumangka-sayaat pay ni Enkarnasion.  
pretty-one SUBJ Dolores, more-pretty even SUBJ Enkarnasion  
'Enkarnasion is prettier than Dolores.'

Further research needs to address the question of whether such comparatives display properties of covert movement in these languages, but data like these are at least suggestive. A second kind of typological variation could emerge from the interaction of the crucial constraints DELETE and STAY with other constraints in the language. For example, if STAY is outranked by a constraint that forces overt movement in contexts that include comparatives, we would expect to find examples of CSD in which movement of the

<sup>12</sup>Although this position is incompatible with Minimalist assumptions about movement, it is consistent with traditional views of QR (see May 1991 and Kennedy 1997a for discussion), which I assume here. The same results could be presumably be obtained in a feature-based approach to LF movement.

compared constituent is overt. As reported in Rivero 1981, Knowles 1984, and Price 1990, Castilian Spanish is just such a language (data from Price 1990:43; Borsley 1984 describes similar facts in Polish equatives):

- (65) a. Mi padre vende más libros que discos compra mi madre.  
 My father sells more books than records buys my mother.  
 \*Mi padre vende más libros que mi madre compra discos.  
 b. 'My father sells more books than my mother buys records.'
- (66) a. La mesa es más larga que ancha es la puerta.  
 The table is more long than wide is the door.  
 \*La mesa es más larga que la puerta es ancha.  
 b. 'The table is longer than the door is wide.'
- (67) a. El crío gatea más cuidadosamente que descuidadamente anda su hermana.  
 The baby crawls more carefully than carelessly walks his sister.  
 b. \*El crío gatea más cuidadosamente que su hermana anda descuidadamente.  
 'The baby crawls more carefully than his sister walks carelessly.'

### 3.2 Hidden Subdeletion

The Spanish data alone do not constitute an argument in favor of ranked and violable constraints in syntax, since the facts could also be accounted for in a standard model by assuming e.g. that the compared constituent in CD and CSD in Spanish bears a strong feature that forces overt movement (in the sense of Chomsky 1995; see Rivero 1981 for an approach along these lines). This type of approach would have no principled explanation for the typological difference between English and Spanish, however. Although the different properties of CSD in the two languages could be derived by assuming that the features on the compared constituent in English are weak, in order to capture the further distinction between CD and CSD in English we would need to make the completely ad hoc stipulation that these features are weak if and only if the compared constituent is not identical to the head, otherwise strong. This proposition is not only falsified by the data discussed in the next section, which show that even in some cases of identity (i.e. CD) the compared constituent does not move overtly, it is also completely devoid of explanatory power. In contrast, provided we can construct a principled characterization of the constraint that forces violation of STAY in Spanish CSD, both the typological difference between English and Spanish and the English-specific difference between CD and CSD can be handled in a straightforward and explanatory way by constraint (re-)ranking.<sup>13</sup>

Recall that the final definition of DELETE requires only that elements of a chain be deleted; it does not specify how this deletion should be accomplished. The prediction of this definition is that if the compared constituent is part of a larger phrase that is targeted for deletion, then the compared constituent should not undergo overt movement. For example, the sentence in (68a), in which the embedded VP is deleted, should be assigned the PF in (68b), in which the compared constituent has remained in situ: since VP-deletion has the effect of deleting the compared constituent from the PF representation, this constituent can satisfy DELETE without undergoing (overt) movement.

<sup>13</sup>One possibility, suggested to me by Norbert Corver (p.c.), would be to formulate this constraint in terms of focus movement, since the compared constituents in CSD must bear contrastive focus.

Comparative (Sub)deletion

- (68) a. Dennis has more tattoos than Michael does.  
 b. Dennis has more tattoos than [CP Michael does [VP ~~have~~ [DP tattoos]]]

Tableau 4 illustrates the analysis. Even though both candidates correspond to identical surface strings, deletion of the compared constituent in (b) is “parasitic” on VP-deletion, simultaneously satisfying both DELETE and STAY. The result is that a “hidden subdeletion” structure is assigned to an example of CD.

Tableau 4			
a.	Dennis has more tattoos than [CP [DP tattoos] Michael does		*i
b.	Dennis has more tattoos than [CP Michael does [VP <del>have</del> [DP tattoos]]]		

Clearly, if evidence for hidden subdeletion can be identified, it would constitute strong support for the claim that the syntax of comparatives involves ranked and violable constraints. An analysis that does not use constraint ranking would predict either that overt movement in CD should occur no matter what, or at least that it should be possible. The analysis I have advocated here, however, predicts that *only* the hidden subdeletion analysis is possible in these contexts, since the alternative “standard” CD analysis (corresponding to the (a) candidate in Tableau 4) is non-optimal.

Two contexts provide evidence for hidden subdeletion. The first involves from multiply-headed comparatives. Recall from the discussion in section 1.2 that multiply-headed CD is unacceptable; this was illustrated by the examples in (69).

- (69) a. \*Max persuaded more people to buy more cars than you persuaded to buy.  
 b. \*Christmas doesn't make as many children as happy as birthdays make.

This constraint is not absolute, however. Multiple CD is possible only if the compared constituents are contained in a larger deleted constituent, an observation made by Izvorski (1995) but as yet unexplained (see also Andrews 1985):

- (70) a. Max persuaded more people to buy more cars than you did.  
 b. Christmas doesn't make as many children as happy as birthdays do.

Within the context of the analysis I have proposed here, these facts constitute evidence for hidden subdeletion. The reason that multiply-headed CD is acceptable in (70) that deletion of the VP allows (in fact requires) the compared constituents to remain in situ. As in Tableau 4 above, VP-deletion without movement generates representations that simultaneously satisfy both STAY and DELETE, as illustrated in (71). Since the compared constituents do not move overtly, the multiply-headed comparative is well-formed.

- (71) a. Max persuaded more people to buy more cars than you did [VP ~~persuaded~~ [DP people] to buy [DP cars]]  
 b. Christmas doesn't make as many children as happy as birthdays do [VP ~~make~~ [DP children] [DegP happy]]

Even stronger evidence for hidden subdeletion comes from the interaction of comparative deletion, VP-deletion, and parasitic gaps. As shown by (72) and (73), VP-deletion and other types of ellipsis can bleed otherwise acceptable parasitic gaps. (Kennedy and Merchant 2000 observe, but do not explain, similar facts in attributive comparatives.)

- (72) a. Mo interviewed more suspects than Art interviewed without arresting  $e_i$ .  
 b. \*Mo interviewed more suspects than Art did without arresting  $e_i$ .
- (73) a. I actually liked more of the films that came out this year than I expected to enjoy before seeing  $e_i$ .  
 b. \*I actually liked more of the films that came out this year than I expected to before seeing  $e_i$ .

This is not a property of VP-deletion in general: a *wh*-phrase extracted out of a deleted VP can license a parasitic gap, as shown by (74).

- (74) I don't know which films $_j$  Hillary enjoyed  $t_j$  after seeing  $e_j$ , but I know which books $_i$  she did [VP enjoyed  $t_i$ ] after reading  $e_i$ .

If deletion of the VP forces a subdeletion-like analysis of CD, however, then the contrasts in (72) and (73) follow. Deletion of the VP prohibits overt movement the compared constituent, as in (75), and the parasitic gap is not licensed.

- (75) a. Mo interviewed more suspects than Art did [VP interviewed [DP suspects] $_i$ ] without arresting  $e_i$ .  
 b. I actually liked more of the films that came out this year than I expected to [VP like [DP films] $_i$ ] before seeing  $e_i$ .

Note that if a hidden subdeletion analysis were possible but not required, then (72b) and (73b) would have analyses involving overt movement, and the parasitic gaps would be licensed. That these sentences are ill-formed shows that overt movement is impossible: when a compared constituent identical to the head can be eliminated without movement, it does not move. This result that follows directly from the analysis I have presented here; it would have to be stipulated in analyses that do not make reference to a notion of optimality.

#### 4. Derived-head Analyses

In terms of empirical coverage, the analysis proposed here fares better than other uniform analyses, which either fail to fully capture the differences between CD and CSD, or run into problems with the Left Branch Constraint. (As we have seen, even though Izvorski's (1995) movement analysis avoids the problems with left branch effects, it cannot explain the parasitic gap data.) At the same time, the analysis provides a more explanatory account of English comparatives than non-uniform approaches, since it straightforwardly captures the similarities between CD and CSD.

Recent work by Winfried Lechner (1999) represents a stronger challenge to the approach I have advocated here, however. This work, which pursues what I will refer to as a "derived head" analysis, treats CD as movement of the compared constituent into the head position of the comparative clause, as shown in the examples in (76). (This type of analysis was first suggested for comparatives in Rivero 1981; cf. the analyses of relative clauses in Schachter 1973, Vergnaud 1974, Kayne 1994, Bhatt (this volume) and related work.)

Comparative (Sub)deletion

- (76) a. Michael has [DP more [NP scoring titles]<sub>i</sub>] than [CP Dennis has *t<sub>i</sub>*]  
 b. Michael's feet are [DegP [AP wide]<sub>i</sub>] er than [CP my feet are *t<sub>i</sub>*]

Although Lechner does not discuss CSD, the most natural assumption is that both the head and the compared constituent are generated in their surface positions, as in (77).

- (77) a. Michael has [DP more [NP scoring titles]<sub>i</sub>] than [CP Dennis has tattoos]<sub>i</sub>  
 b. Michael's feet are [DegP [AP wide] er than [CP my feet are long]<sub>i</sub>

If we add the further assumption that the compared constituent must raise at LF (e.g. for interpretive reasons, as in the analysis proposed here; cf. Rivero 1981), then the various properties of CD and CSD could be explained in exactly the same way that I explained them in section 2.1. The crucial difference between the derived head analysis and the one I have advocated, however, is that the former does not need to assume that syntactic constraints are ranked and violable. Instead, we can make the single assumption that an empty head position in a comparative must be filled; in Lechner 1999, this is implemented by postulating a feature on the head that is eliminated by movement of the compared constituent. The result is that movement is forced in CD by the principles of feature checking.<sup>14</sup>

There are at least three compelling arguments against this type of approach, two of which come from data we have already seen. Recall from the previous section that multiply-headed CD is possible because VP-deletion allows an example like (78a) to be assigned the hidden subdeletion analysis in (78b), in which the compared constituents do not move overtly.

- (78) a. Christmas doesn't make as many children as happy as birthdays do.  
 b. Christmas doesn't make as many children as happy as birthdays do [VP make [DP children] [DegP happy]]

Under the derived head analysis, however, the two overt heads should raise from their base positions in the comparative clause. But if this is the case, there is no difference between well-formed sentences with VP-deletion and unacceptable examples without (such as (69b) above): both involve the same movement of the same constituents.

The interaction of parasitic gaps and VP-deletion in comparatives provides a similar argument against a derived head approach. In section 3.2, the unacceptability of (79a) was shown to follow from the fact that VP-deletion forces a hidden subdeletion analysis of the comparative: since no overt movement occurs, the parasitic gap is not licensed.

- (79) a. \*Mo interviewed more suspects than Art did without arresting *e<sub>i</sub>*  
 b. Mo interviewed more suspects than Art did [VP interview [DP suspects]<sub>i</sub>] without arresting *e<sub>i</sub>*

The problem for the derived head analysis is the same as we saw above. If the overt head of the comparative is raised from a lower position, then the ill-formed example in (79a) should have the same derivational history as a comparable well-formed example that does

<sup>14</sup>As observed by Lechner (1999:33), this analysis derives the locality of "deletion" in comparatives in an elegant way: since the head of the comparative is literally a copy of the compared constituent, it must receive the same interpretation.

not involve VP-deletion, such as (72a) above. The result is that the explanation for the contrast between the two examples disappears.

The only way to save the derived head analysis from these two arguments would be to make essentially the same claim I have made here: that deletion of a phrase that contains the compared constituent somehow licenses a hidden subdeletion analysis. At the same time, it must be the case that only deletion of a containing phrase licenses a hidden subdeletion analysis, since otherwise all examples of CD would be incorrectly predicted to have (at least) an alternative syntactic analysis as CSD. In my proposal, this result follows from principles of optimality, in particular, the emergence of the lower-ranked constraint STAY as the crucial constraint differentiating examples in which DELETE is satisfied in different ways (as well as examples in which DELETE must be violated, such as standard cases of CSD). It is unclear what this result would follow from in a derived head analysis that does not also make similar assumptions about optimality and constraint ordering.

A third argument against a derived head analysis comes from disjoint reference effects in the comparative clause. Lechner (1999) argues that the impossibility of coreference in (80a) provides an argument in favor of a derived head approach, since this type of analysis, in conjunction with the copy theory of movement, derives (80a) from the structure in (80b), which violates Condition C.

- (80) a. \*Louise is prouder of Frank<sub>i</sub> than he<sub>i</sub> is.  
 b. Louise is [DegP [AP prouder of Frank<sub>i</sub>]<sub>j</sub>] than he<sub>i</sub> is [DegP [AP ~~prouder of Frank<sub>i</sub>~~]<sub>j</sub>].

The acceptability of (81a) is a problem for this hypothesis, though, since the structure assigned to it by the derived head analysis should also violate Condition C:

- (81) a. Louise is prouder of Frank<sub>i</sub> than he<sub>i</sub> thinks she is.  
 b. Louise is [DegP [AP prouder of Frank<sub>i</sub>]<sub>j</sub>] than he<sub>i</sub> thinks she is [DegP [AP ~~prouder of Frank<sub>i</sub>~~]<sub>j</sub>].

These facts can be accounted for within the analysis developed in section 2 if we add the assumption that deletion in movement, like deletion in ellipsis, does not require strict form identity between referring expressions and pronouns, but only identity of indices (Fienngo and May 1994), a position that Safir (to appear) adopts for relative clauses on independent grounds. If this is correct, then the occurrence of *Frank<sub>i</sub>* in the head of the comparative would count as identical with a coindexed pronoun *him<sub>i</sub>* in the compared constituent, and deletion of the compared constituent would be licensed, as shown in (82). (Here I show traces as full copies.)

- (82) Louise is prouder of Frank<sub>i</sub> than [CP [DegP ~~prouder of him<sub>i</sub>~~]<sub>j</sub>] he<sub>i</sub> thinks she is [DegP ~~prouder of him<sub>i</sub>~~]<sub>j</sub>].

This explains the acceptability of (81a); the unacceptability of (80a) may then be analyzed as a violation of Condition B, rather than a violation of Condition C, since the structure assigned to this sentence would be (83).

- (83) \*Louise is prouder of Frank<sub>i</sub> than [CP [DegP ~~prouder of him<sub>i</sub>~~]<sub>j</sub>] he<sub>i</sub> is [DegP ~~prouder of him<sub>i</sub>~~]<sub>j</sub>].

Crucially, it must not be the case that there is an alternative syntactic analysis of (80a) in which *him<sub>i</sub>* is replaced with a reflexive pronoun, as in (84), since such a structure would not violate Condition B.

(84) Louise is prouder of Frank<sub>i</sub> than [<sub>DegP</sub> proud of himself<sub>i</sub>]<sub>j</sub> he<sub>i</sub> is [<sub>DegP</sub> proud of ~~himself<sub>i</sub>~~]<sub>j</sub>

In order to rule out the analysis in (84), referring expressions and reflexive pronouns must not count as identical for the purposes of deletion, even if they are coindexed. The fact that (85a) is unacceptable (cf. (85b)) suggests that this is indeed the case, at least with respect to licensing deletion of adjectival predicates.<sup>15</sup>

(85) a. \*Louise is proud of Frank<sub>i</sub>, but he<sub>i</sub> isn't.  
 b. Louise is proud of Frank<sub>i</sub>, but he<sub>i</sub> doesn't think she is.

Regardless of how this issue is resolved, the crucial point to take from this discussion is that since a derived head analysis states that the head of the comparative is literally the same as the compared constituent, it makes the wrong predictions regarding examples like (81a). In contrast, the movement-plus-deletion analysis that I have advocated in this paper can appeal to independently-justified properties of deletion (such as conditions on identity) to explain the data.

## 5. Conclusion

This paper has presented an analysis of the syntax of English comparatives in which comparative deletion involves overt movement plus deletion of a compared constituent, while comparative subdeletion involves covert movement of the same element. This analysis was shown to provide a principled explanation of both the similarities and differences between these two types of comparatives, thus achieving a level of descriptive explanatory adequacy not matched by earlier proposals. At the same time, the derivational distinction between CD and CSD was shown to follow from general assumptions about the relation between movement and deletion in a model in which syntactic constraints are ranked and violable. While the larger implications of this proposal remain to be explored (in particular, with respect to other movement constructions such as relative clauses), its overall success in accounting for the properties of comparatives, a notoriously complex area of English syntax, and the surprising properties of “hidden subdeletion” makes a very strong case for ranked and violable constraints in syntax.

## References

- Abney, Steven. 1987. *The English Noun Phrase in Its Sentential Aspect*. Doctoral dissertation, MIT.
- Andrews, Avery. 1985. *Studies in the Syntax of Relative and Comparative Clauses*. New York: Garland.
- Bhatt, Rajesh. This volume. Adjectival modifiers and the raising analysis of relative

<sup>15</sup>If we adopt e.g. Reinhart and Reuland's (1993) approach to reflexivity, then referring expressions and reflexive pronouns are different sorts of semantic objects – the former refer to entities; the latter are functions on predicates – strengthening the view that they should not count as identical in deletion.

- Borsley, Robert. 1984. *Wh-movement and unbounded deletion in Polish equatives.* *Journal of Linguistics* 17:179-392.
- Bresnan, Joan. 1975. Comparative deletion and constraints on transformations. *Linguistic Analysis* 1.1:25-74.
- Bresnan, Joan. 1973. Syntax of the comparative clause construction in English. *Linguistic Inquiry* 4:275-343.
- Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge: MIT Press.
- Chomsky, Noam. 1977. On *wh-movement*. In P. Culicover et al. (eds.), *Formal Syntax*. Academic Press: New York.
- Chomsky, Noam and Howard Lasnik. 1977. Filters and control. *Linguistic Inquiry* 8:3.
- Corver, Norbert. 1990. *The Syntax of Left Branch Constructions*. Doctoral dissertation, Tilburg University.
- Corver, Norbert. 1993. A note on subcomparatives. *Linguistic Inquiry* 24:773-781.
- Cresti, Diana. This volume. Ellipsis and reconstruction in relative clauses.
- Donati, Caterina. 1998. Another way to build *wh-structures*. Paper presented at the Workshop on Comparatives, ZAS, Berlin, November 27-28, 1998.
- Fiengo, Robert and Howard Lasnik. 1972. On nonrecoverable deletion in syntax. *Linguistic Inquiry* 3.4:528.
- Fiengo, Robert, and Robert May. 1994. *Indices and Identity*. Cambridge: MIT Press.
- Gazdar, Gerald. 1981. A phrase structure syntax for comparative clauses. In T. Hoekstra, J. van der Hulst, and M. Morgan (eds.), *Lexical Grammar*. Oxford: Blackwell.
- Grimshaw, Jane. 1991. Extended projection. Unpublished ms., Brandeis University.
- Grimshaw, Jane. 1987. Subdeletion. *Linguistic Inquiry* 18:659-669.
- Hankamer, Jorge. 1971. Constraints on deletion in syntax. Doctoral dissertation, Yale University.
- Heim, Irene. 1985. Notes on comparatives and related matters. Ms., University of Texas, Austin.
- Hendriks, Petra. 1995. *Comparatives in categorial grammar*. Doctoral dissertation, Univ. of Groningen.
- Higginbotham, James and Robert May. 1981. Questions, quantifiers, and crossing. *The Linguistic Review* 1:41-79.
- Honegger, Mark. 1996. A phonological account of the "adverb effect" and *that*-violations. *Proceedings of the 1996 Mid-America Linguistics Conference*.
- Izvorski, Roumyana. 1995. A solution to the subcomparative paradox. In *The Proceedings of WCCFL XIV*. Stanford: CSLI Publications.
- Kayne, Richard. 1994. *The Antisymmetry of Syntax*. Cambridge: MIT Press.
- Kennedy, Christopher. 1999. *Projecting the adjective: The syntax and semantics of gradability and comparison*. New York: Garland.
- Kennedy, Christopher. 1998. Local dependencies in comparative deletion. In *The Proceedings of WCCFL XVII*. Stanford: CSLI Publications.
- Kennedy, Christopher. 1997a. Antecedent-contained deletion and the syntax of quantification. *Linguistic Inquiry* 28.4:662-688.
- Kennedy, Christopher. 1997b. On the quantificational force of the comparative clause. In Austin, Jennifer and Aaron Lawson (eds.), *Proceedings of ESCOL '97*. Ithaca: CLC Publications.
- Kennedy, Christopher. To appear. Polar opposition and the ontology of 'degrees'. *Linguistics & Philosophy*.
- Kennedy, Christopher and Jason Merchant. 2000. Attributive comparative deletion. *Natural Language and Linguistic Theory* 18.1.

- Klein, Ewan. 1980. A semantics for positive and comparative adjectives. *Linguistics & Philosophy* 4:1-45.
- Knowles, John. 1984. Structural choice in Spanish clausal comparison. *Canadian Journal of Linguistics* 29:1-19.
- Larson, R. 1988. Scope and comparatives. *Linguistics & Philosophy* 11:1-26.
- Larson, R. 1987. "Missing prepositions" and the analysis of English free relative clauses. *Linguistic Inquiry* 18.2:239-266.
- Lechner, Winfried. 1999. *Comparatives and DP Structure*. Doctoral dissertation, University of Massachusetts, Amherst.
- Lees, Robert. 1961. Grammatical analysis of the English comparative construction. *Word* 17.2:171-185.
- May, Robert. 1991. Syntax, semantics, and logical form. In Kasher, A. (ed.). *The Chomskyan Turn*. Cambridge: Basil Blackwell.
- May, Robert. 1985. *Logical Form*. Cambridge: MIT Press.
- Merchant, Jason. 1999. *The Syntax of Silence: Sluicing, Islands, and Identity in Ellipsis*. Doctoral dissertation, University of California, Santa Cruz.
- Napoli, Donna Jo. Comparative Ellipsis: A Phrase Structure Analysis. *Linguistic Inquiry* 14.4:675-694.
- Nissenbaum, Jon. 1998. Derived Predicates and the Interpretation of Parasitic Gaps. In Shahin, Kimary, Susan Blake and Eun-Sook Kim (eds.), *Proceedings of WCCFL17*. Stanford: Stanford Linguistics Association.
- Nissenbaum, Jon. This volume. Covert movement and parasitic gaps.
- Pinkham, Jesse. 1982. *The Formation of Comparative Clauses in French and English*. Doctoral dissertation, Harvard University.
- Price, Susan. 1990. *Comparative Constructions in Spanish and French Syntax*. New York: Blackwell.
- Prince, Alan and Paul Smolensky. 1993. *Optimality Theory: Constraint Interaction in Generative Grammar*. RuCCS Technical Report #1, Rutgers Center for Cognitive Science, Rutgers University. [To appear MIT Press, Cambridge.]
- Postal, Paul. 1998. *Three Investigations of Extraction*. Cambridge: MIT Press.
- Reinhart, Tanya and Eric Reuland. 1993. Reflexivity. *Linguistic Inquiry* 24.4:657-721.
- Rivero, Maria-Luisa. 1981. Wh-movement in comparatives in Spanish. In Cresssey, William and Donna Jo Napoli (eds.), *Linguistic Symposium on Romance Languages 9*. Washington: Georgetown University Press.
- Ross, John R. 1967. *Constraints on variables in syntax*. Doctoral dissertation, MIT.
- Rullmann, Hotze. 1995. *Maximality in the semantics of WH-constructions*. Doctoral dissertation, University of Massachusetts at Amherst.
- Safir, Ken. To appear. Vehicle change and reconstruction in A'-chains. *Linguistic Inquiry*. Stassen, Leon. 1985. *Comparison and Universal Grammar*. Oxford: Basil Blackwell.
- von Stechow, Arnim. 1984. Comparing semantic theories of comparison. *Journal of Semantics* 3:1-77.
- Vergnaud, Roger. 1974. *French Relative Clauses*. Doctoral dissertation, MIT.
- Williams, Edwin. 1977. Discourse and Logical Form. *Linguistic Inquiry* 8.1:101-140.

Department of Linguistics  
Northwestern University  
2016 Sheridan Road  
Evanston, IL 60208

kennedy@nwu.edu