CHRISTOPHER KENNEDY*

COMPARATIVE DELETION AND OPTIMALITY IN SYNTAX

ABSTRACT. This paper investigates the syntax of comparative deletion and comparative subdeletion in English and argues that the apparently paradoxical behavior of these two types of clausal comparative constructions is due to a derivational distinction between them: comparative deletion involves overt movement plus deletion of a compared phrase, while comparative subdeletion involves covert movement of the compared phrase. Although this derivational difference must be stipulated in standard approaches, it follows from general constraints on the relation between movement and deletion in English in a model of syntax in which syntactic constraints are ranked and violable, and well-formedness is determined by evaluating competing representations against the set of constraints, as in Optimality Theory. The analysis receives independent support from the interaction of comparatives and ellipsis, and achieves a higher level of descriptive and explanatory adequacy than alternative analyses that do not make reference to ranked and violable constraints.

1. INTRODUCTION

COMPARATIVE DELETION (CD) constructions are expressions of comparison such as those in (1), which compare two quantities of the same sort of stuff (number of stars, 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.

* I am very grateful to Tonia Bleam, Sandy Chung, Hana Filip, Hajime Hoji, Kyle Johnson, Richard Kayne, Jeff Lidz, Jason Merchant, Jon Nissenbaum, Eric Potsdam, Jane Grimshaw, Uli Sauerland, and Ayumi Ueyama for valuable comments on, and discussion of, the material in this paper. I am particularly indebted to Marcel den Dikken and the three anonymous *NLLT* reviewers for their insightful comments and suggestions. I would also like to thank audiences at NELS 30, University of California at Santa Cruz, Kyushu University, Indiana University, University of Iowa, University of Massachusetts at Amherst, New York University, Rutgers University, Stanford University, and the Zentrum für Allgemeine Sprachwissenschaft, Berlin, for their comments and thoughts on presentations of various components of what you see here. I take full responsibility for any shortcomings that remain in this paper after all of this helpful input. This paper is based upon work supported by the National Science Foundation under Grant No. 0094263.

Natural Language & Linguistic Theory **20:** 553–621, 2002.

Somewhat more exotic, but equally well-attested, are examples of COMPARATIVE SUBDELETION (CSD) such as the sentences in (2), which compare 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 .... (Greg Bear, 1997, *Slant*, New York, Tor, p. 262)

c. My sister drives as carefully as I drive carelessly.

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 CSD, an amount or degree term must be omitted from the constituent that provides the point of comparison with the morphologically marked phrase in the matrix clause (the COMPARED CONSTITUENT and HEAD OF THE COMPARATIVE, respectively, to use the terminology of Bresnan 1977), as shown in (3). Note that this is not a semantic restriction: (3a) could in principle mean something like 'the number of scoring titles that Michael has is greater than the number of tattoos that Dennis has, which is many/two'.

(3)a. Michael has more scoring titles than Dennis has (*two/*many) tattoos.

b. The shapes seem to be longer than they are (*2 inches/*that) thick.

c. My sister drives as carefully as I drive (*so/*very) carelessly.

In the case of CD, the lexical content must be omitted from the compared constituent as well, as illustrated by the examples in (4).

(4)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 (*carefully).

This 'obligatory deletion' requirement is important, as it distinguishes CD from other deletion operations in English, such as ellipsis, which is optional. There is one exception to this generalization about CD, however,

that involves examples in which the compared constituent bears contrastive focus, as in (5), where capitalization indicates focal stress ((5) is from Chomsky 1977, ex. (247); see also Sag 1976, p. 235).

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

b. Watching the Cubs on his satellite dish has been almost as difficult for Beck as watching Beck close games has been difficult for the CUBS. (*Chicago Tribune*, 6.8.99)

Since at least (Lees 1961), most analyses of comparatives in English have hypothesized that CSD structures are basic, 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 such as those in (4).[1] However, although a number of shared properties provide strong support for the hypothesis that CD and CSD have essentially the same syntax, research in the past fifteen years has uncovered empirical distinctions between the two constructions that call this conclusion into question. The goal of this paper is to show that a uniform and highly explanatory analysis of CD and CSD is in fact possible, but only if we adopt a syntactic framework in which constraints are ranked and violable and well-formedness is calculated in terms of principles of optimality.

Specifically, I will propose that all clausal comparatives in English involve Ā-movement of the compared constituent to the specifier of the

[1] A variety of other constituents can also be omitted from the comparative clause resulting in COMPARATIVE ELLIPSIS (CE) structures. Following Napoli (1983), I will assume that examples of CE like (ia)–(ie) involve either CD or CSD plus the various independently-motivated ellipsis operations listed below (see also Lechner 1999). In this paper, I will focus primarily on the syntactic analysis of the more basic CD/CSD configurations, though I will return to a discussion of the relation between comparative deletion and ellipsis in section 5.1.2.

(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* (or base-generation; see Hankamer 1973)

c. The galaxy contains more stars than anyone thought. *Null complement anaphora* (though see Kennedy and Merchant 2000b)

d. The galaxy contains more stars than it does planets. *Pseudogapping*

e. I suspect they have more to fear from us than we from them. *Gapping* (example from the film *Mars Attacks*)

clausal complement of *than* (i.e., SpecCP; see Hankamer 1973), but that the two constructions differ in when this movement applies. The central claims are summarized in (6).

(6) *English Comparative Formation* (Version 1)

i. CD involves overt movement of the compared constituent to the specifier of a clausal complement of *than/as*, plus deletion under identity with the head of the comparative (cf. Hankamer 1971; Chomsky 1977).[2]

ii. CSD involves covert movement of the compared constituent to the specifier of a clausal complement of *than/as*.

The crucial consequence of this analysis is that CD and CSD are structurally identical at LF (in the relevant respects), but differ at PF, predicting that the two types of comparatives should behave the same with respect to LF constraints, and that any differences should be localized to PF, a prediction that I show to be correct.

The paper is organized as follows. In section 2, I provide a detailed overview of the empirical data that any analysis needs to explain. Section 3 then shows how the proposal in (6) provides the basis for an explanation of both the similarities and differences between CD and CSD. Section 4 builds the theoretical justification for the analysis, demonstrating that the proposed distinction between CD and CSD follows from general principles governing movement and deletion in a syntactic framework in which constraints are ranked and violable, as in Optimality Theory. Section 5 discusses new data that provide independent support for the Optimality Theoretic analysis. Finally, section 6 considers an alternative approach that does not make reference to constraint ranking and optimality, and shows that it does not achieve the same level of descriptive and explanatory adequacy as the analysis advocated in this paper.

[2] This proposal is essentially the same in terms of its structural claims as recent versions of the 'matching analysis' of (restrictive) relative clauses, in which an internal head raises to SpecCP and deletes under identity with the external head (see Sauerland 1998 and Cresti 2000; see also Carlson 1977). If this is the correct analysis of (at least some) relative clauses (a point about which there is some debate), then the theoretical machinery that I propose in section 4 to explain *why* we get movement and deletion in cases of identity between the head and compared constituent in CD should extend directly to matching analyses of relative clauses.

2. SYNTACTIC PROPERTIES OF THE ENGLISH COMPARATIVE CLAUSE

2.1. *Evidence for Uniform Analysis of Comparatives*

The hypothesis that CD and CSD involve a single rule of comparative formation is most fully developed in Bresnan (1973, 1975). Bresnan argues that comparatives are derived by an unbounded deletion operation that obligatorily eliminates a degree term from the compared constituent. 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, p. 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 (7) and (8) (where the constituents that count as identical in Bresnan's analysis are underlined and deleted material is struck out).

(7)a. The galaxy contains more stars than the eye can see ~~x many stars~~

b. Sea level was as high as it later became ~~x much high~~

(8)a. Michael has more scoring titles than Dennis has ~~x many~~ tattoos

b. The shapes are longer than they are ~~x much~~ thick

Subsequent analyses, building on Ross's (1967) observation that CD and CSD have properties similar to *wh*-movement constructions, recast Bresnan's approach in terms of movement or binding of a degree term.[3] While there are important differences between these various analyses, 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 have a well-defined set of shared properties. In particular, they should behave the same with respect to constraints on movement or unbounded deletion.

Initial confirmation of this prediction comes from the fact that both CD and CSD require a gap, as observed in section 1, and both CD and CSD

[3] 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 Ā-movement of a degree term and CD as additional deletion of redundant material. Klein (1980) and Gazdar (1981) develop GPSG implementations of Bresnan's approach using the SLASH notation for unbounded dependencies, in which the head of the comparative can introduce either a S/XP or S/Deg constituent, i.e., a clause missing a phrase (CD) or a degree term (CSD).

are ill-formed when the gap is embedded in an extraction island (see Ross 1967; Huddleston 1967; Chomsky 1977, and for a more recent discussion, Postal 1998). This is illustrated by the examples in (9)–(12).

(9)a. *Complex NP islands*

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.

(10) *Wh-islands*

a. *The shapes were longer than I wondered whether they would be.

b. *The shapes were longer than I wondered whether would be thick.

(11) *Adjunct islands*

a. *My sister drives as carefully as I avoid accidents when I drive.

b. *My carefully as I get into accidents when I drive carelessly.

(12) *Sentential subjects*

a. *There are more stars in the sky than that the eye can see is certain.

b. *There are more stars in the sky than that the eye can see planets is certain.

Crucially, comparatives do permit the gap to be embedded in non-island complement clauses, as shown by (13).[4]

(13)a. Michael has more scoring titles than Kim says he has.

b. Michael has more scoring titles than Kim says Dennis plans to get tattoos.

A second piece of evidence for a uniform analysis of CD and CSD comes from crossover effects. (14) demonstrates that CD shows both

[4] Bresnan (1975, p. 59 ff. 10) points out that embedded occurrences of subdeletion are in general somewhat less felicitous than corresponding examples of comparative deletion, but are improved by maximizing parallelism between matrix and comparative clauses. Bresnan attributes the lowered felicity to a processing factor, a conclusion with which I agree.

strong and weak crossover effects, and (15) makes the same point for CSD (Bresnan 1975; Chomsky 1977).

(14)a. More Democrats$_i$ voted than they$_{*i/j}$ expected to vote.

b. More Democrats$_i$ voted than their$_{*?i/j}$ friends expected to vote.

(15)a. More Democrats voted than they$_{*i/j}$ expected Republicans$_i$ to vote.

b. More Democrats voted than their$_{*?i/j}$ friends expected Republicans$_i$ 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 these 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. 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. (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.) Both the CD and CSD options in (16a) and (17a), for example, can be assigned the interpretations paraphrased in (16b) and (17b), which differ only in terms of the sort of stuff compared.

(16)a. Michael has more scoring titles than Dennis has (tattoos).

b. *the number of Michael's scoring titles > the number of Dennis' scoring titles/tattoos*

(17)a. Michael's hands are as wide as my feet are (long)

b. *the width of Michael'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 they should also have the same types of interpretations.

2.2. *Evidence Against a Uniform Analysis*

The hypothesis that CD and CSD should be assigned the same syntactic analysis has received a strong challenge in recent years with the identification of a set of facts that clearly differentiate between the two types of comparatives, suggesting 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 assumption that CD is derived through unbounded deletion or some combination of Ā-movement and ellipsis, while CSD is different.[5] Five sets of data crucially distinguish CD and CSD.

First, in languages that prohibit preposition stranding, such as Czech, CD behaves like movement constructions and obeys this constraint, but CSD does not. This is illustrated by the contrast between (18) and (19); (20) demonstrates the unacceptability of preposition stranding in a comparable Czech question. (Thanks to Hana Filip for supplying the Czech data; this point was originally made by Corver (1990) for Dutch.)

(18)a. *Bydlel jsem ve více městech než ty jsi
live.PAST.1SG aux in more city.PL.GEN than you AUX
bydlel v.
live.PAST.2SG in
I have lived in more cities than you have lived in.

(19) Chci bydlet ve více amerických městech
want.1SG.PRES live.INF in more American city.PL.GEN
než jsem bydlel v europských městech.
than AUX lived.PAST.1SG in European city.PL.LOC
I want to have lived in more American cities than I have lived in European cities.

[5] 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 have distinct but homonymous degree morphemes, which select for syntactically distinct comparative clauses; Chomsky (1977) and Knowles (1984) claim that the constituent targeted by movement in CSD is different from the moved element in CD, while Rivero (1981) argues that it is the landing site of the moved constituent in CSD that differentiates it from CD.

(20) *Kterých městech Václav bydlel v?
which city.PL.LOC Vaclav live.PAST.3SG in
Which city does Vaclav live in?

Second, CD shows COMP-trace effects in English, but CSD does not (Bresnan 1977; Grimshaw 1987), as shown by the contrast between the (a) and (b) sentences in (21) and (22).

(21)a. More books were published than the editor said (*that) would be.

b. More boys flunked than I predicted (*that) would pass.

(22)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):

(23)a. I thought there was more meat than there is/*'s.

b. John was more upset then than he is/*'s now.

c. She was as happy about it then as she is/*'s now.

(24)a. There's more meat than there's rice.

b. John was more upset then than he's angry now.

c. She was as happy about it then as she's sad.

Fourth, CD licenses parasitic gaps, a fact that has been taken as further evidence for its status as a type of Ā-movement construction (see Postal 1998 for recent discussion), but CSD does not (Grimshaw 1987). This is illustrated by the contrasts between the examples in (25) and those in (26), where e denotes a parasitic gap.

(25)a. I threw away more books than I kept without reading *e*.

b. Jerome followed more suspects than Arthur interrogated without arresting *e*.

(26)a. *I threw away more books than I kept magazines without reading *e*.

b. *Jerome followed more leads than Arthur interrogated suspects without arresting *e*.

Note that examples of CSD in which the parasitic gap corresponds to a degree term are also impossible, as shown by (27a), which would have the interpretation in (27b).

(27)a. *I threw away more books than I kept magazines without reading *e* novels.

b. *the number n such that I threw away n books > the number m such that I kept m magazines without reading m novels*

(27a) is ruled out on independent grounds. Munn (2001) shows that parasitic gaps correspond to individual denoting expressions (see also Cinque 1990 and Postal 1993), but as (27b) makes clear, the gap in (27a) denotes a degree/amount.

This fact raises the question of whether parasitic gaps constitute a relevant difference between CD and CSD, since if CSD involved movement of a degree term, we would expect parasitic gaps to be impossible for independent reasons. Below I will argue that CSD involves (covert) movement of the whole compared constituent, however, and I will show in section 3.3 that CSD actually licenses parasitic gaps of the sort in (26) in exactly the same contexts that *wh-in situ* does (as documented in Nissenbaum 2000). I therefore assume that the unacceptability of the examples in (26) must be explained in terms of the grammar of comparatives and not (only) in terms of the grammar of parasitic gaps.

A fifth contrast comes from so-called MULTIPLY-HEADED COMPARATIVES, such as the CSD sentences in (28).

(28)a. Christmas makes as many children as happy as it makes adults unhappy.

b. 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), since

they involve multiple instances of comparison. (28b), for example, has the interpretation in (29).

(29) [*the number of men that Max persuaded to buy cars* > *the number of women that you persuaded to buy trucks*] AND [*the number of cars that Max persuaded men to buy* > *the number of trucks that you persuaded women to buy*]

Thus (28b) is false if Max persuaded more men to buy cars than you persuaded women to buy trucks, but the number of cars bought by the men was fewer than the number of trucks bought by the women (as in a context in which, for example, each woman bought several trucks, while each man bought only one).

The syntactic properties of these constructions are relatively clear, however: while examples of multiply-headed CSD like those in (28) are well-formed, parallel examples of multiply-headed CD are unacceptable (Corver 1990, 1993; Hendriks 1995):

(30)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.

'Mixed' multiply-headed comparatives – comparatives involving both CD and CSD – are acceptable, however, as illustrated in (31)

(31)a. Christmas makes as many people as happy as it makes unhappy.

b. Max persuaded more people to buy more cars than you persuaded to buy trucks.

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 Ā-movement of a null degree operator from inside the nominal or adjectival projection to the specifier of the complement of *than/as*, which we may assume to be CP (Hankamer 1973). This is illustrated in (32a) and (32b), where movement chains are represented as sequences consisting of one overt element and some number of deleted copies (as in Chomsky (1993) and subsequent versions of the 'copy and delete' theory of movement). Here I assume adjectival pro-

jections to be Degree Phrases (DegPs); see Abney (1987), Corver (1990) Grimshaw (1991) and Kennedy (1999).

(32)a. Michael has more scoring titles than [$_{CP}$ *Op* Dennis has [$_{DP}$ ~~*Op*~~ tattoos]]

b. The shapes are longer than [$_{CP}$ *Op* they are [$_{DegP}$ ~~*Op*~~ 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 (33), which violate the Left Branch Constraint (Ross 1967).

(33)a. *How many does Dennis have [$_{DP}$ ~~how many~~ tattoos]]?

b. *[$_{CP}$ How were the shapes [$_{DegP}$ ~~how~~ thick]]?

To get around this problem, Chomsky (1977, p. 123) suggests that subdeletion may involve feature movement only, and so may somehow avoid the Left Branch Constraint (a proposal recently revived by Donati 1998). Corver (1990) provides compelling arguments that the moved elements in (33) (and therefore in (32) as well) are heads, however, in which case the actual reason that (33a) and (33b) are unacceptable is that they violate the Head Movement Constraint. If this is correct, then Chomsky's proposal is untenable: movement of formal features alone in (32) would violate Generalized Pied Piping, which requires overt movement of formal features on a head X to take along XP as well (see Chomsky 1995, pp. 262–264). The apparent absence of left branch effects in CSD thus remains a problem for most movement-based approaches to comparatives.

Izvorski (1995) presents a movement account that does not suffer from this problem, because it claims that the term targeted by movement in CSD (and possibly CD as well) is not syntactically parallel to the degree heads *how (many/much)* in (33), but rather to the amount adjuncts *in what quantity* and *to what degree* in (34a) and (34b).[6]

[6] Izvorski develops this analysis specifically for CSD, but also suggests that CD could be given the same analysis if we assume that the redundant lexical material in the compared constituent is targeted by ellipsis. There are (at least) two reasons to believe that CD cannot be handled in this way, however.

First, an example like (ia) would have the structure in (ib), where the compared DP is elided.

(i)a. Dennis got more tattoos than Michael got.

b. ... than [~~in what quantity~~ Michael got ~~tattoos~~]

(34)a. In what quantity does Dennis have tattoos?

b. To what degree were the shapes thick?

This analysis succeeds in explaining many of the properties of CSD (for example, the fact that it is sensitive to islands but does not show COMP-trace effects), but it runs into a number of empirical problems. The most serious is that it fails to account for the fact that the compared constituent in CSD must be a bare nominal (a mass or plural NP). Although an overt degree term is incompatible with an amount adjunct, as in (35) (as Izvorski 1995, p. 205 observes), the DP that introduces the object(s) to be measured may have a determiner, as shown by (36).

(35)a. *In what quantity did many Republicans vote for this bill?

b. *In what quantity did Dennis buy several new shirts?

(36)a. In what quantity did the Republicans vote for this bill?

b. In what quantity did Dennis buy some new shirts?

The problem is that English doesn't allow this sort of DP-ellipsis elsewhere (nor does it allow NP-ellipsis without an overt determiner), as shown by (ii), which is structurally parallel to (ib) except that the amount operator is overt.

(ii) *I found out in what quantity Dennis got tattoos at the same time that I found out in what quantity Michael got ~~tattoos~~.

In other words, this analysis would have to stipulate that deletion of DP is possible (in fact obligatory; see section 1) only in comparatives; this follows in analyses in which CD directly targets the compared constituent. Of course, the burden on such analyses is to explain how CD and ellipsis are different; see the discussion of this point in section 4.3 below.

Second, if CD and CSD involve exactly the same type of Ā-movement operation, they should have the same licensing properties for parasitic gaps. As we have already seen, this is not the case: CD licenses parasitic gaps; CSD does not. More to the point, movement of the sort shown in (34a) does not license parasitic gaps at all, as shown by (iii).

(iii) *In what quantity did you throw away books without reading?

If CD involved this sort of movement, then it too should fail to license parasitic gaps. While Izvorski's analysis does appear to make the correct predictions for CSD regarding parasitic gaps, we will see in section 3.3 that CSD can license parasitic gaps in certain contexts.

CSD strictly requires the compared constituent to be bare, however:

(37)a. More Democrats abstained than (*the) Republicans voted.

b. Dennis bought more new ties than he bought (*some) new shirts.

2.3. *Partitive Comparatives*

I conclude this section by looking at a class of comparatives that have received somewhat less attention in the literature than the 'standard' CD and CSD constructions discussed above: comparatives with partitive syntax, which I will call PARTITIVE COMPARATIVES (see Bresnan 1975; Grimshaw 1987). These constructions come in both CD and CSD variants, as shown in (38).

(38)a. I met as many of the students as you met.

b. I met as many of the students as you met of the teachers.

With respect to their syntactic properties, partitive CD and CSD are mainly parallel to their standard CD/CSD counterparts. Like other comparatives, both partitive CD and CSD are sensitive to islands, as shown by (39a)–(39d). ((39e) shows that both types of partitive comparative can be embedded in a non-island context.)

(39)a. *I met as many of the students as you had wondered whether you would be able to meet (of the teachers).

b. *I met as many of the students as you were excited because you met (of the teachers)

c. *I met as many of the students as you were introduced to someone who met (of the teachers).

d. *I met as many of the students as that you would be able to meet (of the teachers) was certain.

e. I met as many of the students as you had hoped to be able to meet (of the teachers).

Similarly, both partitive CD and partitive CSD show strong and weak crossover effects:

(40)a. *I was able to meet more of the senators than they$_i$ thought I would be able to meet [Ø/of the representatives]$_i$.

b. *I was able to meet more of the senators than their$_i$ aides thought I would be able to meet [Ø/of the representatives]$_i$

Turning to the differences between CD/CSD documented in section 2.2, we see that partitive CD and CSD show parallel behavior to standard CD/CSD with respect to parasitic gap licensing. While partitive CD licenses parasitic gaps, partitive CSD does not:

(41)a. We invited more of the students to the party than you invited (*of the teachers) after meeting at the open house

b. We accepted as many of the domestic applications as they rejected (*of the foreign applications) after closely examining.

The parallel behavior of partitive comparatives and other comparatives breaks down when we look at the other contexts discussed in section 2.2, however, in the sense that partitive CD and CSD behave alike in environments where standard CD and CSD behave differently. First, as shown by (42), both partitive CD and partitive CSD forbid contraction of a preceding auxiliary:

(42)a. There isn't as much of the fish left as you said there is (of the rice).

b. *There isn't as much of the fish left as you said there's (of the rice).

Second, both partitive CD and partitive CSD show COMP-trace effects:

(43)a. *More of the old books were read than I thought that (of the new books) were read.

b. *As many of the students registered as you said that (of the teachers) were available.

There may be independent reasons for the ill-formedness of the partitive CSD variants in these examples, however. As shown by (44), partitive CSD is also bad in examples in which the complementizer is omitted.

(44)a. More of the old books were read than I think (*of the new books) were read.

b. As many of the students registered as you said (*of the teachers) were available.

Bresnan (1975) attributes this restriction to Kuno's (1973) Incomplete Subject Constraint, which prohibits movement out of a subject NP or clause when what would be left over is an incomplete subject/clause. If this is correct, then we cannot necessarily conclude from the data in (43) that partitive CSD shows COMP-trace effects. We can, however, conclude that partitive CSD and standard CSD are syntactically distinct, since only the latter allows the compared constituent to occur in an embedded subject position. I return to this point below.

Third, both partitive CD and partitive CSD appear to be impossible in multiply-headed comparatives, as shown by (45)–(46). (45b) also runs afoul of Incomplete Subject Constraint, but (46b) does not, so we can safely conclude that multiply headed partitive comparatives are impossible.

(45)a. *Max persuaded more of the men to buy more of the cars than you persuaded to buy.

b. *Max persuaded more of the men to buy more of the cars than you persuaded of the women to buy of the trucks.

(46)a. *Max sold as many of the men as many of the cars as you sold.

b. *Max sold as many of the men as many of the cars as you sold of the women of the trucks.

Furthermore, 'mixed' multiply-headed partitive comparatives appear to be impossible. Compare the pairs in (47)–(48), where the (a) examples are mixed multiply-headed standard comparatives (cf. (31) above) and the (b) examples are mixed multiply-headed partitive comparatives.

(47)a. Max persuaded more men to buy more cars than you persuaded to buy trucks.

b. *Max persuaded more of the men to buy more of the cars than you persuaded to buy of the trucks.

(48)a. Max sold as many men as many cars as you sold trucks.

b. *Max sold as many of the men as many of the cars as you sold of the trucks.

The conclusion to be drawn from this collection of facts is that, with the exception of parasitic gap licensing, all partitive comparatives – both the CD and CSD variants – have the properties of standard (non-partitive) CD constructions. In fact, Grimshaw (1987), following Taraldsen (1978), argues that examples of what I have been calling partitive CSD are in fact syntactically instances of CD, in that they target a full DP, with obligatory extraposition of the *of*-PP. (Grimshaw points to the parallel behavior of partitive *how many* questions in contexts such as those discussed above to support her proposal.) On this view, an example like (49) has a representation along the lines of (50a), rather than (50b), where trace notation is used to represent extraposition and crossed out structure indicates the target of CD.

(49) I met as many of the students as you met of the teachers

(50)a. I met as many of the students as you met ~~[$_{DP}$ many t_i]~~ [$_{PP}$ of the teachers]$_i$

b. I met as many of the students as you met [$_{DP}$ ~~many~~ [$_{PP}$ of the teachers]]

In section 4.2.3, I will develop an analysis of partitive comparatives that builds on the Grimshaw/Taraldsen extraposition proposal, and I will show that the properties of partitive comparatives can be explained in terms of the interaction of the larger analytical framework I will develop for English comparatives and the syntax of partitive DPs more generally.[7]

[7] One difference between the two types of partitive comparatives is that partitive CSD cannot target a constituent inside a prepositional phrase (Baltin 1978):

(i)a. John gave flowers to more of the men than he gave cigars to.

b. *John gave flowers to more of the men than he gave cigars to of the women. (cf. John gave flowers to more men than he gave cigars to women.)

3. Overt and Covert Movement in Comparatives

3.1. *Overview of the Proposal*

The facts discussed in section 2 indicate that there are compelling arguments in favor of assigning similar syntactic analyses to CD and CSD, but there are also empirical differences between the two types of constructions that must be derived in order to maintain a minimum level of descriptive adequacy. One conclusion that could be drawn 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 2.2 (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 differ in the level of representation at which these operations apply. This is the claim of the proposal presented in section 1, which is repeated below.

(51) *English Comparative Formation* (Version 1)

i. CD involves overt movement of the compared constituent to the specifier of the complement of *than/as* plus deletion under identity with the head of the comparative.

ii. CSD involves covert movement of the compared constituent to the specifier of the complement of *than/as*.

Partitive *how many* questions show the same restriction:

(ii)a. How many of the men did John give cigars to?

b. *How many did John give cigars to of the men? (cf. How many did John meet of the men?)

This looks like a P-stranding effect, but this is unlikely given that English in general does not forbid preposition stranding. At the very least, this can be taken as further evidence (along with (43) and (44) above) that partitive CSD and standard CSD are syntactically distinct. This restriction may be explainable in terms of extraposition, though the facts are not completely clear. Examples like the following do show that extraposition of a partitive PP out of the complement of a preposition is degraded, but more work needs to be done here.

(iii)a. *Kim read to a few yesterday of the children that had requested a bedtime story

b. ?Kim read a few yesterday of the papers that had been assigned for the class.

This proposal is stated in the terms of the Principles and Parameters framework (Chomsky 1993, 1995, etc.), in which the computational system generates two 'interface representations': one that provides the input to the semantics (LF), and one that provides the input to the phonology (PF). (I briefly address the question of how the analysis could be recast in other syntactic frameworks in the conclusion.) On this view, overt movements are those that are represented at PF (and typically at LF as well, though not necessarily; see Aoun and Benmamoun 1998), and covert movements are those that are represented only at LE. A sentence ⟨LF, PF⟩ is well-formed iff it (i) is generated by the computational system, and (ii) satisfies the interface constraints on LF and PF representations.

In section 4,1 will argue that these constraints must be ranked and violable, and that well-formedness with respect to them is determined by an optimality metric. For the immediate purpose of evaluating the empirical adequacy of (51), however, I will focus on a more general result of the proposal for English comparatives: CD and CSD have structurally identical LF representations but structurally distinct PF representations. This leads to the following two predictions:

1. CD and CSD should behave the same with respect to constraints on LFs.
2. All syntactic differences between the two types of comparatives should be localized to PF.

As the following two sections will demonstrate, the similarities and differences between CD and CSD that were discussed section 2 break down in exactly this way. (In what follows, I focus on standard comparatives, and address partitive comparatives in detail in section 4.2.3.)

3.2. *The Similarities*

Let us begin with semantics. The analysis derives the semantic similarity between CD and CSD by supporting a single compositional semantic analysis of both constructions. (Unlike in mixed analyses, it is not necessary to posit multiple lexical entries for the comparative morphemes, as in, e.g., Kennedy (1998, 1999).) Consider first the case of adjectival comparatives like those in (52), which have the LFs in (53a) and (53b). Here I take the impossibility of overt degree morphology on the compared constituent (see (3) in section 1) as an indication that this phrase is headed by a null degree morpheme, notated Deg^0_C.

(52)a. Michael's hands are wider than my feet are.

b. Michael's hands are wider than my feet are long.

(53)a. Michael's hands are wider than [$_{CP}$ [$_{DegP}$ Deg^0_C wide] my feet are ~~[$_{DegP}$ Deg^0_C wide]~~]

b. Michael's hands are wider than [$_{CP}$ [$_{DegP}$ Deg^0_C long] my feet are ~~[$_{DegP}$ Deg^0_C long]~~]

Following Kennedy (1999, 2001), I assume that gradable adjectives denote functions from objects to degrees and combine with degree morphology to generate properties of individuals (cf. Bartsch and Vennemaun 1973). Within this framework, the degree morphemes that head the compared constituent and the head of the comparative can be assigned the interpretations in (54a) and (54b), respectively (*less* and *as* differ from *more* only in the nature of the ordering relation they impose), where G is a function from objects to degrees, Q is a function from properties to truth values (the semantic value of a clausal constituent with an extracted DegP), and *max* is a maximality operator that returns the maximal element of an ordered set of objects.[8]

(54)a. $\text{Deg}^0_C = \lambda G \lambda Q.max\{d \mid Q(\lambda x.G(x) \succeq d)\}$

b. er/more = $\lambda G \lambda d \lambda x.G(x) \succ d$

Semantic composition in the comparative clause derives a definite description of a maximal degree. The crucial steps in the composition of the comparative clause in (53a) are illustrated with the tree in (55); (53b) is analyzed in exactly the same way, except that *wide* is replaced with *long*. Here I assume that traces (uninterpreted copies) are assigned a category-specific type (in this case, $\langle e, t \rangle$ for a predicative expression) and abstracted over (Heim and Kratzer 1998), and I ignore tense.

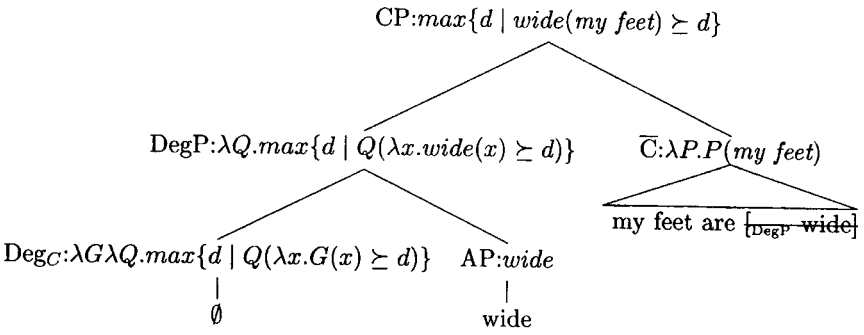
[8] The argument for Maximality in comparatives comes from von Stechow (1984) (see also Lerner and Pinkal 1992; Rullmann 1995; Gawron 1995 and Kennedy 1997). Von Stechow observes that treating the comparative clause as a simple definite description (as originally proposed in Russell 1905) won't work for (i).

(i) Kim can jump as far as Lee can jump.

The problem is that there is no unique degree d such that Lee can jump d-far (there are potentially quite a few such degrees), so a definite description would fail to denote. What (i) means is that Kim can jump farther than the *maximal* degree d such that Lee can jump d-far, hence the maximality operator in (54a). The interpretation of *max* is given in (ii).

(ii) $[\![max(P)]\!] = \iota d \in P.\forall d' \in P : d \succeq d'$

(55)



The comparative clause supplies the 'standard of comparison' argument for the comparative morpheme, which establishes a relation between two degrees. The interpretations that are ultimately assigned to (53a) and (53b) are as in (56), which are truth-conditionally equivalent to the informal representations given above in (17b). (See Kennedy (1999) for a detailed discussion of the compositional semantics of degree predicates.)

(56)a. $wide(michael's\ hands) \succ max\{d \mid wide(my\ feet) \succeq d\}$

b. $wide(michael's\ hands) \succ max\{d \mid long(my\ feet) \succeq d\}$

Nominal comparatives can be analyzed in essentially the same way. The examples in (57) are assigned the LFs in (58) (where D^0_C is the null head of the compared constituent).

(57)a. Michael has more scoring titles than Dennis has.

b. Michael has more scoring titles than Dennis has tattoos.

(58)a. Michael has more scoring titles than [$_{CP}$ [$_{DP}$ D^0_C scoring titles] Dennis has ~~[$_{DP}$ D^0_C scoring titles]~~]

b. [Michael has more scoring titles than [$_{CP}$ [$_{DP}$ D^0_C tattoos] Dennis has [~~$_{DP}$ D^0_C tattoos~~]]

Building on the semantic similarity between the vague determiner *many* and gradable adjectives (see Klein 1980 for discussion), I assume that part of the meaning of nominal degree morphology is a function MANY from plural objects to amounts.[9] The head of the compared constituent and the

[9] This assumption is implicit in most syntactic analyses of nominal comparatives (as illustrated by the underlying structures that Bresnan (1973) assigns to comparatives; see

comparative determiner *more* can then be assigned the meanings in (59), where X is a variable over pluralities, P is a (plural) NP meaning, Q is a function from plural objects to truth values (the semantic value of a clausal constituent with an extracted plural DP), and *max* is the maximality operator.

(59)a. $\mathrm{D}^0_C = \lambda P \lambda Q.max\{n \mid \exists X[P(X) \wedge Q(X) \wedge \textsc{many}(X) \geq n]\}$

b. $\text{more} = \lambda P \lambda m \lambda Q.\exists Y[P(Y) \wedge Q(Y) \wedge \textsc{many}(Y) > m]$

As in the case of adjectival comparatives, the comparative clause is interpreted as a description of a maximal amount, and supplies the standard of comparison for the comparative morpheme. The interpretations assigned to the LFs in (58) are given in (60), which are equivalent to the informal characterizations of the truth conditions of these sentences given above in (16b).

(60)a. $\exists Y[titles(Y) \wedge have(michael, Y) \wedge \textsc{many}(Y) > max\{n \mid \exists X[titles(X) \wedge have(dennis, X) \wedge \textsc{many}(X) \geq n\}]$

b. $\exists Y[titles(Y) \wedge have(Michael, Y) \wedge \textsc{many}(Y) > max\{n \mid \exists X[tattoos(X) \wedge have(dennis, X) \wedge \textsc{many}(X) \geq n\}]$

Turning to syntax, the crucial similarities between CD and CSD are sensitivity to extraction islands and crossover effects, two sets of phenomena that (in models that assume multiple levels of representation) have been analyzed as involving (at least) constraints on LF representations (see Higginbotham 1980; Huang 1982; May 1985; Chomsky 1995, and others). If CD and CSD have structurally identical LF representations, it follows that they should have the same range of (un-)acceptability in these contexts.

Note that given the semantic analysis of the compared constituent that I have proposed here, this element must move at LF: an *in situ* analysis is not possible. This is because the quantificational force of the comparative clause (the maximality operator) is introduced by the degree morphology on the compared constituent, not by a higher operator (cf. Baker's (1970)

(7) above), and is explicitly implemented (in different ways) in the semantic analyses of comparatives developed in, e.g., Cresswell (1977), von Stechow (1984), Gawron (1995) and Hackl (2001) (see also Grosu and Landman 1998 for an even more refined semantic analysis of degree morphology). As shown by Cresswell (1977), this sort of approach generalizes to comparatives with mass nouns as well.

analysis of questions and related work). In order to generate the right interpretation of the comparative clause, then, the compared constituent must take scope over the rest of the clause.[10]

Before moving to a discussion of the differences between CD and CSD, I want to say a bit more about the assumption that the head of the compared constituent is a null morpheme. This is arguably a weak point in the analysis, since deletion or movement accounts of CSD appear to derive the obligatory presence of a 'gap' in this position. In fact, however, the traditional movement and deletion analyses also incorporate stipulations about null material. In a movement analysis, it must be stipulated that the moved degree term is phonologically null (as in, e.g., Chomsky 1977); in a deletion analysis, it must be stipulated that the bound degree head (*x-many/much* in Bresnan's (1973, 1975) representations; see (7)) never occurs overtly. If it could occur overtly, then we would expect a sentence like (61a) to allow an interpretation like (61b).

(61)a. Every quantity of beans should be served with much rice.

b. Every quantity of beans should be served with that much/the same quantity of rice.

As (61b) indicates, the degree use of *that* gives an approximation of the bound reading (an observation that Lees (1961) uses as the basis for his analysis of the comparative). Since this term cannot occur overtly in the comparative clause, however, a deletion analysis must 'derive' deletion by writing it into the rule.

If we treat the degree head of the compared constituent as a designated (null) lexical item, as suggested here, then the fact that the head position cannot be filled by an overt degree morpheme follows from the semantic

[10] In priniciple, the maximality operator could be introduced by the comparative morpheme itself, as in (ia), in which case it would be necessary to analyze the comparative clause as a set of degrees, as in (ib).

(i)a. $\text{er/more} = \lambda G \lambda D \lambda x . G(x) \succ max(D)$

b. $\text{Deg}_C^0 = \lambda G \lambda Q . \{d \mid Q(\lambda x . G(x) \succeq d)\}$

Assuming that the canonical syntactic expression of such meanings in English is the *wh*-construction (Cooper 1983; Jacobson 1995; Heim and Kratzer 1998), we still expect the compared constituent to undergo movement (or, in non-transformational frameworks, to participate in a corresponding relation). Since the proposed syntactic analysis is compatible with either view of where maximality in comparatives comes from, I will not attempt to resolve this issue here.

analysis presented above, in which the head of the compared constituent introduces the semantic property of maximality. If some other degree morpheme were inserted (e.g., *seven*, *very*, *how*, etc.), the comparative clause would be assigned the wrong meaning and the construction would be (semantically) ill-formed. Note also that in some languages (e.g., Greek, Bulgarian), the compared constituent is headed by an overt degree term (see note 20), which is exactly the sort of cross-linguistic variation we expect to see.

3.3. *The Differences*

3.3.1. *P-Stranding, COMP-Trace Effects, Contraction*

Turning now to the properties that differentiate CD and CSD, there is good evidence that the first three, P-stranding effects, COMP-trace effects, and contraction, involve conditions on overt movement only. When we look at expressions that are hypothesized to undergo covert Ā-movement – quantificational DPs and *in situ wh*-phrases – we see that they are acceptable in these environments. (The former provide a more reliable test, since there is some debate about whether *in situ wh*-phrases must raise at LF or not.)

For example, (62) shows that in Czech, quantificational DPs and *wh*-phrases may occur as objects of prepositions, despite the impossibility of P-stranding in overt movement constructions (see (20) above).

(62)a. Václav bydlel v každém evropském hlavním městě.
Vaclav live.PAST.3SG in every European capital
Vaclav has lived in every European capital.

b. Kdo bydlel ve kterých městech?
who live.PAST.3SG in which city.PL.LOC
Who lived in which city?

COMP-trace effects seem problematic at first, since it has been claimed that *wh-in situ* shows the same sensitivity to this constraint as overt movement. Kuno and Robinson (1972), for example, point out that (63) is ungrammatical:

(63) *I don't know who expects that who will visit Mary.

Likewise, the example in (64) does not have an interpretation in which the universal quantifier *each professor* takes scope over the existential *some student*, suggesting that the former cannot raise from its base position.

(64) Some student expected that each professor would talk about comparatives.

There is good reason to believe that whatever is going on in these examples should not be explained in terms of COMP-trace effects, however. First, even if we eliminate the complementizers in these examples, (63) remains ungrammatical and (64) still forbids a wide scope interpretation of *every professor*. (The former observation is in fact made by Kuno and Robinson, who are interested in establishing a broader, clause-mate restriction on the interpretation of multiple *wh*-questions.)

Second, if we look at a different matrix predicate, such as *make sure*, the facts change: (65) is a perfectly grammatical multiple question, and (66) has a reading in which *each speaker* takes scope over *some student* (Farkas and Giannakidou 1996).

(65) Which student made sure that which speaker got home safely?

(66) Some student made sure that each speaker got home safely.

These facts clearly indicate that the COMP-trace filter does not apply to covert movement constructions, since (63)–(64) and (65)–(66) do not differ with respect to the COMP/subject configurational relation. Whatever restrictions are active in examples like (63) and (64) must therefore stem from other factors (see Farkas and Giannakidou 1996 for relevant discussion).

Finally, the examples in (67) show that contraction is possible before an *in situ wh*-phrase.

(67)a, Who said there's how much rice?

b. Which team's how likely to win?

Quantificational DPs occur infrequently in predicative position, but when they do, contraction is possible, as shown by the examples in (68). ((68b) contains a quantificational possessive, which triggers QR of the entire DP of which it is a subconstituent, according to Barker (1995).)

(68)a. The new president's everything we expected him to be.

b. He's everyone's worst nightmare.

While the principles underlying these constraints certainly deserve explanations in their own right (see Anderson 2000 for a recent PF-based account of COMP-trace effects; see also Chomsky and Lasnik 1977 and Honegger 1996), for the purposes of this paper it is enough to observe that if CSD involves covert movement, then it should have properties similar

to quantifiers and *in situ wh*-phrases. The facts discussed above verify this prediction.

3.3.2. *Multiply-Headed Comparatives*

The syntactic properties of multiply-headed comparatives can also be traced to the overt/covert movement distinction. Recall from the discussion in section 2.2 that multiply-headed CSD constructions are well-formed, but multiply-headed CD is impossible:

(69)a. Christmas makes as many children as happy as it makes adults unhappy.

b. *Christmas makes as many children as happy as birthdays make.

The acceptability of multiply-headed CSD follows from the proposed analysis. If CSD involves covert movement, then multiply-headed comparatives are syntactically parallel to multiple *wh*-questions or sentences with more than one quantificational DP, which have been claimed to involve multiple instances of covert Ā-movement (Higginbotham and May 1981; Huang 1982; May 1985, etc.).

However, although English allows multiple instances of covert movement to SpecCP, multiple instances of overt movement to SpecCP are in general disallowed, as illustrated by (70).

(70) *Which children how happy will Christmas make?

Multiply-headed CD should therefore also be impossible, since the derivation of a sentence like (69b) would be completely parallel to (70).

The constraint against multiple Ā-movement in English is not absolute, however. As pointed out by Baltin (1982), there are some well-formed cases of multiple Ā-movement, such as (71).

(71) He's a man to whom liberty we could never grant.

The prediction of my analysis is that multiple CD should be possible in the same context. That is, all other things being equal, multiply-headed CD and multiple *wh*-movement should pattern together. (72), the relevant test, is in fact more acceptable than the examples of multiple-CD discussed above.[11]

(72) ?Clinton granted more pardons to more questionable applicants than any other president has ever granted.

[11] There is an independent problem with (72): the second instance of CD deletes the preposition *to*, which is typically not allowed:

(i) Clinton granted pardons to more applicants than he granted favors ?(to).

More generally, the analysis makes the following cross-linguistic prediction: all other things being equal, if a language allows multiple instances of overt Ā-movement, it ought to allow multiply-headed CD constructions. The 'all other things being equal' caveat is crucial here since languages differ widely in the syntactic resources used to express comparison (see section 5.2). However, initial support for this prediction comes from Japanese, a language that allows multiple scrambling. In this language multiply-headed CD is possible, as illustrated by (73). (I am grateful to Hajime Hoji for bringing this fact to my attention.)

(73) John-wa kimi-ga kaw-ase-ta-yori(mo) motto
John-TOP you-NOM buy-CAUSE-PAST-THAN more
ooku-no dansei-ni motto ooku-no kuruma-o
QUANTIFY-GEN male-DAT more QUANTITY-GEN car-ACC
kaw-ase-ta
buy-CAUSE-PAST
John made more men buy more cars than you made buy.

Although multiple instances of CD are in general ruled out in English, the analysis correctly predicts that 'mixed' multiply-headed comparatives such as (74a) should be well-formed, since such constructions involve only one instance of overt movement, namely the one that targets the deleted constituent, as shown by (74b).

(74)a. Max persuaded more people to buy more cars than you persuaded to buy trucks.

b. Max persuaded more people to buy more cars than [$_{CP}$ ~~[$_{DP}$ people]~~ you persuaded ~~[$_{DP}$ people]~~ to buy [$_{DP}$ trucks]]

The acceptability of mixed multiply-headed comparatives, and the fact that in multiple questions in English, one *wh*-phrase can (and must) remain *in situ* suggest that it ought to be possible to find examples of multiply-headed CD in which one of the compared constituents remains *in situ* and undeleted. As we will see in section 5.1.1, such constructions are in fact grammatical, and moreover provide independent evidence for the Optimality Theoretic analysis that I will present in section 4.

We may conclude therefore that the relative acceptability of (72) compared to other cases of multiple CD bears out the predictions of the analysis.

3.3.3. *Parasitic Gaps*

The difference between CD and CSD regarding parasitic gap licensing can also be explained in terms of the overt/covert movement distinction. That CD licenses parasitic gaps is expected, since the movement postulated to occur in a comparative like (75a) is exactly the same as the movement in a question like (75b); the only difference is that in (75a), the moved constituent is deleted.

(75)a. I threw away more books than [$_{CP}$ ~~[$_{DP}$ books]$_i$~~ I kept ~~[$_{DP}$ books]$_i$~~ without reading e_i]

b. [$_{CP}$ [$_{DP}$ How many books]$_i$ did you keep ~~[$_{DP}$ how many books]$_i$~~ without reading e_i]

The fact that CSD does not license parasitic gaps also follows, given the well-known constraint that parasitic gaps are dependent on overt Ā-movement (Engdahl 1983; see Nissenbaum 1998 for a new analysis of this requirement). Since the compared constituent in (76a), like the *in situ Wh*-phrase in (76b), does not move overtly, it should fail to license a corresponding parasitic gap.

(76)a. *I threw away more books than [$_{CP}$ I kept [$_{DP}$ papers]$_i$ without reading e_i]

b. *[$_{CP}$ 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 in Nissenbaum (2000), covert Ā-movement licenses a parasitic gap if there already exists another parasitic gap chain that is licensed by overt movement. This is illustrated by the contrasts in (77) (from Nissenbaum 2000, (2a)–(2b)).

(77)a. Which senator$_i$ did you persuade to borrow which car$_j$ after getting an opponent of e_i to put a bomb in e_i

b. *Which senator$_i$ did you persuade to borrow which car$_j$ after putting a bomb in e_j?

As pointed out to me by Jon Nissenbaum (personal communication), CSD also licenses parasitic gaps in precisely the same context, providing further support for the claim that it involves covert Ā-movement:

(78)a. I persuaded as many senators to buy as many cars as you persuaded to buy 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 you persuaded to buy trucks$_j$ after putting bombs in e_j.

The examples in (78) are mixed multiply-headed comparatives, in which one of the compared constituents undergoes overt movement (CD) and the other remains *in situ* (CSD). Just as with the multiple wh-questions in (77), the *in-situ* compared constituent licenses a parasitic gap only if the moved compared constituent also licenses a parasitic gap.

3.3.4. *The Left Branch Constraint*

Finally, the analysis avoids the problems with the Left Branch Constraint that arise in other movement analyses, even though it claims that CD and CSD involve Ā-movement. In the analysis defended here, 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 (79) are structurally analogous not to the questions in (80), as on standard movement analyses, but rather to those in (81), which are perfectly well-formed.

(79)a. Michael has more scoring titles than Dennis has (tattoos).

b. Michael's hands are wider than your feet are (long).

(80)a. *How many does Dennis have tattoos?

b. *How (much) are your feet long?

(81)a. How many tattoos does Dennis have?

b. How long are your feet?

The end result is that the Left Branch Constraint, however it is formalized (see Kennedy and Merchant 2000a for a recent proposal), does not come into play. I therefore agree with Izvorski (1995) that the reason that CSD

does not trigger Left Branch effects is that it does not involve movement of a left branch degree term.

4. Optimality in Syntax

4.1. *Movement, Identity, and Deletion*

The previous section demonstrated that the empirical similarities and differences between comparative deletion and subdeletion in English follow from a syntactic analysis in which the compared constituent moves overtly in CD and covertly in CSD. This approach achieves a level of descriptive adequacy not matched by earlier uniform analyses of comparatives, which do not satisfactorily explain the differences between CD and CSD. At the same time, it achieves a higher level of explanatory adequacy than non-uniform analyses, since it straightforwardly derives the similarities between the two constructions. However, the assumption that is crucial to achieving these results – that CD involves overt movement and CSD covert movement – is, at this point, a stipulation.

In fact, an analysis similar in spirit to the one proposed here is considered by Bresnan (1975), who rejects it precisely because of its apparently stipulative and *ad hoc* nature (Borsley 1984, p. 281 makes a similar objection). In particular, Bresnan (1975, p. 63) objects that:

> To guarantee that only the maximally recoverable constituent is moved [in comparatives], 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.

In particular, Bresnan is concerned about the impossibility of examples like the following ((82a) is from Bresnan 1977, example (29a)):

(82)a. *She has more boyfriends than books she has.
(cf. She has more boyfriends than she has books.)

b. *She has more boyfriends than I have boyfriends.
(cf. She has more boyfriends than I have.)

In Bresnan's analysis, these facts are explained by the assumption that comparatives involve deletion, not movement (which rules out (82a)), and the fact that the Relativized A-over-A Condition requires deletion of as much identical material as possible (ruling out (82b)).

The challenge for a movement analysis is to explain why we get movement and deletion whenever the compared constituent is identical to the

head (CD), but neither movement nor deletion when the head and compared constituent are not identical (CSD). My goal in this section is to demonstrate that this result follows from the interaction of general constraints on movement and deletion in an Optimality Theoretic framework (Prince and Smolensky 1993), in which syntactic constraints are ranked and violable and well-formedness is determined by evaluating competing syntactic representations against the constraint hierarchy. In short: *deletion is good and overt movement is bad, but it's better to delete than to avoid overt movement*.

Four constraints are involved in the analysis. First, following Grimshaw (1997) and Ackema and Neeleman (1998), I assume a constraint STAY which forbids movement. For the purpose of this paper, I will assume that STAY is violated by any single instance of movement, as in Grimshaw (1997).[12] Since I have assumed that movement is a copying operation, STAY can be defined as in (83), where α_{i+1} and α_i are copies of a single syntactic expression α.

(83) STAY: $^*[\alpha_{i+1} \ldots \alpha_i]$

Next we need a constraint that favors deletion. As already mentioned, I am adopting the basic architectural assumptions of the Principles and Parameters framework: I assume that the output of the computational system is a pair of syntactic representations ⟨LF, PF⟩. If we further assume the Late Insertion model of morphology developed in Halle and Marantz (1993) (see also Anderson 1992; Beard 1995; Bobaljik 1995; Zwart 1997; Lidz 1998), in which lexical insertion is a Post-syntactic operation (the 'interpretation' of a PF representation), deletion can be construed as an indication that a constituent in the PF representation should remain uninterpreted, and therefore unpronounced.

To make things precise, let us assume that constituents in a syntactic representation are feature structures of the sort common to work in constraint-based formalisms like Head-Driven Phrase Structure Grammar (Pollard and Sag 1994); see Veenstra (1998) for a formalization of P&P feature structures in these terms. In particular, let us assume that every terminal node has a MORPH feature that indicates how the node is to be interpreted by the morphological component (Veenstra's WORD feature), whose value is determined as a function of syntactic composition. (For example, tense or agreement features may be added/deleted by movement.)

[12] Ackema and Neeleman (1998) calculate violates of STAY in terms of the distance between a moved expression and its base position, building in a minimality requirement. This is arguably a better approach, but since length of movement is not a feature in the data I am discussing here, I will stick with the simpler version of the constraint in (83).

The view of deletion presented above can then be implemented by allowing the computational system to generate representations in which the value of a node's MORPH feature is NIL. For notational convenience, I will represent a feature/value combination [MORPH: NIL] on a node α by striking out α: $\overline{[_{\alpha} \ldots]}$.

As a starting point, I will make the strongest possible assumption about deletion: the grammar contains a constraint DELETE which is violated by any non-NIL MORPH feature as expressed in (84).

(84) DELETE: $\overline{[_{\alpha} \ldots]}$

DELETE can be thought of as a grammaticization of the functional notion of 'economy of effort' familiar from work on speech production (see e.g., Lindblom 1990): the best output of the syntax is one that requires the least amount of work for the production system. This constraint is necessarily violated by just about every well-formed sentence, but this is not a problem: it simply illustrates the fact that constraint violations are tolerated in order to satisfy higher ranked constraints.

In this case, the crucial higher ranked constraint is one that requires deletions to be recoverable. Such a constraint played a fundamental role in early work in generative syntax, in particular in the analysis of 'long distance' deletion rules, later reanalyzed as movement. (See in particular the discussion in Chomsky 1965, pp. 144–145, 182–183.) Given the return to a theory of movement that crucially incorporates a deletion operation (the 'copy and delete' theory advocated in Chomsky (1993) and a wide range of subsequent work in recent years), the implementation of a recoverability constraint once again becomes a central issue. Here I will adopt a formalization of recoverability as in (85), and I will assume it to be undominated (Fiengo and Lasnik 1972).[13]

(85) RECOVERABILITY: For any constituent α, if $\overline{[_{\alpha} \ldots]}$, then there is a constituent $\beta \neq \alpha$ such that β is recoverable and ID(α, β), where ID is a grammatical identity relation.

(85) requires any deleted constituent to stand in a 'grammatical identity relation' with some other constituent; I assume the set of grammatical identity relations to include at least the copy relation that holds between chain

[13] On the surface, it would seem that RECOVERABILITY has to be undominated. This is an issue that deserves closer scrutiny, however, as there are certain phenomena – e.g., 'pro-drop' in languages without agreement morphology, deletion of prepositions in certain relative clauses – which may be insightfully explained in terms of violations of RECOVERABILITY. I am grateful to Rodrigo Gutiérrez Bravo and Judith Aissen for discussion of this point.

elements and the identity relation that holds between an elided constituent and its antecedent. The former is a local relation, so in the case of deletion in movement chains (85) can be evaluated by looking at a single syntactic representation. On the surface, however, it appears that in at least some cases of deletion in ellipsis, RECOVERABILITY can be evaluated only by examining the syntactic representations of other sentences in the discourse (see Hankamer 1979, pp. 291–293). Introducing this sort of 'global search' requirement into the evaluation metric for particular sentences is clearly undesirable, but as we will see in section 4.3, with the proper formulation of the identity relation involved in ellipsis, this can be avoided.[14] Note that (85) is vacuously satisfied by any non-deleted constituent.

The three constraints discussed above are general; the fourth is specific to comparatives. Building on the analysis of questions in Ackema and Neeleman (1998), I assume a constraint C-SCOPE that governs the syntax of the comparative clause.

(86) C-SCOPE: The compared constituent must occupy the specifier of the complement of *than/as*.

I assume this constraint to make the analysis clear, but C-SCOPE should more properly be thought of as a specific instantiation of a more general constraint that requires expressions to occupy syntactic positions appropriate for their semantic types. In the case of comparatives, the compared constituent must be in SpecCP in order to derive the right interpretation for the comparative clause, in the same way that, e.g., the 'null operator' in a relative must be in SpecCP in order to derive the right interpretation for a relative clause.

With these constraints in hand, the grammatical organization needed to explain the properties of English comparatives is as follows. First, since I have assumed a syntactic framework with two levels of representation, I allow for the possibility that the different interfaces have different constraint rankings, with corresponding structural differences in optimal outputs.[15]

[14] I am assuming here that the licensing conditions for deletion in movement chains differ from the licensing conditions for deletion in ellipsis, even though the end result in both movement and ellipsis (non-pronunciation of syntactic material) is realized by the same mechanism (deletion of phrasal material). That movement and ellipsis are subject to different licensing conditions is well-established: in Williams' (1977) terms, the former is part of S(entence)-grammar and the latter is part of D(iscourse)-grammar (see also Hankamer 1979).

[15] Whether the entire set of constraints that apply at LF or PF are the same (just under different rankings), or whether the two interface levels have disjoint sets of constraints is a question that I will not address here. What seems most likely is that constraints governing

To derive the result that all comparatives involve movement of the compared constituent to SpecCP at LF, then, the grammar must have the LF constraint ranking in (87).

(87) *LF Ranking*
C-SCOPE ≫ STAY

Again, if C-SCOPE is actually a specific instance of a more general constraint that says that expressions need to be in a syntactic position appropriate for their type, this ranking amounts to saying that covert movement for semantic reasons is licensed in English, a common assumption in the P&P framework (see Heim and Kratzer 1998).

Such movement is generally not licensed overtly, however, indicating that STAY outranks C-SCOPE at PF. This ranking forbids any overt movement in comparatives, so in order to derive movement in CD, some other constraint must outrank STAY. I claim that this constraint is DELETE, giving the PF ranking in (88).

(88) *PF Ranking*
DELETE ≫ STAY ≫ C-SCOPE

The result of this ranking is that overt movement should in general be dispreferred, but, all other things being equal, syntactic representations in which movement feeds deletion will be preferred to representations which avoid movement but fail to delete. (At the same time, representations that effect deletion without movement should be best of all, a point that I will return to in section 5.1.2.) In the next section, I will show that this ranking of PF constraints derives the overt/covert movement distinction between CD and CSD.

4.2. *The Grammar of Comparison in English*

4.2.1. *Comparative Deletion*

Let us begin by examining comparative deletion. Recall that the desired output is a PF representation in which the compared constituent moves and deletes. Looking at the relevant candidate PFs for an example like (89), however, we see that there is a problem. (In the following tableaux,

purely syntactic relations, such as STAY, C-SCOPE, and so forth should be shared, but constraints that are 'interface specific' (such as DELETE) should be localized to particular interfaces.

I will indicate only relevant violations of DELETE, which is actually violated by any overt material, as noted above. I will also generally omit the undominated RECOVERABILITY constraint.)

(89) The galaxy contains more stars than the eye can see.

Tableau I. *Comparative deletion* (wrong)

		DELETE	STAY
a.	than [$_{CP}$ the eye can see [$_{DP}$ stars]]	*!	
b.	√than [$_{CP}$ the eye can see [$_{DP}$ stars]]		
c.	than [$_{CP}$ [$_{DP}$ stars] the eye can see [$_{DP}$ stars]]	*!*	*
d.	than [$_{CP}$ [$_{DP}$ stars] the eye can see ~~[$_{DP}$ stars]]~~	*!	*
e.	than [$_{CP}$ ~~[$_{DP}$ stars]~~ the eye can see [$_{DP}$ stars]]	*!	*
f.	than [$_{CP}$ ~~[$_{DP}$ stars]~~ the eye can see ~~[$_{DP}$ stars]]~~		*!

As the tableau indicates, the optimal candidate is not the desired (f), but rather (b), in which the compared constituent is deleted *in situ*, in satisfaction of both DELETE and STAY. This is in fact the structure assigned to (89) by early transformational analyses of CD (such as Lees 1961; Chomsky 1965; Bresnan 1973 and Carlson 1977), but within the larger theoretical framework adopted here, this is clearly the wrong result: even though (b) and (f) are string identical, the range of evidence examined in sections 2 and 3 indicates that CD involves overt movement. In order to derive (f) as the optimal output, then, the (b) candidate must be ruled out by some higher-ranking constraint(s) which have the effect of making movement a necessary condition for deletion of the compared constituent in CD.

In fact, if we take a closer look at the way in which deletion is licensed, we can find arguments that exactly this conclusion is correct. I have assumed here that deletion is part of the set of operations that build candidate PF representations. Actual instances of deletion in particular PFs still need to be licensed, however, and in English, the two most common modes of licensing phrasal deletion are movement and ellipsis. (For arguments that ellipsis involves deletion, rather than copying or recovery of semantic content, see Merchant 2001; Kennedy and Merchant 2000a; Kennedy to appear) In order to determine which structure is optimal in Tableau I, then, we also need to ask how the various deletions indicated in the tableau are licensed.

Let us consider first whether the deletions indicated in Tableau I are licensed by ellipsis. (90) shows that the NP complement of certain overt determiners may be deleted when its content is recoverable.

(90)a. Kim could see stars, but I couldn't see [$_{DP}$ any ~~[$_{NP}$ stars]~~]

b. Kim heard about Lee's discovery before I heard about [$_{DP}$ Pat's ~~[$_{NP}$ discovery]~~].

Examples like (91) indicate that null determiners do not license NP-ellipsis, however. Similarly, full DPs cannot be elided, even when the information they convey is recoverable, as illustrated by (92).

(91) *Kim could see stars, but I couldn't see [$_{DP}$ ~~[$_{NP}$ starts]~~].

(92)a. *Kim heard about Lee's discovery before I heard about ~~[$_{DP}$ Lee's discovery]~~.

b. *Kim knows an astronomer who married ~~[$_{DP}$ an astronomer]~~.

Assuming that these facts reflect the influence of some highly ranked constraint(s) in the grammar of English (which I will claim to be RECOVERABILITY in section 4.3), we may conclude that deletion of the compared constituent in candidate (b) in Tableau I cannot be licensed by the principles of ellipsis: (b) cannot be derived by NP-ellipsis, since the head of the compared constituent is null, and it cannot be derived through DP-ellipsis, since this is not possible in English.[16]

Deletion of a DP is licensed when the DP is part of a movement chain, however, assuming the copy and delete theory of movement. On this view, movement is not literal displacement of a constituent, but is rather a complex operation consisting of a copying component, which builds a chain, and a deletion component, which eliminates redundant copies from the chain at PF. The structure of a question like (93a) is therefore (93b), where the lower copy of the chain is deleted

(93)a. How many stars can you see?

b. [$_{DP}$ How many stars] can you see ~~[$_{DP}$ how many stars]~~

[16] Note that I am not saying that the principles of ellipsis are *never* active in licensing deletion in comparatives, only that they cannot license deletion of the compared constituent in CD. If deletion of the compared constituent could be licensed by the principles of ellipsis, then indeed a candidate like (b) in Tableau I should be optimal (all other things being equal). In fact, this turns out to be the case, as we will see in section 5.1.2.

Typically, all of the lower members of a chain are deleted (in English), but the head of the chain remains. Within the optimality theoretic framework adopted here, the fact that exactly one copy must remain follows from the interaction of RECOVERABILITY and DELETE: the latter maximizes deletion, while the former is violated only if no overt copy remains in the structure.[17]

Deletion of the highest copy in a chain is in principle an option, however, and would in fact be optimal if doing so would not violate RECOVERABILITY. In the case of CD in particular, if a grammatical identity relation could be established between the copy of the compared constituent in SpecCP and the head of the comparative, then RECOVERABILITY would be satisfied and DELETE would actually force maximal deletion of all chain copies.[18] Ideally, we would like to avoid introducing a new identity relation into the grammar, but there is reason to believe that there is in fact an identity relation that holds between the head of the comparative and the copy of the compared constituent in SpecCP that has properties distinct from both movement and ellipsis.

Evidence that this relation cannot be reduced to ellipsis comes from locality effects. As shown in Kennedy (1998, 1999) (see also Williams 1977 and Hazout 1995), comparative deletion obeys a locality requirement that is not seen in ellipsis constructions. This is illustrated by the following pair of examples.

(94)a. I usually buy books, but today I bought records when I couldn't find any. (~~books/records~~)

b. ?I usually buy books, but today I bought more records than I bought. (*~~books/records~~)

(94a) shows that an elided constituent can find its antecedent nonlocally: *books* can license deletion of the nominal complement of *any*. The deleted nominal can also be licensed by the local phrase *records*, but this reading is dispreferred since it results in an anomalous interpretation. In contrast,

[17] The constraints I have adopted here do not require the highest copy to be retained in the PF form, as opposed to some other chain element, however. I assume that this aspect of chain pronunciation is controlled by some other (set of) constraint(s). See Runner (1995), Bobaljik (1995, 1999), Pesetsky (1997), Franks Bošković (2000) for relevant discussion.

[18] Such a relation between an external head and a constituent in SpecCP of a modifier clause is proposed in the recent analyses of relative clauses developed by Cresti (2000) and Sauerland (1998) (see also Chomsky 1965, p. 182). Cresti takes this relation to be ellipsis, and Saurland shows that it does has properties similar to ellipsis but different from movement (though he does not address the question of what exactly the licensing relation is), but as we will see below, there is reason to think that it is unique.

(94b) shows that deletion of the compared constituent can only be licensed locally (by the head of the comparative) even when this results in an anomalous (in this case, contradictory) interpretation. A sensible interpretation of (94b), in which the compared constituent is understood as identical to a nonlocal phrase, is unavailable.

An alternative approach would be to reduce the relation between the head of the comparative and the chain in the comparative clause to movement; i.e., to postulate that the head of the comparative is derived from the clause-internal position of the gap as proposed in Rivero (1981), Lechner (1999) and Kayne (1994) (cf. 'head raising' analyses of relative clauses). This sort of analysis would have the additional advantage of deriving the locality effects observed in (94b), as pointed out by Lechner (1999), since the head and the compared constituent would literally be the same phrase. There are a number of empirical arguments against this sort of approach, however. Since the arguments require some detailed discussion, and since such an analysis represents a more general challenge to the approach advocated here, I will postpone discussion of it until section 6.

Given these considerations, I conclude that there is a unique, local grammatical identity relation that compares the head of the comparative and an element in SpecCP of the comparative clause.[19] If this is correct, then the optimal candidate in comparatives in which the compared constituent and the head are identical – i.e., CD constructions – will be one in which the compared constituent moves to SpecCP, licensing deletion of all copies of the chain. This is illustrated in Tableau II, which is a revised look at the competing PF representations of the comparative clause in (89). (The (b) candidate from Tableau I has been eliminated on the assumption that it violates a higher-ranked constraint, as argued above.)

Tableau II. *Comparative Deletion* (right)

		DELETE	STAY
a.	than [$_{CP}$ the eye can see [$_{DP}$ stars]]	*!	
b.	than [$_{CP}$ [$_{DP}$ stars] the eye can see [$_{DP}$ stars]]	*!*	*
c.	than [$_{CP}$ [$_{DP}$ stars] the eye can see ~~[$_{DP}$ stars]~~]	*!	*
d.	than [$_{CP}$ ~~[$_{DP}$ stars]~~ the eye can see [$_{DP}$ stars]]	*!	*
e.	✓ than [$_{CP}$ ~~[$_{DP}$ stars]~~ the eye can see ~~[$_{DP}$ stars]~~]		*

[19] I also expect that this is the same relation that holds between the head of a relative clause and a moved internal head, and that the analysis of comparatives that I develop here should extend to (at least some) relative clauses, though a full exploration of this issue is beyond the scope of this paper. See Sauerland (1998) and Kennedy (in preparation) for relevant discussion.

An important result of this analysis is that CD is analyzed as ordinary Ā-movement: it is not necessary to postulate a special rule of 'comparative deletion' or 'comparative movement' in order to generate the right output. The movement operation in CD appears special only because all copies created by movement are deleted, a result that follows from the assumption that DELETE is a general constraint requiring as much material as possible to be eliminated from the PF representation, and the hypothesis that the compared constituent in SpecCP can establish an identity relation with the head of the comparative, satisfying RECOVERABILITY.[20]

Note that the analysis also makes deletion in CD obligatory, explaining facts like (4) and Bresnan's (82b). Moreover, maximality of deletion follows directly from the architecture of the OT framework; it does not have to be stipulated in a separate constraint, as was the case in Bresnan's original analysis with the Relativized A-over-A condition. This feature of the analysis sets it apart from approaches that rely on ellipsis to handle the elimination of redundant material, since ellipsis is optional (and does not have a maximality requirement, though maximal deletions are often preferred). This is clearly a point in the proposal's favor, but it also raises a new question: if ellipsis also involves deletion, and deletion is subject to the same constraints discussed here, why is it optional? I address this question in section 4.3.

4.2.2. *Subdeletion*

The crucial difference between subdeletion and comparative deletion is that in the former, the compared constituent is not identical to the head. It follows that any deletion of the compared constituent would violate RE-

[20] Strictly speaking, it should be the case that the D^0/Deg^0 head of the compared constituent in SpecCP is not deleted. (I am grateful to Eric Potsdam for bringing this point to my attention.) Since its meaning differs from the morphologically comparative D^0/Deg^0 on the head of the comparative, its content is not recoverable (see the semantic analyses of degree morphology in section 3.2). Because this item is phonologically null in English, we have the effect of full deletion. If it had content, however, we would expect it to remain in the phonological representation (like *which*, in *which*-relatives). This is exactly what we see in, e.g., Bulgarian ((i), from Izvorski 1995) and Greek ((ii), from Rudin 1984).

(i) Ivan izpi poveče vino ot-kolkoto bjahme kupili.
Ivan drank more wine from-DEG were-1PL bought
Ivan drank more wine than we bought.

Exis perisotera vivlia ap' osa exo ego.
have-2PL more books than DEG have-1SG I
You have more books than I have.

COVERABILITY. If RECOVERABILITY is an undominated constraint, the result is that STAY emerges as the crucial factor when evaluating candidate PF representations of CSD. All relevant candidates equally violate DELETE, therefore the optimal representation is one in which the compared constituent remains *in situ*. This is illustrated by Tableau III, which shows the competing candidates for the comparative clause in (95).

(95) Michael has more scoring titles than Dennis has tattoos.

Tableau III. *Subdeletion*

	DELETE	STAY
a. √ than [$_{CP}$ Dennis has [$_{DP}$ tattoos]]	*	
b. than [$_{CP}$ [$_{DP}$ tattoos] Dennis has ~~[$_{DP}$ tattoos]~~]	*	*!
c. than [$_{CP}$ [$_{DP}$ ~~[$_{DP}$ tattoos]~~ Dennis has [$_{DP}$ tattoos]]	*	*!
d. than [$_{CP}$ [$_{DP}$ tattoos] Dennis has [$_{DP}$ tattoos]]	**!	*

Viewed from the perspective of a model that evaluates syntactic well-formedness in terms of 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 grammaticality because all relevant candidates equally violate some higher-ranked constraint(s).[21] In particular, the (b) candidate (movement without deletion), which is parallel to Bresnan's (82a), is ruled out by STAY.

[21] An apparent problem for this analysis is that it appears to incorrectly predict that comparatives should allow *wh-in situ* when the compared constituent is a *wh*-word, as in some dialects (see Hankamer 1979). This is incorrect.

(i)a. John bought more books than what Bill bought.

b. *John bought more books than Bill bought what.

There are two possible solutions to this problem. The first is that *what* in these examples is an overt realization of the Deg^0 head of the compared constituent, in which case examples like (ia) are parallel to the Bulgarian and Greek comparatives discussed in note 20: the structure of (ia) is (ii), which is optimal because movement licenses deletion of the lexical component of the compared constituent.

(ii) John bought more books than [$_{CP}$ [$_{DP}$ what ~~[$_{NP}$ books]~~] Bill bought ~~[$_{DP}$ what [$_{NP}$ books]]~~]

If, on closer scrutiny, *what* turns out to be phrasal in these examples, a second solution is that the contrast between (ia) and (ib) is a fact about *wh*-XPs, rather than a fact about movement in general. That is, the reason that movement must occur in (ia) is because *what* introduces specific syntactic requirements that can only be met through movement. This is a standard assumption about *wh*-XPs, which is independently necessary to capture the

Given that the (a) and the (c) candidates in Tableau 4.2.2 are string-equivalent, it is worth considering an even more general analysis of CD and CSD, in which both constructions involve overt movement of the compared constituent, but in CSD it is only the lower copy that is pronounced (as in the theories of covert movement advocated in, e.g., Bobaljik 1995, 1999 and Pesetsky 1997; cf. Brody 1995). On this view, the role of the constraint system is to determine which copy of a chain (if any) is pronounced, rather than whether a constituent is overtly moved or not.

Recent work by Bosković (2000) argues against this approach, however (for CSD at least; it may be correct for other cases of 'covert' movement). Boskovic' argues that in certain contexts, multiple *wh*-questions in Romanian have exactly the structure of the (c) candidate: the lowest copy in a *wh*-chain is pronounced instead of the highest one (due to phonological constraints which prohibit realization of the higher copy; see also Franks 1998). As evidence that movement has actually occurred, Bošković shows that in just these environments, the apparently unmoved *wh*-phrase licenses a parasitic gap. This result entails that if CSD were analyzed in the same way (i.e., if the (c) structure were correct), it should also license parasitic gaps, contrary to fact. I therefore conclude that the (a) candidate is in fact the correct PF representation of CSD.

A final point to make about this analysis of subdeletion is that it explains why a compared constituent that is lexically identical to the head but contrastively focused is not deleted, as in (96) (see (5)). (Focus can actually fall on either the head or the compared constituent, as pointed out in Sag 1976, pp. 235–236.)

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

Assuming that the occurrence of the adjective *high* in the comparative clause in (96) bears a focus feature, this type of example can be explained in exactly the same way as the more typical cases of subdeletion: deletion of the compared constituent would entail deletion of the focus feature, and so would violate RECOVERABILITY, since the head and the compared constituent would not be identical.

---

fact that they must undergo overt Ā-movement in English, while other XPs (e.g., quantificational DPs) undergo Ā-movement only covertly. This hypothesis can be implemented either by ranking STAY below specific constraints that make reference to *wh*-XPs (such as Grimshaw's (1997) OPSPEC or Ackema and Neeleman's (1998) Q-MARKING).

4.2.3. *Partitive Comparatives*

The data discussed in 2.3 demonstrated that partitive comparatives in both their CD and CSD forms behave like standard CD constructions in not showing a CD/CSD dichotomy. That partitive CD behaves like standard CD is unsurprising: a sentence like (97a) can be given exactly the same analysis as (97b): the compared constituent raises to SpecCP and is deleted in accord with the principles described in section 4.2.1.

(97)a. Kim read more of the books than Lee read.
Kim read more books than Lee read.

Partitive CSD does not behave like standard CSD, however; instead, it looks more like standard CD. In particular, partitive CSD shows anti-contraction effects, the compared constituent cannot occur in subject position at all, and multiply-headed partitive CSD is ungrammatical. This is illustrated by the examples in (98); compare the standard CSD examples in (99).

(98)a. *There's more of the fish than there's of the rice in the pot.

b. *More of the old books were read than I thought (that) of the new books were read.

c. *Max sold as many of the men as many of the cars as you sold of the women of the trucks.

(99)a. There's more fish than there's rice in the pot.

b. More old books were read than I thought (that) new books were read.

c. Max sold as many men as many cars as you sold women trucks.

These facts would follow if all partitive comparatives are syntactically instances of CD, in that they involve overt movement of the compared constituent. This is the position advocated by Grimshaw (1987), who claims that an example like (100a) involves extraposition of the *of*-PP plus movement of a partitive DP, as shown in (bob). (For clarity, I will use trace notation to indicate extraposition.)

(100)a. I read more of the abstracts than I read of the articles.

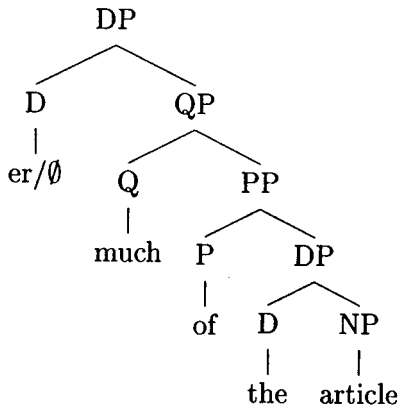
b. . . . than [$_{CP}$ ~~[$_{DP}$ much t_i]~~ I [$_{VP}$ [$_{VP}$ read ~~[$_{DP}$ much t_i]~~ [$_{PP}$ of the article]$_i$]

The challenge is to explain why this should be so. At the same time, we need to ensure that examples like (101), in which the compared constituent moves but only part of it deletes, are not generated (cf. Bresnan 1975, p. 63).[22]

(101) *There aren't as many of us as of them there are.
(cf. There aren't as many of us as there are of them.)

In fact, these results can be made to follow from the syntax of partitives, plus one additional assumption about deletion in movement. First, I assume the basic structure of partitive comparatives to be one in which D^0 selects a functional projection headed by an amount term (*many*, *much*, *several*, etc.), which in turn selects a PP headed by *of*. On this view, the structure of the compared constituent in an example like (100a) is (102), where D^0 is either the comparative morpheme or the null head of the compared constituent, and QP is an arbitrary label for whatever functional category is involved in partitives (cf. Bresnan 1973; Corver 1997).

(102)



[22] Bresnan (1975) is actually worried about examples like (i).

(i) *There isn't as large a number of men as of women there is.
(cf. There isn't as large a number of men as there is of women.)

This example is somewhat different, since it involves an attributive adjectival comparative. The explanation is similar to the one that I present below, however. Assuming that the *of*-PP is the complement of *number*, the material that needs to be deleted to derive (i) – a *large number* (or *large a number*, if the adjective must invert; see Kennedy and Merchant 2000a) – is not a constituent. The only way to delete this string is to extrapose *of women* and raise the DP [*a large number t*], which can be deleted under identity with the head of the comparative.

Second, we need to assume that deletion of material in SpecCP targets phrases. Support for this assumption comes from the fact that so-called null operators, which in the current system can be analyzed as instances of deletion, do not pied-pipe lexical material (see Browning 1987 and Grosu 1994). Assuming the structure in (102), the result is that the elimination of the compared constituent in a partitive comparative is an all-or-nothing operation: if QP is deleted, then the PP complement of Q must be deleted along with it. The only way to avoid deleting the PP is therefore to extrapose it from the compared constituent.

If these assumptions are correct, then the constraint system proposed so far will conspire to derive the results we want in partitive CSD. DELETE favors movement of the compared constituent at the cost of a STAY violation, since this feeds deletion of the Q element *much/many*. At the same time, RECOVERABILITY rules out representations in which an *of*-PP distinct from the one on the head is deleted. Since deletion of the entire QP in SpecCP is the only option, this means that the PP must be left behind in examples of partitive CSD, even though extraposition incurs a second violation of (low-ranked) STAY. The evaluation of the comparative clause in (103) is shown in Tableau IV, with RECOVERABILITY added to the tableau for completeness. Again, I show only those violations of DELETE that are incurred by elements of the compared constituent.

(103) I read more of the abstract than I read of the article.

Tableau IV. *Partitive CSD*

	REC	DELETE	STAY
a. than [$_{CP}$ I read [$_{DP}$ much of the article]]		****!	
b. than [$_{CP}$ [$_{DP}$ much of the article] I read ~~[$_{DP}$ much of the article]~~]		****	*
c. than [$_{CP}$ ~~[$_{DP}$ much of the article]~~ I read ~~[$_{DP}$ much of the article]~~]	*		*
d. than [$_{CP}$ [$_{DP}$ much t_i] I read ~~[$_{DP}$ much t_i]~~ [$_{PP}$ of the article]$_i$]		****!	**
e. $\surd$ than [$_{CP}$ ~~[$_{DP}$ much t_i]~~ I read ~~[$_{DP}$ much t_i]~~ [$_{PP}$ of the article]$_i$]		***	**

According to this analysis, partitive 'CSD' is syntactically CD, in that it involves overt movement of a DP; this explains the fact that all partitive comparatives have CD-like properties (anti-contraction effects, no multiply-headed partitive comparatives). At the same time, examples of partitive CSD obligatorily involves extraposition of the *of*-PP. If this expression is not extraposed, as when it appears in subject position, the result is correctly predicted to be ungrammatical.[23]

[23] A question that remains to be answered is why extraposition does not save an example like (ia); (ib) appears to be just as bad.

4.3. *Ellipsis*

In section 4.1, I made the assumption that DELETE is a general constraint that is violated by any overt material. This is appropriate from a theoretical standpoint, as it grounds the constraint in a functional principle of reduction of effort on the production side. This position makes a strong prediction, however: DELETE should force elimination of any material that satisfies RECOVERABILITY, since representations in which recoverable material is eliminated are preferred to competing representations in which it is not. This result derives maximality of deletion in CD, but if it is correct, why is deletion apparently optional in ellipsis?

The problem is illustrated by the examples in (104). In addition to the basic CD structure in (104a), which does not involve ellipsis, there are two ellipsis variants: one involving VP-deletion (104b), and one involving stripping (104c) (assuming for the sake of argument that an ellipsis analysis of this sentence is at least possible; see Hankamer 1973; Napoli 1983; Lechner 1999).

(i)a. *More of the boys stayed than of the girls left.

b. *More of the boys stayed than left of the girls.

This is particularly puzzling considering the fact that the corresponding partitive *how*-questions, which should be syntactically parallel to the comparatives, behave as expected: the *of*-PP can be left behind only if it extraposes:

(ii)a. *How many did you say of the girls left?

b. How many did you say left of the girls?

A second issue that needs to be addressed in a future investigation of partitive comparatives is the fact that partitive CSD does not license parasitic gaps, as observed in section 2.3. If the analysis of partitive CSD presented above is correct, then it is not clear why this should be so, though it presumably tells us something interesting about parasitic gaps. Of relevance here is the fact that partitive *how many* questions behave in the same way as partitive comparatives: parasitive gaps are licensed only if the of-PP moves undergoes along with *how many*.

(iii)a. How many did you invite of the teachers after meeting *(them) at the open house?

b. How many of the teachers did you invite after meeting (them) at the open house?

(104)a. Michael won more scoring titles than Dennis won.

b. Michael won more scoring titles than Dennis did ~~[$_{VP}$ win]~~.

c. Michael won more scoring titles than Dennis ~~[$_{I}$ won]~~.

.

In fact, once it is taken into account that constituents targeted for ellipsis must be explicitly marked as such, the problem of optionality disappears. A deletion analysis of ellipsis must account for the 'communication' between LF and PF that ensures that only a constituent that stands in a relatively strict identity relation to some other constituent in the discourse may be elided. Merchant (2001) handles this problem by postulating an 'E-feature' (see also Kennedy and Merchant 2000a) that may optionally be assigned to certain heads (in English, I^0, D^0, C^0; see Merchant (2001) for detailed arguments that the identity relation in ellipsis is unique and must therefore be annotated in the syntactic representation). In Merchant's analysis, this feature has different interpretations at the two interface levels. At LF, it is interpreted as a specification that the sister of the marked constituent stand in an appropriate identity relation to some other constituent in the discourse. (For Merchant, this is a semantic identity relation, though other characterizations are also possible.) At PF, it is interpreted as an instruction to delete, which in the current system would be a requirement that the MORPH feature on the complement be NIL.

Within the framework of PF constraints proposed here, Merchant's E-feature can be analyzed exclusively in terms of its semantic contribution: explicit marking of a grammatical identity relation. If we assume that the interpretation of the E-feature is a specification that its complement XP is identical (in whatever sense is the right one for ellipsis) to some other constituent in the discourse, then the syntactic representation alone indicates that deletion of XP satisfies RECOVERABILITY. In contrast, deletion of an XP that is not the complement of an E-marked head (nor part of a movement chain), should violate RECOVERABILITY, since the syntactic representation provides no indication that XP is part of a grammatical identity relation. In other words, within the broader grammatical framework presented here, once we have explicit marking of a grammatical identity relation, deletion comes for free.

This result is illustrated in Tableau V, where (a)–(b) and (a′)–(b′) represent two different inputs: one in which a head bears the E-feature and one in which it does not. In the former case, deletion is obligatory; in the latter, it is ruled out. The underlying logic is that if the syntactic representation of a sentence explicitly indicates that a constituent stands in a grammatical identity relation to some other constituent, as in the case of ellipsis (or

sequences of copies in movement chains), then there is no need to include XP in the pronounced form. Deletion of an E-marked XP is therefore just as obligatory as deletion of a copy in a movement construction; the apparent optionality of ellipsis is really optionality of E-marking.[24]

Tableau V. *Deletion and Ellipsis*

		RECOVERABILITY	DELETE
a.	$\ldots\ Y^0_E\ [_{XP}\ \ldots\ ]\ldots$		*!
b.	$\surd\ \ldots\ Y^0_E\ [\sout{_{XP}\ \ldots\ ]}\ldots$		
a′.	$\surd\ \ldots\ Y^0\ [_{XP}\ \ldots\ ]\ldots$		*
b′.	$\ldots\ Y^0\ [\sout{_{XP}\ \ldots\ ]}\ldots$	*!	

An important part of this analysis is that even in cases of ellipsis, RECOVERABILITY is evaluated locally: it is not necessary to invoke a search of the discourse representation to check whether an elided constituent is identical to some other constituent (see the discussion of this issue in section 4.1). Instead, it is only necessary to look to the governing head to check whether it bears the E-feature: if it does, then deletion satisfies RECOVERABILITY; if it does not, then deletion violates this constraint. This does not mean that ellipsis is not anaphoric, however. It just means that the anaphoricity of ellipsis comes not from the deletion operation, but rather from the meaning of the E-feature, which states that the complement of an E-marked head has the same meaning as some other constituent in the discourse (see Merchant (2001) for a formalization of this idea in terms of Schwarzschild's 1999 approach to focus and accent).[25]

[24] An anonymous reviewer correctly points out that the E-feature is not a deletion feature on this view, but rather an identity feature, in that its meaning is an identity requirement. On the surface, this looks similar to the role played by indices, but this is incorrect. Indices are standardly interpreted as variables, which means that the index itself does not denote an identity relation. Rather, the effect of identity comes from how the larger structure in which the variables appear is evaluated. An assignment function assigns the same value to all like-named free variables, and an operator assigns the same value to all like-named variables that it binds.

Note also that I am not claiming that all cases of 'surface anaphora' involve the E-feature (or its equivalent): this is a property of ellipsis constructions specifically. For example, I do not assume that *do so* constructions, which are similar to ellipsis constructions in many ways, involve the E-feature. In fact, if the analysis of *do so* defended in Kehler and Ward (1999) is correct, we must assume that this expression is typically *not* recoverable, since it introduces unique presuppositions into the discourse. An occurrence of *do so* would be recoverable only if introduced by a head bearing the E-feature, which would further require another occurrence of *do so* in the discourse.

[25] The analysis therefore claims that a sentence like (ia), with the structure in (ib), is perfectly well-formed when uttered out of the blue (contra Hankamer and Sag 1976), but

Finally, this approach to ellipsis also allows us to explain the impossibility of the (b) candidate in Tableau I, in which the compared constituent is deleted *in situ*, in terms of RECOVERABILITY. As pointed out by Merchant (2001, pp. 60–61), the E-feature allows us to account for language-internal and cross-linguistic differences in the types of constituents that may be elided in terms of independently necessary mechanisms of feature licensing. If we assume that languages may impose restrictions on which heads may bear the E-feature – in English, certain types of C^0, I^0, and D^0 – then we can explain why only certain constituents elide. In particular, if we assume that neither V^0 nor null D^0 are possible hosts for the E-feature in English, gaps corresponding to full DPs would have representations parallel to (b′) in Tableau V above, and so would be ruled out by RECOVERABILITY.

4.4. *Summary*

To summarize, the rule of English comparative formation that emerges from the analysis presented here can be stated as in (105).

(105) *English Comparative Formation* (final)
Move the compared constituent to the specifier of the complement of *than*.

(105) is in effect a restatement of C-SCOPE, and is of course not a grammatical rule in the standard sense, but rather a description of a movement operation which I assume to be driven by independent properties of the compared constituent. In particular, I assume this movement to be driven by the semantic requirements of the comparative: given the semantic analysis of degree morphology adopted in section 3.2, the compared constituent must move as stated in (105) in order for the comparative clause to be properly interpreted as a definite description of a degree.

Crucially, within the analytical framework developed in the previous sections, this is all that needs to be said about the syntax of comparatives specifically: the distinction between overt and covert movement and the obligatoriness of deletion in CD follow from the different rankings of STAY and C-SCOPE at LF and PF and the relative ranking of RECOVERABILITY,

is anomalous because the semantic requirements imposed by the E-feature (roughly, 'find an antecedent VP in the discourse') cannot be met.

(i)a. ?Kim can.

b. Kim can$_E$ ~~[$_{VP}$ juggle]~~

DELETE and STAY at PF. All other things being equal, whenever movement provides the sole means of generating a representation that maximizes satisfaction of DELETE – whenever the compared constituent is identical to the head (CD) – it must occur. Conversely, whenever movement does not optimize deletion – whenever the compared constituent is not identical to the head (CSD) – it must not occur.

This result is quite different from what we would obtain in, for example, a syntactic framework in which movement is driven exclusively by feature strength, as in most work in the Minimalist Program. In order to capture the distinction between CD and CSD in this type of system, we would need to make the *ad hoc* stipulation that the features on the compared constituent are strong when it is identical to the head and weak when it is not. This proposition is not only falsified by data to be discussed in the next section, which show that even in some cases of identity (CD) the compared constituent does not move, it is also completely devoid of explanatory power. Such an approach would indeed be subject to Bresnan's (1975) objections, since it would be " ... accidental that the moved constituents undergo deletion and that the elements moved just happen to be those which would be maximally recoverable if deleted".

5. More Evidence for Optimality

5.1. *Identity without Movement*

This section provides language-internal support for the Optimality Theoretic analysis of comparatives presented here by showing that there are contexts in which a compared constituent is identical to the head of the comparative, but it does not move overtly. On the surface, such cases appear to violate the principles that were introduced in the previous section to derive the relation between identity and movement in comparatives. However, this apparent contradiction is precisely why these facts constitute an argument in favor of an OT analysis. As we will see below, the principles that normally force movement in such contexts are violated precisely because doing so results in a representation that better satisfies the overall set of constraints on PF representations.

5.1.1. *Multiply-Headed Comparatives*

A surprising piece of evidence in favor of the analysis of subdeletion developed in section 4.2.2 comes from a set of data involving multiply-headed comparatives. We have already seen that multiply-headed CD is ruled out because it would require two instances of overt Ā-movement.

For any input containing a pair of identical comparisons, however, there are other possible outputs that must be considered in addition to the one that corresponds to the structure with multiple Ā-movements. In particular we need to consider outputs in which only one of the two compared constituents is moved and deleted, while the second remains *in situ* and overt. (I am very grateful to Marcus Hiller for suggesting this to me.)

The examples in (106) and (107) illustrate the various possibilities we need to consider, ranging from movement and deletion of both compared constituents, to no movement or deletion at all.

(106)a. *University officials convinced more researchers to join more committees than they had ever convinced to join before.

b. *University officials convinced more researchers to join more committees than they had ever convinced researchers to join before.

c. ?University officials convinced more researchers to join more committees than they had ever convinced to join committees before.

d. *University officials convinced more researchers to join more committees than they had ever convinced researchers to join committees before.

(107)a. *Max persuaded more men to buy more cars than you persuaded to buy.

b. *Max persuaded more men to buy more cars than you persuaded men to buy.

c. ?Max persuaded more men to buy more cars than you persuaded to buy cars.

d. Max persuaded more men to buy more cars than you persuaded men to buy cars.

The interesting examples are the (c) sentences, in which the higher of the two compared constituents is moved and deleted and the lower one remains *in situ*. While these sentences are somewhat degraded, presumably because of the redundancy in the comparative clause and their overall complexity, they contrast quite clearly with the other examples, which are completely unacceptable.

In fact, this pattern of data is exactly what the analysis predicts. Even though the (c) structures forego deletion of a phrase that could be eliminated without violating RECOVERABILITY, the other candidates are less optimal: the (a) candidate violates whatever constraint rules out multiple *wh*-movement, the (b) candidate violates Superiority, and the (d) candidate violates DELETE twice. Tableau VI provides a schematic illustration of how the competing structures are evaluated, where CC_1 and CC_2 are the compared constituents, and SUPERIORITY and *MULT-*wh* are abbreviations for whatever constraint or set of constraints is responsible for these restrictions. (I have left these constraints unordered with respect to each other for simplicity.)

Tableau VI. *Multiply-headed CD*

	SUPERIORITY	*MULT-*wh*	DELETE	STAY
a. than $[_{CP}$ ~~CC_1~~ ~~CC_2~~ … ~~CC_1~~ ~~CC_2~~		*!		**
b. than $[_{CP}$ ~~CC_2~~ … CC_1 ~~CC_2~~]	*!		*	*
c. ✓ than $[_{CP}$ ~~CC_1~~ … ~~CC_1~~ CC_2]			*	*
d. than $[_{CP}$ … CC_1 CC_2]			**!	

What is particularly interesting about these facts is that they clearly illustrate both constraint violation and optimization. Even though the (c) candidate fails to (move and) delete a compared constituent under identity with the corresponding head, the structure is optimal because the only way to fully satisfy DELETE – movement – would incur a violation of some higher-ranked constraint. At the same time, (107c) and (106c) demonstrate that a simple rule for comparatives of the form '(move and) delete a compared constituent when it is identical to the head' would be too strong. This is a problem for any analysis of comparatives stated in terms of absolute constraints, regardless of whether it is formulated in terms of feature strength (an approach we have already seen to be stipulative) or unbounded deletion of identical material (as in Bresnan 1975). This pattern of data is an expected result in a model in which well-formedness is determined by an optimality metric, however.

5.1.2. *Hidden Subdeletion*

A second context in which we find identity without movement involves the interaction of comparatives and ellipsis. Recall that the definition of DELETE was formulated to be as general as possible: it rules out any overt expressions. If this is the correct characterization of this constraint, then the constraint should not care how deletion is licensed: the result is what is important, not the means that achieve the result. In section 4.2.1, I observed that there are at least two ways to license deletion of phrasal material in English: movement and ellipsis. I also claimed that in examples of CD like

(108), movement is the only means of deleting the compared constituent, because *in situ* deletion of an argument violates RECOVERABILITY (see section 4.3). Thus, even though the representations in (109a) and (109b) equally satisfy DELETE, and (109b) better satisfies STAY, (109a) must be the actual structure of (108).

(108) Dennis has more tattoos than Michael has.

(109)a. Dennis has more tattoos than [$_{CP}$ ~~[$_{DP}$ tattoos]~~ Michael has ~~[$_{DP}$ tattoos]~~]

b. *Dennis has more tattoos than [$_{CP}$ Michael has ~~[$_{DP}$ tattoos]~~]

This reasoning, though, predicts that if the compared constituent were part of a larger phrase that *could* be legally deleted, then it should remain in its base position. In other words, the ranking of DELETE over STAY doesn't simply favor syntactic representations in which movement occurs in order to license deletion, it also entails that syntactic representations in which deletion can be licensed *without* movement are the best of all.

To check this prediction, we must look at cases of comparative ellipsis: CD constructions in which additional phrasal material is targeted by ellipsis. Consider, for example, the sentence in (110), which is just like (108) except that the embedded VP is the target of VP-deletion.

(110) Dennis has more tattoos than Michael does.

The two crucial candidate representations of this sentence are one in which the compared constituent moves and deletes and one in which it remains *in situ*, as shown in Tableau VII.

Tableau VII. *Hidden Subdeletion*

	DELETE	STAY
a. than [$_{CP}$ ~~[$_{DP}$ tattoos]~~ Michael does ~~[$_{VP}$ have ~~[$_{DP}$ tattoos]~~]~~]		*
b. √ than [$_{CP}$ Michael does ~~[$_{VP}$ have ~~[$_{DP}$ tattoos]~~]~~]		

Although the candidates correspond to identical surface strings, and so equally satisfy DELETE, they differ with respect to the lower ranked constraint STAY. In particular, the compared constituent in (b) is contained in the elided VP, and so is eliminated without movement. As a result, this candidate is predicted to be optimal.

The prediction of the analysis, then, is that examples of comparative ellipsis should have what I will call a HIDDEN SUBDELETION structure:

one in which the compared constituent remains *in situ*, even though it is identical to the head. Clearly, if evidence for hidden subdeletion can be identified, it would provide a powerful tool for distinguishing between different analyses of comparatives. A standard approach using hard constraints should predict either that overt movement (or the equivalent) occurs in CD regardless of whether or not the compared constituent is contained in an elided phrase (a prediction we saw to be incorrect in the previous section), or at least that movement should always be an option. In contrast, the Optimality Theoretic analysis that I have advocated here clearly predicts that the hidden subdeletion structure is the only possible representation in these contexts, since the alternative structure with movement is less optimal.

Two contexts provide evidence for hidden subdeletion. The first involves multiply-headed comparatives. While multiply-headed CD is generally impossible (*modulo* the observations in the previous section), this constraint is not absolute. In particular, multiple CD is possible 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). This is illustrated by the contrasts in (111) and (112).

(111)a. *Max persuaded more people to buy more cars than you persuaded to buy.

b. Max persuaded more people to buy more cars than you did.

(112)a. *Christmas doesn't make as many children as happy as birthdays make.

b. Christmas doesn't make as many children as happy as birthdays do.

These facts follow if VP-deletion forces a hidden subdeletion structure in CD. If deletion of the VPs in (111b) and (112b) forces the compared constituent to remain *in situ*, as argued above, then the PF representations assigned to these examples are as shown in (113).

(113)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]]~~

Crucially, neither structure involves multiple instances of overt Ā-movement, so they are correctly predicted to be well-formed.

A second piece of evidence for hidden subdeletion comes from the interaction of VP-deletion and parasitic gaps in comparatives. As shown by the contrasts in (114) and (115), VP-deletion and other types of ellipsis can bleed otherwise acceptable parasitic gaps in comparatives. (Kennedy and Merchant (2000a) observe, but do not explain, similar facts in attributive comparatives.)

(114)a. Mo interviewed more suspects than Art interviewed without arresting *e*.

b. *Mo interviewed more suspects than Art did without arresting *e*.

(115)a. I actually liked more of the films that came out this year than I expected to enjoy before seeing *e*.

b. *I actually liked more of the films that came out this year than I expected to before seeing *e*.

Note that this is not a property of VP-deletion in general: Ā-movement out of a deleted VP in relative clauses and questions can license a parasitic gap, as shown by the examples in (116).

(116)a. The books that Hillary threw away after reading *e* are the same ones that Max did before reading *e*.

b. Hillary bought the same car that I did after seeing advertised *e* on TV.

Again, the contrasts in (114) and (115) follow if VP-deletion forces the hidden subdeletion analyses of these comparatives shown in (117). Since the compared constituents do not move, they do not license parasitic gaps.

(117)a. *Mo interviewed more suspects than Art did ~~[$_{VP}$ interview [$_{DP}$ 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.

The parasitic gap facts are particularly important, because they demonstrate that a hidden subdeletion analysis in instances of comparative ellipsis

is not just an option, it is in fact the only possible analysis. If a hidden subdeletion structure were possible but not required, then (114b) and (115b) would have alternative parses in which the compared constituent moves and deletes, licensing the corresponding parasitic gaps. 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 follows directly from an Optimality Theoretic analysis, but it would have to be stipulated in frameworks that do not incorporate some kind of optimality metric.

5.2. *Cross-Linguistic Variation in Comparatives*

Within Optimality Theory, differences between languages are accounted for in terms of constraint re-rankings. Indeed, showing that typological differences can be straightforwardly explained in these terms is one of the most important types of argument for the approach in general. In the case of comparatives, however, it is difficult to use this metric, as the syntactic means of expressing comparison in the world's languages vary quite dramatically. (See Stassen (1985) for a comprehensive survey of comparative constructions in the languages of the world.) In particular, many languages do not appear to use $\bar{A}$-movement constructions to express comparison. Since the constraints I have posited for English are constraints governing the expression and derivation of $\bar{A}$-chains, the principles that I have proposed here will not necessarily apply to the analysis of comparatives in other languages. It follows that a full exploration of the cross-linguistic predictions of the analysis first requires a comprehensive syntactic analysis of the structural aspects of comparison in a variety of languages, a task that is beyond the scope of this paper. I will therefore limit myself to a discussion of what sort of variation we should expect to find, plus a brief examination of a couple of cases that appear to fit in with the expectations of the analysis.

The most obvious question that needs to be answered is whether the constraint re-ranking STAY $\gg$ DELETE is manifested in some language. All other things being equal, this type of constraint ordering would result in a language in which the compared constituent is either not moved and not deleted, or else not moved but deleted in accord with some other principle operative in the language. For example, if a language allowed DP ellipsis, we would expect nominal comparatives to look just like English superficially, but to not involve overt movement. (Such a language would therefore (in principle) allow multiple CD and not license parasitic gaps in comparatives.) Alternatively, if DP ellipsis were not an option, we would expect only covert movement in comparatives, and some (possibly

reduced) expression of the compared constituent in CD. Whether these options are manifested remains to be seen.

There are other possible constraint re-rankings to consider which do appear to be manifested cross-linguistically. For example, if STAY were ranked below C-SCOPE at PF, the result would be overt movement in sub-deletion. As reported in Rivero (1981), Knowles (1984), and Price (1990), Castilian Spanish is just such a language. This is illustrated by the following examples from Price (1990, p. 43) (see Borsley 1984 for similar facts in Polish equatives).[26]

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

b.*Mi padre vende más libros que mi madre compra discos.
my father sells more books than my mother buys records

(119)a. La mesa es más large que ancha es la puerta.
the table is more long than wide is the door
The table is longer than the door is wide.

b.*La mesa es más large que la puerta es ancha.
the table is more long than the door is wide

[26] Rivero (1981, p. 192) presents the following examples as evidence that this is indeed Ā-movement. (i) shows that CSD in Spanish can operate over a long distance, and (ii) shows that it is sensitive to the Complex NP Constraint.

(i) Aqui hay tantos libros como revistas espero que prometas
Here there-are as-many books as journals I-hope that you-will-promise
que comprarás tú.
that you-will-buy you
Here there are as many books as I hope you will promise to buy journals.

(ii)* Aqui hay tantos libros como revistas espero tu promesa de que
Here there-are as-many books as journals I-hope your promise of that
comprarás tú.
you-will-buy you
Here there are as many books as I am waiting for to your promise to buy journals.

(120)a. El crío gatea más cuidadosamente que descuidadamente
the baby crawls more carefully than carelessly
anda su hermana.
walks his sister
The baby crawls more carefully than his sister walks carelessly.

b.*El crío gatea más cuidadosamente que su hermana anda
the baby crawls more carefully than his sister walks
descuidadamente.
carelessly

A similar type of case would be one in which the relative ranking of DELETE and STAY is as in English, but STAY is ranked above the constraints that force overt movement of *wh*-XPs (see Ackema and Neeleman 1998 for an OT analysis of the typology of wh-movement); this would give us a *wh-in situ* language with overt movement in comparatives. Japanese appears to be such a language, since according to Kikuchi (1989) (see also Ishii 1991), CD in Japanese has the canonical properties of overt Ā-movement: it requires a gap, it may occur across a bridge verb, it is sensitive to islands, it shows crossover effects, and it licenses parasitic gaps. This is illustrated by the following set of data from Kikuchi (1989).[27]

(121) Tom-wa John-ga *t*/*sore/*sorera/hon-o yonda yorimo
Tom-TOP John-NOM t/it/them/book-ACC read than
hon-o takusan yonda.
book-ACC many read
Tom read more books than John read *t*/*it/*them/*the books.

[27] A caveat: Ayumi Ueyama (personal communication) informs me that the facts reported by Kikuchi are not fully representative of the class of comparative constructions in Japanese, and that it is unclear whether Kikuchi's generalizations hold across the board. It is possible that Japanese (like English, for that matter) has different means of expressing the same concept (cf. *Kim is taller than Lee* and *Kim's height exceeds Lee's height*), which involve different functional vocabularies and are therefore subject to different constraints.

(122) John-ga *t* yonda to iwarete iru to minna-ga omotte
John-NOM t read C is-said ASP C everyone-NOM think
iru yorimo Mary-wa takusan hon-o yonde ita
ASP than Mary-TOP many book-ACC read ASP
Mary has read more books than everyone thinks that it is said that John read.

(123) * John-ga *t* yonde ita tokini zisin-ga okita
John-nom t read ASP when earthquake-NOM happened
yorimo Paul-wa harukani takusanno hon-o yonde ita
than Paul-TOP far many book-ACC read ASP
Paul has read more books than an earthquake happened when John was reading.

(124) * Zibun-tati-gai rakudaisita koto-ga *t* odorokasita yorimo
self-PL-NOM flunked fact-NOM t surprised than
harukani takusanno gakusei-o Bill-ga rakudaisita
far many students-ACC Bill-NOM flunked
koto-ga odorokasita
fact-NOM surprised
The fact that Bill flunked surprised far more students than the fact that they$_i$ flunked surprised t_i.

(125) Ronbun-nituite ieba Bill-wa John-gai London-de e_i kaita
article-about say Bill-TOP John-NOM London-at e_i wrote
ato Paris-de t_i happyousita yorimo ookuno ronbun-o
after Paris-at t_i published than many article-ACC
America-de kaite ita
America-at write ASP
As for the articles, Bill wrote more articles in America than John had published t_i in Paris after he wrote e_i in London.

6. THE HEAD RAISING ANALYSIS

Perhaps the strongest challenge to the analysis of comparatives that I have developed here comes from recent work by Lechner (1999), which analyzes CD as movement of the compared constituent from its base position in the comparative clause into the head position (see also Rivero 1981 and Kayne 1994, and the analyses of relative clauses in Schachter 1973; Vergnaud 1974; and Carlson 1977). This type of approach, traditionally referred to as the RAISING ANALYSIS, assigns typical instances of comparative deletion the structures in (126).

(126)a. Michael has [$_{DP}$ more [$_{NP}$ scoring titles] than [$_{CP}$ Dennis has ~~[$_{NP}$ scoring titles]~~]]

b. Michael's feet are [$_{DegP}$ [$_{AP}$ wide] er than [$_{CP}$ my feet are ~~[$_{AP}$ wide]~~]]

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 (127).

(127)a. Michael has [$_{DP}$ more [$_{NP}$ scoring titles] than [$_{CP}$ Dennis has tattoos]]

b. Michael's feet are [$_{DegP}$ [$_{AP}$ wide] er than [$_{CP}$ my feet are long]]

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 3: CD would involve overt movement of the compared constituent, and CSD would involve covert movement of the compared constituent.

The crucial difference between a raising analysis and my proposal is that the former would not need to assume that syntactic constraints are ranked and violable. Instead, it would derive the derivational difference between CD and CSD from the plausible assumption that an empty head position in a comparative must be filled. In Lechner (1999), this hypothesis is implemented by postulating a feature on the head that is eliminated by movement of the compared constituent, with the result that overt movement is forced in CD by principles of feature checking.

There are at least three compelling arguments against this type of approach, two of which come from data we have already seen. (See Carlson

(1977) for additional arguments against a raising analysis of comparatives.) Recall from the previous section that multiply-headed CD is possible in an example like (128a) because VP-deletion forces the hidden subdeletion analysis in (128b), in which the compared constituents do not move overtly.

(128)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]]~~

If the head of the comparative were derived from an internal position, however, then the structure assigned to (128a) would presumably be (129), not (128b).

(129) Christmas doesn't make as many [$_{NP}$ children] as [$_{AP}$ happy] as birthdays do ~~[$_{VP}$ make [$_{NP}$ children] [$_{AP}$ happy]]~~

If this were the correct analysis, however, there would be no difference between the well-formed multiply-headed structures with VP-deletion and the unacceptable examples without ellipsis (such as (112a) above): both would involve the same movements of the same constituents. The analysis should therefore predict either that VP-deletion should not save multiple-CD, or that multiple-CD should always be possible. As we have seen, both predictions are wrong.

The interaction of parasitic gaps and VP-deletion in comparatives provides a similar argument against the raising analysis. In section 5.1.2, the unacceptability of (130) 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.

(130) *Mo interviewed more suspects than Art did without arresting e_i

The problem for the raising analysis is that if the overt head of the comparative were raised from a lower position, the structure assigned to (130) would be (131).

(131) Mo interviewed [$_{DP}$ more [$_{NP}$ suspects]$_i$ than Art did ~~[$_{VP}$ interview [$_{NP}$ suspects]$_i$]~~ NA without arresting e_i]

(130) should therefore have the same syntactic analysis as a comparable comparative that does not involve VP-deletion (such as (114a) above), with

the result that the analysis incorrectly predicts that a parasitic gap should be possible.

The only way to save the raising analysis from these two problems would be to make essentially the same claim that I made in section 5.1.2: deletion of a phrase that contains the compared constituent forces a hidden subdeletion analysis (i.e., it forces base-generation of the head, rather than raising). In my proposal, this result follows from principles of optimality, in particular, the emergence of the lower-ranked constraint STAY as the crucial factor deciding between representations that equally satisfy DELETE. It is unclear what this result would follow from in a raising analysis that does not make essentially the same set of assumptions about optimality.[28]

A third argument against a raising analysis comes from disjoint reference effects in the comparative clause. Lechner (1999) argues that the impossibility of coreference in (132a) provides an argument in favor of the raising approach, since this type of analysis, in conjunction with the copy theory of movement, assigns this sentence the structure in (132b), which violates Condition C.

(132)a. *Louise is prouder of Frank$_i$ than he$_i$ is.

b. Louise is [$_{\text{DegP}}$ [$_{\text{AP}}$ prouder of Frank$_i$] than he$_i$ is ~~[$_{\text{AP}}$ proud of Frank$_i$]~~]

If this is the correct analysis, however, the acceptability of (133a) comes as a surprise, since (133b) also violates Condition C:

(133)a. Louise is prouder of Frank$_i$ than he$_i$ thinks she is.

b. Louise is [$_{\text{DegP}}$ [$_{\text{AP}}$ prouder of Frank$_i$] than he$_i$ thinks she is ~~[$_{\text{AP}}$ proud of Frank$_i$]~~]

The contrast between (132a) and (133a) represents a serious challenge to the raising analysis, since the very same assumption that rules out the

[28] In section 5.1.2, I pointed out that VP-deletion in relative clauses does not bleed parasitic gaps, as shown by (i).

(i) Hillary bought the same car that I did after seeing advertised *e* on TV.

This difference between relative clauses and comparatives suggests that the raising analysis may in fact be correct for the former, as claimed by Carlson (1977) (for amount relatives specifically). See Kayne (1994), Grosu and Landman (1998), Sauerland (1998), Bhatt (2000), and Bianchi (2000) for recent arguments in favor of a raising analysis for relatives, but see also Borsley (1997) for an opposing view.

former – that the head originates inside the comparative clause – should also rule out the latter.

The contrast between these two examples can be accommodated within the analysis I have developed here, however, if we add the assumption that deletion of chain copies under identity with the head of the comparative does not require strict form identity, but only a weaker requirement of identity of reference, as argued independently for relative clauses in Sauerland (1998).[29] If this is correct, then (133a) could have a representation in which the compared constituent actually contains a pronoun that corefers with the occurrence of *Frank* in the head, as in (134), which is perfectly well-formed.

(134) Louise is prouder of Frank$_i$ than [$_{CP}$ ~~[$_{DegP}$ proud of him$_i$]~~ he$_i$ thinks she is ~~[$_{DegP}$ proud of him$_i$]~~]

The unacceptability of (132a) may then be analyzed as a violation of Condition B, rather than Condition C, since the structure assigned to it would be (135).

(135) *Louise is prouder of Frank$_i$ than [$_{CP}$ ~~[$_{DegP}$ proud of him$_i$]~~ he$_i$ is ~~[$_{DegP}$ proud of him$_i$]~~]

Crucially, it must not be the case that there is an alternative syntactic analysis of (132a) in which the compared constituent contains a reflexive pronoun, as in (136), since such a structure would not violate the principles of the Binding Theory.

(136) Louise is prouder of Frank$_i$ than [$_{CP}$ ~~[$_{DegP}$ proud of himself$_i$]~~ he$_i$ is ~~[$_{DegP}$ proud of himself$_i$]~~]

In fact, there are good reasons to believe that (136) is not a possible analysis of (132a). A number of researchers have argued that reflexive predicates are semantically distinct from their non-reflexive counterparts (see in particular Bach and Partee 1980 and Reinhart and Reuland 1993). If this is

[29] Sauerland argues that deletion of relative clause internal chain copies under identity with an external (nonderived) head requires only identity up to 'vehicle change', in Fiengo and May's (1994) terms. Safir (1999) argues for the same conclusion, but because he ends up adopting a raising analysis of relative clauses, he also extends the proposal to apply in the actual copying component of movement. Such an assumption drastically weakens the notion of 'copy', however, and should be met with some scepticism. Here I maintain the stronger position that copies created by movement are identical to each other in all respects.

correct, then the head and compared constituent in (136) are not identical, and therefore deletion of the chain in (136) violates RECOVERABILITY.

Regardless of how this issue is resolved, the crucial point to take away from this discussion is that since a raising analysis states that the head of the comparative is literally the same as the compared constituent, it makes the wrong predictions regarding examples like (133a). In contrast, the movement-plus-deletion analysis that I have advocated in this paper can appeal to independently motivated properties of deletion rules (such as conditions on identity) to explain the data.

A final, and more general, question that remains unexplained on a raising analysis is why deletion is obligatory in CD. If the head of the comparative can optionally be base generated, rather than derived – an assumption that is necessary to accommodate subdeletion structures – then it is unclear what rules out examples like (137).

(137) *Michael has more scoring titles than Dennis has scoring titles

To paraphrase Bresnan, it becomes accidental on this analysis that maximally recoverable constituents must be deleted.

7. 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 and explanatory adequacy not matched by earlier proposals. At the same time, the overt/covert movement distinction between CD and CSD was shown to follow from general assumptions about the relation between movement, identity and deletion in a model in which syntactic constraints are ranked and violable. While the larger implications of this proposal must be further explored, its overall success in accounting for the (apparently paradoxical) properties of comparatives and in uncovering and explaining new facts (such as the hidden subdeletion data) make it a strong argument for optimality in syntax.

I have formulated the analysis of comparatives in terms of a syntactic framework that posits two levels of syntactic representation, but this assumption is arguably not necessary. The two syntactic features that are most important to the analysis are the 'gap' property (whether an expression that has a semantic value has a phonological value) and the

'displacement' property (whether an expression is interpreted in the position in which it is pronounced). Using this terminology, the difference between CD and CSD in English can be restated as follows: both CD and CSD involve displacement, but only CD creates a gap. Since all generative syntactic frameworks have some mechanism(s) for handling displacement and gaps, it should be possible to transfer the core of the analysis itself to another framework. Whether such an alternative approach is viable will depend crucially on whether the differences between CD and CSD that I accounted for in terms of overt vs. covert movement – in particular, parasitic gap licensing and the distribution of multiply-headed comparatives – can be satisfactorily explained without reference to this distinction. This question is particularly relevant given the complex interaction of overt movement and ellipsis (the 'hidden subdeletion' data) discussed in section 5.1, which crucially showed that even comparatives that normally require a gap in the position of the compared constituent do not require one – in fact, forbid one – when this position is included in the target of ellipsis.

REFERENCES

Abney, Steven P. 1987. *The English Noun Phrase in Its Sentential Aspect*, unpublished Ph.D. dissertation, MIT, Cambridge, MA.

Ackema, Peter and Ad Neeleman. 1998. 'Optimal Questions', *Natural Language & Linguistic Theory* **16**, 443–490.

Anderson, Catherine. 2000. 'Why 'that' Happens: A PF Account of Variability in Complementizer Deletion', unpublished manuscript, Northwestern University.

Anderson, Stephen R. 1992. *A-Morphous Morphology*, Cambridge University Press, Cambridge.

Andrews, Avery. 1985. *Studies in the Syntax of Relative and Comparative Clauses*, Garland, New York.

Aoun, Joseph and Elabbas Benmamoun. 1998. 'Minimality, Reconstruction, and PF Movement', *Linguistic Inquiry* **29**, 569–597.

Bach, Emmon and Barbara Partee. 1980. 'Anaphora and Semantic Structure', in J. Kreiman and A. Ojeda (eds.), *Paper from the Parasession on Pronouns and Anaphora*, Chicago Linguistics Society, Chicago, IL, pp. 1–28.

Baker, C. Lee: 1970, 'Notes on the Description of English Questions', *Foundations of Language* **6**, 197–219.

Baltin, Mark. 1978. 'PP as a Bounding Node', in M. Stein (ed.), *Proceedings of NELS 8*, CSLA, Amherst, MA, pp. 33–40.

Baltin, Mark. 1982. 'A Landing Site Theory of Movement Rules', *Linguistic Inquiry* **13**, 1–38.

Barker, Chris. 1995. *Possessive Descriptions*, CSLI Publications, Stanford, CA (1991 Ph.D. dissertation, University of California, Santa Cruz).

Bartsch, Renate and Theo Vennemann. 1973. *Semantic Structures: A Study in the Relation between Syntax and Semantics*, Athäenum Verlag, Frankfurt.

Beard, Robert. 1995. *Lexeme-Morpheme Based Morphology*, SUNY Press, Albany.

Bhatt, Rajesh. 2000. 'Adjectival Modifiers and the Raising Analysis of Relative Clauses', in M. Hirotani (ed.), *Proceedings of NELS 30*, GSLA, Amherst, MA, pp. 159–173.
Bianchi, Valentina. 2000. 'The Raising Analysis of Relative Clauses: A Reply to Borsley', *Linguistic Inquiry* **31**, 123–140.
Bobaljik, Jonathan D. 1995. *Morphosyntax: The Syntax of Verbal Inflection*, unpublished Ph.D. dissertation, MIT, Cambridge, MA.
Bobaljik, Jonathan D. 1999. 'A-Chains at the Interface: Copies, Agreement and "Covert Movement" ', unpublished manuscript, McGill University.
Borsley, Robert D. 1997. 'Relative Clauses and the Theory of Phrase Structure', *Linguistic Inquiry* **28**, 619–647.
Bošković, Željko. 2000. 'What Is Special about Multiple Wh-Fronting?', in M. Hirotani, A. Coetzee, N. Hall, and J. Kim (eds.), *Proceedings of NELS 30*, GLSA, Rutgers University, Amherst, MA, pp. 83–108.
Bresnan, Joan. 1973. 'Syntax of the Comparative Clause Construction in English', *Linguistic Inquiry* **4**, 275–343.
Bresnan, Joan. 1977. 'Variables in the Theory of Transformations', in P. Culicover, T. Wasow, and A. Akmajian (eds.), *Formal Syntax*, Academic Press, New York, pp. 157–196.
Brody, Michael. 1995. *Lexico-Logical Form: A Radically Minimalist Theory*, MIT Press, Cambridge, MA.
Browning, Marguerite. 1987. 'Null Operator Constructions', unpublished Ph.D. dissertation, MIT, Cambridge, MA.
Carlson, Greg N. 1977. 'Amount Relatives', *Language* **58**, 520–542.
Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*, MIT Press, Cambridge, MA.
Chomsky, Noam. 1977. 'On Wh-Movement', in P. Culicover, T. Wasow, and A. Akajian (eds.), *Formal Syntax*, Academic Press, New York, pp. 71–132.
Chomsky, Noam. 1993. 'A Minimalist Program for Linguistic Theory', in K. Hale and J. Keyser (eds.), *The View from Building 20, Essays in Linguistics in Honor of Sylvain Bromberger*, MIT Press, Cambridge, MA, pp. 1–52.
Chomsky, Noam. 1995. *The Minimalist Program*, MIT Press, Cambridge, MA.
Chomsky, Noam and Howard Lasnik. 1977. 'Filters and Control', *Linguistic Inquiry* **8**, 425–504.
Cinque, Gugliemo. 1999. *Types of A′-Dependencies*, MIT Press, Cambridge, MA.
Cooper, Robin. 1983. *Quantification and Syntactic Theory*, Reidel, Dordrecht.
Corver, Norbert. 1990. *The Syntax of Left Branch Extractions*, unpublished Ph.D. dissertation, Tilburg University.
Corver, Norbert. 1993. 'A Note on Subcomparatives', *Linguistic Inquiry* **24**, 773–781.
Corver, Norbert. 1997. 'Much-Support as a Last Resort', *Linguistic Inquiry* **28**, 119–164.
Cresswell, M. J. 1977. 'The Semantics of Degee', in B. Partee (ed.), *Montague Grammar*, Academic Press, New York, pp. 261–292.
Cresti, Diana. 2000. 'Ellipsis and Reconstruction in Relative Clauses', in M. Hirotani, A. Coetzee, N. Hall, and J. Kim (eds.), *Proceedings of NELS 30*, GLSA, Rutgers University, Amherst, MA, pp. 153–164.
Donati, Caterina. 1998. 'Another Way to Build Wh-Structures', paper presented at the Workshop on Comparatives, Zentrum für Allgemeine Sprachwissenschaft, Berlin, November 27–28.
Engdahl, Elisabet. 1983. 'Parasitic Gaps', *Linguistics and Philosophy* **6**, 5–34.

Farkas, Donka and Anastasia Giannakidou. 1996. 'How Clause-Bounded is the Scope of Universals?', in T. Galloway and J. Spence (eds.), *Proceedings of SALT 6*, Ithaca, New York, pp. 35–52.

Fiengo, Robert and Howard Lasnik. 1972. 'On Nonrecoverable Deletion in Syntax', *Linguistic Inquiry* **3**, 528.

Fiengo, Robert and Robert May. 1094. *Indices and Identity*, MIT Press, Cambridge, MA.

Franks, Steven. 1998. 'Clitics in Slavic', paper presented at the Comparative Slavic Morphosyntax Workshop, Indiana University.

Gawron, Jean-Mark. 1995. 'Comparatives, Superlatives and Resolution', *Linguistics and Philosophy* **18**, 333–380.

Gazdar, Gerald. 1981. 'A Phrase Structure Syntax for Comparative Clauses', in T. Hoekstra, J. van der Hulst, and M. Morgan (eds.), *Lexical Grammar*, Blackwell, Oxford, pp. 165–179.

Grimshaw, Jane. 1987. 'Subdeletion', *Linguistic Inquiry* **18**, 659–669.

Grimshaw, Jane. 1991. 'Extended Projection', unpublished manuscript, Brandeis University.

Grimshaw, Jane. 1997. 'Projection, Heads and Optimality', *Linguistic Inquiry* **28**, 373–422.

Grosu, Alexander. 1994. *Three Studies in Locality and Case*, Routledge, London.

Grosu, Alexander and Fred Landman. 1998. 'Strange Relatives of the Third Kind', *Natural Language Semantics* **6**, 125–170.

Hackl, Martin. 2001. *Comparative Quantifiers*, unpublished Ph.D. dissertation, MIT, Cambridge, MA.

Halle, Morris and Alec Marantz. 1993. 'Distributed Morphology and the Pieces of Inflection', in K. Hale and S. J. Keyser (eds.), *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, MIT Press, Cambridge, MA, pp. 111–176.

Hankamer, Jorge. 1971. *Deletion in Coordinate Structures*, unpublished Ph.D. dissertation, Yale University.

Hankamer, Jorge. 1973. 'Why There are Two 'than's in English', in C. Corum, T. C. Smith-Stark, and A. Weiser (eds.), *Proceedings of the 9th Annual Meeting of the Chicago Linguistics Society*, Chicago, IL, pp. 179–191.

Hankamer, Jorge. 1979. *Deletion in Coordinate Structures*, Garland Press, New York (1971, Ph.D. dissertation, York University).

Hankamer, Jorge and Ivan Sag. 1976. 'Deep and Surface Anaphora', *Linguistic Inquiry* **7**, 391–428.

Hazout, Ilan. 1995. 'Comparative Ellipsis and Logical Form', *Natural Language & Linguistic Theory* **13**, 1–37.

Heim, Irene. 1985. 'Notes on Comparatives and Related Matters', unpublished manuscript, University of Texas, Austin.

Heim, Irene and Angelika Kratzer. 1998. *Semantics in Generative Grammar*, Blackwell, Oxford.

Hendriks, Petra. 1995. *Comparatives and Categorial Grammar*, unpublished Ph.D. dissertation, University 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*-t Violations', in C. Pye (ed.), *Proceedings of the 1996 Mid-America Linguistics Conference*, Department of Linguistics, University of Kansas, Lawrence, pp. 168–180.

Huang, C. T. James. 1982. *Logical Relations in Chinese and the Theory of Grammar*, unpublished Ph.D. dissertation, MIT.

Huddleston, Rodney. 1967. 'More on the English Comparative', *J. Linguistics* **3**, 91–102.

Ishii, Yasuo. 1991. *Operators and Empty Categories in Japanese*, unpublished Ph.D. dissertation, University of Connecticut, Connecticut.

Izvorski, Roumyana. 1995. 'A Solution to the Subcomparative Paradox', in J. Camacho, L. Chouieri, and M. Watanabe (eds.), *The Proceedings of WCCFL XIV*, Stanford, CA, pp. 203–219.

Jacobson, Pauline. 1995. 'On the Quantification Force of English Free Relatives', in E. Bach, E. Jelinek, A. Kratzer, and B. H. Partee (eds.), *Quantification in Natural Languages*, Kluwer, Dordrecht, pp. 451–486.

Kayne, Richard S. 1994. *The Antisymmetry of Syntax*, MIT Press, Cambridge, MA.

Kehler, Andreq and Gregory Ward. 1999. 'On the Semantics and Pragmatics of 'Identifier *So*', in K. Turner (ed.), *The Semantics/Pragmatics Interface from Different Points of View*, Vol. 1 of *Current Research in the Semantics/Pragmatics Interface*, Elsevier, New York, pp. 233–256.

Kennedy, Christopher. 1997. 'On the Quantificational Force of the Comparative Clause', in J. Austin and A. Lawson (eds.), *Proceedings of ESCOL '97*, Ithaca, New York, pp. 70–81.

Kennedy, Christopher. 1998. 'Local Dependencies in Comparative Deletion', in *The Proceedings of WCCFL XVII*, Stanford, pp. 375–389.

Kennedy, Christopher. 1999. *Projecting the Adjective: The Syntax and Semantics of Gradability and Comparison*, Garland, New York (1997, Ph.D. dissertation, University of California, Santa Cruz).

Kennedy, Christopher. 2001. Polar Opposition and the Ontology of 'Degrees', *Linguistics and Philosophy* **24**, 33–70.

Kennedy, Christopher. In preparation. 'Comparative and Relative Clauses Compared', unpublished manuscript, Northwestern University.

Kennedy, Christopher. To appear. 'Ellipsis and Syntactic Representation', in K. Schwabe and S. Winkler (eds.). *Interpreting (Omitted) Structure*, John Benjamins, Amsterdam.

Kennedy, Christopher and Jason Merchant. 2000a. 'Attributive Comparative Deletion', *Natural Language and Linguistic Theory* **18**, 89–146.

Kennedy, Christopher and Jason Merchant. 2000b. 'The Case of the "Missing CP" and the Secret Case', Jorge Hankamer Webfest (http://ling.ucsc.edu/Jorge/kennedy_merchant.html).

Kikuchi, Akira. 1989. 'Comparative Deletion in Japanese', unpublished manuscript, Yamagata University.

Klein, Ewan. 1980. 'A Semantics for Positive and Comparative Adjectives', *Linguistics and Philosophy* **4**, 1–45.

Knowles, John. 1984. 'Structurai Choice in Spanish Clausal Comparison', *Canadian J. Linguistics* **29**, 1–19.

Kuno, Susumo. 1973. 'Constraints on Internal Clauses and Sentential Subjects', *Linguistic Inquiry* **4**, 363–385.

Kuno, Susumo and Jane Robinson. 1972. 'Multiple Wh Questions', *Linguistic Inquiry* **3**, 463–487.

Lechner, Winfried: 1999. *Comparatives and DP-Structure*, unpublished Ph.D. dissertation, University of Massachusetts, Amherst, MA.

Lees, Robert B. 1961. 'Grammatical Analysis of the English Comparative Construction', *Word* **17**, 171–185.

Lerner, Jean-Yves and Manfred Pinkal. 1992. 'Comparatives and Nested Quantification', Technical report, Universität des Saarlandes, CLAUS Report No. 21.

Lidz, Jefrey. 1995. 'Valency in Kannada: Evidence for Interpretive Morphology', in A. Dimitriedis, H. Lee, C. Moisset, and A. Williams (eds.), *Penn Working Papers in Linguistics*, Vol. 5, Penn Linguistics Club, Philadelphia, pp. 37–64.

Lindblom, Bjorn. 1990. 'Explaining Phonetic Variation: A Sketch of the H & H Theory', in W. Hardcastle and A. Marchal (eds.), *Speech Production and Speech Modelling*, Kluwer, Dordrecht, pp. 403–439.

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

Merchant, Jason. 2001. *The Syntax of Silence: Sluicing, Islands, and the Theory of Ellipsis*, Oxford University Press, Oxford.

Munn, Alan. 2001. 'Explaining Parasitic Gap Restrictions', in P. Culicover and P. Postal (eds.), *Parasitic Gaps*, MIT Press, Cambridge, MA, pp. 369–392.

Napoli, Donna Jo. 1983. 'Comparative Ellipsis: A Phrase Structure Analysis', *Linguistic Inquiry* **14**, 675–694.

Nissenbaum, Jon. 1998. 'Derived Predicates and the Interpretation of Parasitic Gaps', in K. Shahin, S. Blake, and E.-S. Kim (eds.), *Proceedings of WCCFL 17*, Stanford, CA, pp. 507–521.

Nissenbaum, Jon. 2000. 'Covert Movement and Parasitic Gaps', in M. Hirotani, A. Coetzee, N. Hall, and J. Kim (eds.), *Proceedings of NELS 30*, GLSA, Rutgers University, Amherst, MA, pp. 541–556.

Pesetsky, David. 1997. 'Optimality Theory: Movement and Pronunciation', in D. Archangeli and D. T. Langendoen (eds.), *Optimality Theory: An Overview*, Basil Blackwell, Oxford, pp. 134–170.

Pinkham, Jesse. 1982. *The Formation of Comparative Clauses in French and English*, unpublished Ph.D. dissertation, Indiana University, Indiana.

Pollard, Carl and Ivan Sag. 1994. *Head-Driven Phrase Structure Grammar*, University of Chicago, Press, Chicago.

Postal, Paul. 1993. 'Parasitic Gaps and across the Board Phenomena', *Linguistic Inquiry* **24**, 735–754.

Postal, Paul. 1998. *Three Investigations of Extraction*, MIT Press, Cambridge, MA.

Price, Susan. 1990. *Comparative Constructions in Spanish and French Syntax*, Blackwell, New York.

Price, Alan and Paul Smolensky. 1993. 'Optimality Theory: Constraint Interaction in Generative Grammar', Technical report, Rutgers University, RuCCS Technical Report # 1, Rutgers Center for Cognitive Science.

Reinhart, Tanya and Eric Reuland. 1993. Reflexivity', *Linguistic Inquiry* **24**(4), 657–720.

Rivero, Maria-Luisa. 1981. 'Wh-Movement in Comparatives in Spanish', in W. Cressy and D. J. Napoli (eds.), *Linguistic Symposium on Romance Languages 9*, Georgetown University Press, Washington, D.C., pp. 177–196.

Ross, John R. 1967. *Constraints on Variables in Syntax*, Ph.D. dissertation, MIT, Cambridge, MA; published as *Infinite Syntax*, Ablex Publishing, Norwood, NJ (1986).

Rudin, Catherine. 1984. 'Comparatives and Equatives in Bulgarian and the Balkan Languages', in K. Shangriladze and E. Townshend (eds.), *Papers from the V. congress of Southeast European Studies*, Columbus, OH, pp. 328–337.

Rullmann, Hotze. 1995. *Maximality in the Semantics of Wh-Constructions*, unpublished Ph.D. dissertation, University of Massachusetts, Amherst, MA.

Runner, Jeffrey. 1995. *Noun Phrase Licensing and Interpretation*, unpublished Ph.D. dissertation, University of Massachusetts, Amherst, MA.
Russell, Bertrand. 1905. 'On Denoting', *Mind (New Series)* **14**, 479–493.
Safir, Ken. 1999. 'Vehicle Change and Reconstruction in A-Bar Chains', *Linguistic Inquiry* **30**, 587–620.
Sag, Ivan. 1976. *Deletion and Logical Form*, unpublished Ph.D. dissertation, MIT, Cambridge, MA.
Sauerland, Uli. 1998. *The Meaning of Chains*, unpublished Ph.D. dissertation, MIT, Cambridge, MA.
Schachter, Paul. 1973. 'Focus and Relativization', *Language* **49**, 19–46.
Schwarzschild, Roger. 1999. 'GIVENness, AVOIDF and Other Constraints on the Placement of Accents', *Natural Language Semantics* **7**, 141–177.
Stassen, Leon. 1985. *Comparison and Universal Grammar*, Basil Blackwell, Oxford.
Taraldsen, Knut Tarald. 1978. 'The Scope of Wh-Movement in Norwegian', *Linguistic Inquiry* **9**, 623-640.
Veenstra, Mettina. 1998. *Formalizing the Minimalist Program*, unpublished Ph.D. dissertation, Rijksuniversitet Groningen.
Vergnaud, Jean Roger. 1974. *French Relative Clauses*, unpublished Ph.D. dissertation, MIT, Cambridge, MA.
von Stechow, Arnim. 1984. 'Comparing Semantic Theories of Comparison', *J. Semantics* **3**, 1–77.
Williams, Edwin. 1977. 'Discourse and Logical Form', *Linguistic Inquiry* **8**, 101–139.
Zwart, Jan Wouter. 1997. *Morphosyntax of Verb Movement*, Kluwer, Dordrecht.

Received 26 July 2000
Revised 25 October 2001

Department of Linguistics
Northwestern University
Evanston, IL 60208
USA
<kennedy@northwestern.edu>