

The View from the Periphery: The English Comparative Correlative

Peter W. Culicover
Ray Jackendoff

The English comparative correlative construction (e.g., *The more you eat, the fatter you get*) embeds like an ordinary CP, and each of its clauses displays an ordinary long-distance dependency. However, the connection between the two clauses is not ordinary: they are connected paratactically in syntax, but the first clause is interpreted as if it were a subordinate clause. The construction's mixture of the general and the idiosyncratic at all levels of detail challenges the distinction between "core" and "periphery" in grammar and the assumption that some level of underlying syntax directly mirrors semantic structure.

Keywords: syntax-semantics interface, comparative construction, long-distance dependency, core grammar, English syntax

In memory of Jim McCawley

1 Introduction: Two Dogmas of Generative Grammar

1.1 *Shun the Periphery*

For more than thirty years it has been widely accepted within generative grammar that the only syntactic phenomena truly worth studying are those that are maximally general—those that require minimal language-specific detail in their formulation. The motivation for this view is that these general phenomena are likely to reflect something fundamental and universal about the human capacity for language. By idealizing away from language-specific idiosyncrasies and by formulating maximally general rules, constraints, or relations on structures, generative grammar is seen to provide some preliminary insights into the nature of the language faculty. This perspective on the goals of syntactic theory appears to have reached its zenith in Chomsky's Minimalist Program

As this article was going to press, we learned of the untimely death of our valued friend and colleague Jim McCawley. Of all the generative semanticists of the late 1960s, Jim was the one who took interpretive semantics most seriously, inviting Jackendoff to speak at the University of Chicago as early as 1968 and serving as a friendly kibitzer in a number of Jackendoff's courses at Linguistic Institutes. Over the years he was a continual source of interesting comments on our work, delightful papers of his own on myriad topics, and obscure information on obscure composers. The present article draws heavily on his work; and he offered many helpful suggestions when Jackendoff presented a version of this work at the University of Chicago in 1997. It is therefore altogether appropriate to dedicate this article to his memory.

We are also grateful to two *LI* reviewers for useful comments on earlier versions of this article. Naturally any errors are our responsibility.

(Chomsky 1995), in which such syntactic phenomena as English extraposition and heavy NP shift are hypothesized to fall outside the scope of syntactic theory per se, because they cannot be described using the restrictive minimalist machinery that Chomsky takes to constitute syntax.

We find this construal of “syntax” overly narrow. We do recognize that truly general syntactic phenomena have a special status in virtue of their generality and possible universality. However, in our view an empirically adequate syntactic theory should be able to account for the full range of phenomena that actually exist in natural language. Universal Grammar (UG) must allow not only for the necessary and ubiquitous, but also for the possible and more limited. (Similar remarks are made by McCawley (1988a) and by Kay and Fillmore (1999).)

This broader approach, we think, comports better with generative grammar’s overall goal of accounting for how children can acquire language. Although it may be that certain aspects of language are universal and hence wired directly into the language faculty, idiosyncratic aspects are after all acquired too. Hence, there must be a basis in the language faculty not only for the most general and universal properties of language, but also for the most idiosyncratic and specific.

1.2 Mirror Semantics in Covert Syntax

A second theme in generative grammar over the past thirty years has been to seek some nonsurface (or covert) level of syntactic structure that directly encodes structural aspects of meaning. One interpretation of the Standard Theory (Chomsky 1965) was that Deep Structure was such a level; this interpretation, embodying the strongest version of the Katz-Postal hypothesis (Katz and Postal 1964), led directly to the syntactic exuberance of Generative Semantics (see Jackendoff 1972 for an elaboration of this argument). Chomsky’s reaction was the Extended Standard Theory (Chomsky 1972), which emphasized the independence of semantic structure from syntactic structure. However, this trend was reversed in the middle 1970s, when the level of Logical Form (LF) was introduced into generative theory. At first this level was intended to encode only such semantic notions as quantifier scope (May 1985), but later on in Government-Binding (GB) Theory we find much broader statements such as “PF and LF constitute the ‘interface’ between language and other cognitive systems, yielding direct representations of sound on the one hand and meaning on the other . . .” (Chomsky 1986:68).

Again there is an alternative perspective, one in which syntactic structure has its own autonomous properties and in which the syntactic structure of a sentence corresponds only partially to its semantic structure. Under this perspective it becomes a pervasive issue in linguistic theory to determine which aspects of a grammatical phenomenon are due to its syntax and which to its semantics. Crucial evidence for deciding this issue emerges when one can motivate mismatches between syntactic and semantic structure—situations in which the two levels independently capture generalizations that within a single level would necessarily conflict. In a number of recent works (Jackendoff 1992, 1997, Culicover and Jackendoff 1995, 1997) we have explored these issues of syntactic versus semantic responsibility, particularly with respect to binding, quantification, and extraction, using as our vehicles constructions somewhat off the beaten track—constructions that display considerable idiosyncrasy and are not generally considered part of “core gram-

mar.’’ In particular, all of the phenomena studied there display significant mismatches between syntactic and semantic embedding.

Both these themes—the importance of ‘‘peripheral’’ phenomena and the possibility of mismatch between semantics and syntax—are illustrated by the construction of English with which we will be concerned here. McCawley (1988b, 1998) calls this construction, exemplified in (1a), the *comparative conditional* (CC). It has a meaning roughly paraphrased by (1b).

- (1) a. The more you eat, the less you want.
- b. If/When/As you eat more, you want correspondingly less.

We find the paraphrase with *as* the closest. In particular (as pointed out by an *LI* reader), *if*-clauses but not *as*-clauses may be counterfactual (2a) or contain superlatives (2b). The CC construction follows the *as*-clauses in this respect (2c–d).

- (2) a. If/*As you had eaten more, you would want less.
- b. If/*As you eat the most, you want the least.
- c. *The more you would want, the less you would eat.
- d. *The most you want, the least you eat.

Thus, a better name for the construction than McCawley’s would be the *comparative correlative* (still CC).

This construction was noted first in the generative literature by Ross (1967:sec. 6.1.2.6). It is treated briefly by Fillmore, Kay, and O’Connor (1988), who argue that it is idiosyncratic and meaning-bearing. McCawley (1988b) expands their argument, showing how the construction inherits properties from both the comparative construction and the conditional. McCawley also examines parallel constructions in German and Mandarin, showing that they have similar properties relative to their host languages. Beck (1997) proposes a formal semantics-style account of the construction’s meaning, concentrating on German and English; she also cites examples from Dutch, French, Maltese, and Korean.¹ In investigating the syntax of the CC construction in more detail, we will essentially confirm and refine the conclusions of Fillmore, Kay, and O’Connor and of McCawley, showing how they bear on our more global concerns.

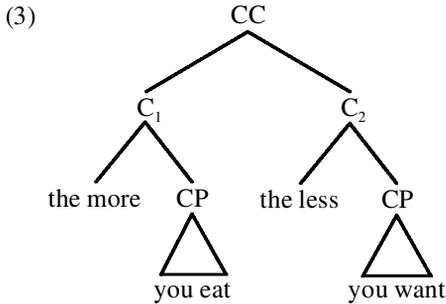
Sections 2 through 4 deal with the relation between the two clauses *the more you eat* and *the less you want*. Section 5 deals with the internal structure of the clauses. Section 6 investigates binding and quantification relations between the two clauses, leading to further reflections on the syntactic and semantic relations between them. Section 7 looks at extraction from the entire construction, a phenomenon that shouldn’t exist, but does. Section 8 sums up and offers final reflections.

¹ Beck bases her analysis on parallels with *if*-conditionals but does not mention *as*-correlatives. We suspect that her analysis could be easily extended to the latter.

2 The Basic Configuration of the Comparative Correlative Construction

We call the two parts of the CC construction the *clauses* of the construction: the first clause (*the more you eat* in (1a)) is C_1 and the second (*the less you want*) is C_2 .

For the moment, let us call the node dominating the entire construction simply CC. Then we get an overall structure along the lines of (3).



In (3) CP dominates *you eat* and *you want*. This is motivated by the fact that the complementizer *that* can occur in either clause, perhaps with a sense of a more colloquial register.²

- (4) a. The more (that) you eat, the less (that) you want.
 b. The angrier (that) Sue gets, the more (that) Fred admires her.

Now it might be that CC is a completely idiomatic category. However, it can be embedded in a complement construction, suggesting that actually $CC = IP$.

- (5) a. I think that the more you eat, the less you want.
 b. I'm not shocked by the idea that the more you eat, the less you want.
 c. It is obvious that the more you eat, the less you want.
 d. It is not entirely clear if/whether the more you eat, the less you want.
 e. I want to explain exactly why the more you eat, the less you want.

Another option is that in (3) $CC = CP$ and that therefore the use of the category CC in (5) is an instance of CP recursion. Still another possibility is that the CC construction is semantically propositional, whatever its syntactic category, and by virtue of this alone it can appear as the complement of CP. In any event we are faced with the uncomfortable choice of either inventing an entirely new syntactic category just for the CC construction, as in (3), or else permitting an unconventional expansion of IP or CP into two CPs. We will tentatively call CC an IP or CP and explore some possible ways to alleviate its apparent anomaly.

If CC is IP or CP, what is its internal structure? We will be exploring two main possibilities:

²This fact is also noted by Fillmore, Kay, and O'Connor (1988:508).

Hypothesis A (Paratactic hypothesis)

C_1 and C_2 are paratactic clauses forming an IP or CP.

(6) [_{IP/CP}[C_1 the more you eat] [C_2 the less you want]]

This is the “what you see is what you get” hypothesis: C_1 and C_2 are main clauses jammed together paratactically in an idiosyncratic configuration, anomalous from the point of view of X-bar theory.

Hypothesis B (Left-subordinate hypothesis)

C_2 is a main clause, C_1 a left adjunct.

(7) [_{IP/CP}[C_1 the more you eat] [_{IP/} C_2 the less you want]]

This gives the construction a somewhat more canonical structure, with a main clause and an adjunct. An immediate problem is the node labeled *IP/C*₂, which does not have the Spec-*I'* structure of an IP. We could instead take C_2 to be a CP. C_2 would then be an instance of CP recursion, as would C_1 . Hence, the structure would reduce to a version of the paratactic hypothesis. So this analysis is not without its own anomalies from the outset.

(Interestingly, Beck (1997) calls C_1 a subordinate clause, presumably because of its semantics, which she treats along the lines of *if*-clauses. However, when she draws a syntactic tree, it is symmetrical, like (6); she does not comment on the apparent mismatch—which is the point of the present investigation.)

Three other logical possibilities might occur to the systematic investigator:

Hypothesis C

C_1 and C_2 are conjoined by an empty conjunction.

(8) [_{IP}[_{IP/} C_1 the more you eat] Conj [_{IP/} C_2 the less you want]]

This is a potential refinement of the paratactic hypothesis. It eliminates the apparent anomaly of paratactic clauses at the cost of introducing a necessarily empty conjunction. However, it still does not solve the node-labeling problem of the left-subordinate hypothesis: *IP/C*₁ and *IP/C*₂ are both IPs that directly dominate CPs. As far as we can tell, then, this is at best a notational variant of the paratactic hypothesis and can be disregarded.³

³ In Italian, however, it is possible to have *e* ‘and’ between C_1 and C_2 .

(i) Più si mangia (e) più si ingrassa.
 more si eats (and) more si gets-fat
 ‘The more you eat, the fatter you get.’

This does not, in our view, constitute evidence about the correct analysis for the English construction, which differs from the Italian one in a number of respects. The most obvious difference is that English uses the idiosyncratic *the more* whereas Italian uses the adverbial *più* ‘more’ in clause-initial position.

Hypothesis D

C_1 is a main clause, C_2 a right adjunct.

- (9) [_{IP/CP}[_{IP/C₁} the more you eat] [_{C₂} the less you want]]

This is the mirror image of the left-subordinate hypothesis and shares all its virtues and vices so far.

Hypothesis E

C_1 and C_2 are arguments of an empty main verb.

- (10) [_{IP}[_{C₁} the more you eat] I [_{VP} V [_{C₂} the less you want]]]

In this hypothesis CC is headed by an empty I and an empty V; C_1 is the specifier and C_2 the complement of the empty V; C_1 and C_2 are both subordinate clauses. This has the virtue of giving CC a relatively normal IP structure. The price is an empty verb and empty I whose only purpose is to make this normal structure possible.

3 Evidence for the Left-Subordinate Hypothesis

We could easily rule out hypothesis E just on methodological grounds, because of its unmotivated verb and I. However, there is also simple empirical evidence against both it and hypothesis D. This evidence shows that if anything is the main clause of CCs, it has to be C_2 , as proposed in the left-subordinate hypothesis.

First, when CC is the complement of a verb or adjective that governs the subjunctive, subjunctive morphology may appear on C_2 (11a). (Since in our dialect the subjunctive is only optional in these contexts, the indicative is also possible.) The subjunctive may not appear on C_1 (11b).

- (11) {It is imperative that
{I demand that }
a. the more John eats, the more he pay(s).
b. *the more John eat, the more he pay(s).

It is also possible to form a tag question based on C_2 (12a), but not one based on C_1 (12b–c).

- (12) a. The more we eat, the angrier you get, don't you.
b. *The more we eat, the angrier you get, don't we.
c. *The more we eat, don't we, the angrier you get.

Hypothesis E incorrectly predicts that the subjunctive should appear on the empty verb, which is the main verb of the construction, and that tags should echo the empty I and contain a pronoun counterpart of the whole clause C_1 . Hypothesis D incorrectly predicts that the subjunctive should appear on C_1 and that tags too should be based on C_1 . Thus, these data rule out both hypotheses D and E: they appear to show that C_2 is a main clause and C_1 is a subordinate clause.

The paraphrase relationship illustrated in (1) and (2) further suggests that C_1 is interpreted as a subordinate clause; more evidence to this effect will appear in what is to follow, especially

section 6. Hence, under the default assumption that syntactic structure parallels semantics and that the antecedent in a CC is semantically subordinate,⁴ we should prefer the left-subordinate hypothesis. Moreover, the left-subordinate hypothesis correctly predicts the use of the subjunctive in (11) and the form of the tag in (12).⁵

Both clauses of a CC begin with a constituent bearing comparative morphology; let us call this clause type a *CC-clause*. A variant of the construction, which we will call the *CC' construction*, is shown in (13). Its second clause is a CC-clause, but its first clause retains the comparative phrase in its normal position.

(13) Mary got angrier and angrier, the more pictures she looked at.

The paraphrases in (14) show that the first clause of a CC', the one that bears normal order, plays a role parallel to that of C₂ in a CC.

- (14) a. The more pictures she looked at, the angrier Mary got. [= (13)]
 b. If/When/As she looked at more pictures, Mary got correspondingly angrier. [≈ (13)]

We will therefore call the first clause of a CC' C₂ and the second C₁, in order to preserve the semantic parallel. The fact that C₂ is in the form of a normal clause suggests again that it is a main clause and C₁ is subordinate.

This hunch is verified by the subjunctive and tag versions of the CC' construction; its first clause parallels the behavior of C₂ in a CC.

- (15) {It is imperative that
 {I demand that }
 a. John pay more, the more he eats.
 b. *John pays more, the more he eat.
- (16) a. You get angrier, the more we eat, don't you.
 b. *You get angrier, the more we eat, don't we.

We next note that it is possible to construct direct questions based on CCs. In such questions *the* in C₂ is replaced by *how much*, and inversion takes place in C₂ if necessary. The interpretation closely parallels that of sentences with *when/as*-clauses such as (17d). (The complementizer *that* cannot appear in C₂.)

⁴ There are of course other possible semantic accounts of CCs, which we will disregard here. For a first approximation, our assumption in the text concurs with Beck (1997) in seeing C₁ as semantically subordinate. In fact, though, our argument is valid under *any* semantic analysis of the CC construction: the point is that this meaning is expressed in two apparently distinct ways in syntax, one symmetrical and one asymmetrical. Hence, at least one of these must either present a syntax-semantics mismatch or have a covert syntax matching the other.

An *LI* reviewer points out that Pesetsky's (1989) analysis of counterfactual subject complements such as *To go there would have been stupid* presents another possible example of conditional semantics expressed in a non-*if*-clause format.

⁵ In Swedish, according to Verner Egeland (personal communication), the second clause in the corresponding construction displays verb-second effects whereas the first clause does not. To the extent that crosslinguistic evidence is telling in such constructions, this is a further piece of evidence that C₂ is a main clause and C₁ is subordinate.

- (17) a. The harder (that) it has rained, how much faster a flow (*that) appears in the river?
 b. The harder (that) it rains, how much faster (*that) do you run?
 c. The harder (that) it rains, how much faster a flow (*that) do you see in the river?
 d. When/As it rains harder, how much faster a flow (*that) appears in the river?

In principle, C_1 might also yield question forms. But it turns out that a question of the requisite form has the semantics of a CC' . For instance, (18a) approximately paraphrases (18b), the conditional paraphrase expected of a CC' . In other words, the questioned clause in (18a) is C_2 , even though it comes first.

- (18) a. How much harder has it rained, the faster a flow you see in the river?
 b. How much harder has it rained, when you see a faster flow in the river? [\approx (18a)]

Again this shows that C_2 is the clause with main clause force, in both the CC and CC' constructions.

Consider also the imperative construction. Standard imperative form, where the subject is omitted, is impossible in a CC (19a). With the proper kind of generic subject in C_2 and the proper pragmatics, though, a moderately acceptable imperative can be formed (19b). As might be now be expected, though, C_1 cannot have imperative form (19c). Again, this parallels the semantics of imperative conditionals (whose subject is however optional (19d)). We note also that a CC' can form an imperative on C_2 ; as befits the normal clause structure of C_2 , the subject can be omitted (19e). Evidently the subject is obligatory only in CC -clauses.

- (19) a. *The more John eats, the tighter keep your mouth shut about it.
 b. ?The more John eats, the tighter everyone keep your mouth shut about it, OK?/if you would.
 c. *The more everyone eat, the more John keeps his big mouth shut about it, OK?
 d. If/When/As John eats more, (everyone) keep your mouth shut tighter, OK? [parallels (19b)]
 e. (Everyone) keep your mouth shut tighter, the more John eats, OK?

Within the left-subordinate hypothesis, the facts in (18) and (19) follow automatically from standard assumptions. Within the paratactic hypothesis, we must say that the possibility of questions and imperatives follows from the semantics of main clause force, which is usually reflected in the syntax—but not here.

In a further surprising variant of the CC construction, C_2 (and more marginally, C_1 as well) takes the form of an accusative-gerundive complement.

- (20) a. ?I can well imagine the more he eats, the fatter him getting.
 b. ??I can well imagine the more him eating, the fatter him getting.

Like imperative CC s, gerundive CC s must have an overt subject present. There is no form (21b) parallel to the subjectless gerundive (21a).⁶

⁶ (20) presents an additional challenge. Usually a gerundive construction permits no fronted subordinate constituent

- (21) a. Bill can well imagine getting fat.
 b. *Bill can well imagine the more he eats, the fatter getting.

The CC' construction also appears in gerundives. Because its C_2 is a normal clause rather than a CC-clause, it can be subjectless.

- (22) Fred can well imagine (Joe) getting fatter, the more he eats.

The accusative-gerundive appears to be an isolated instance of a tenseless CC. For instance, a CC cannot appear as an infinitival complement, with or without an overt subject in C_2 . (23a–b) explore a number of conceivable configurations, all of which prove grossly ungrammatical. On the other hand, a CC', with the normal clause first, is acceptable here (23c).

- (23) a. *It is important (for) the more you (to) eat, the more careful ((for) you) to be.
 b. *It is important (for) the more to eat, the more careful to be.
 c. It is important [(for you) to be more careful, the more you eat].

To sum up, a wide variety of constructions in which CCs can be embedded suggest that the CC construction is asymmetric, with C_2 behaving as a main clause. More evidence will appear in section 6.

4 Evidence for the Paratactic Hypothesis

As mentioned in section 2, the paratactic hypothesis on the face of it seems unlikely, given the absence of parataxis from core constructions of English as well as the subordinate clause interpretation of C_1 extensively documented in section 3. Nonetheless, we will argue eventually that the paratactic hypothesis is correct and that the CC construction is an example of a syntax-semantics mismatch: paratactic (i.e., quasi-coordinate) syntax with conditional semantics. This section presents some initial plausibility arguments for such a position. What we take to be the crucial facts, however, are postponed till section 7.

We begin by noticing the existence of a colloquial construction with a conditional meaning similar to that of CCs, in which parataxis is about the only conceivable relation between the clauses. (We mark rising intonation with \uparrow and falling intonation with \downarrow .)

(except possibly with exaggerated intonation breaks), as seen in (i). In particular, a fronted conditional clause is disallowed (id).

- (i) a. ?*I can well imagine on Tuesday him leaving. [i.e., he leaves on Tuesday]
 b. *I can well imagine quickly Mary answering the question.
 c. ?*I can well imagine with a hatchet Mary destroying the Jeep.
 d. ?*I can well imagine if he eats more, him getting fat.

Hence, it does not appear that C_1 can be a (normal) fronted subordinate clause in (20). A possible escape is provided by the paratactic hypothesis, in which C_1 is not a subordinate adjoined constituent. However, formulating the constraints on the accusative-gerundive construction to admit this possibility is a challenge that we must leave for future research.

- (24) Mary listens to the Grateful Dead↓, she gets depressed↓. [≈ If/When Mary listens to the Grateful Dead, she gets depressed.]

This construction can be embedded, though not as freely as the CC construction; compare (25) with (5). (We find it helpful to put an intonation break after *that*, here indicated by a comma, in order to prevent a garden path interpretation of the first clause.)

- (25) a. ?I really think that, Mary listens to the Grateful Dead↓, she gets depressed↓.
 b. ?I'm not shocked by the idea that, Mary listens to the Grateful Dead↓, she gets depressed↓.
 c. ?It is becoming perfectly clear to me that, Mary listens to the Grateful Dead↓, she gets depressed↓.
 d. *It is not entirely obvious if/whether, Mary listens to the Grateful Dead↓, she gets depressed↓.
 e. ?I want to explain exactly why, Mary listens to the Grateful Dead↓, she gets depressed↓.

This construction thus gives some independent motivation for the existence in English of paratactic syntactic structures whose interpretation parallels that of sentences with preposed adjuncts. This lends the paratactic hypothesis for the CC construction a somewhat greater patina of legitimacy than it might have had in isolation.

In turn, (24) shows a semantic alliance with a coordinate construction involving what we elsewhere (Culicover and Jackendoff 1997) call ‘left-subordinating *and*’ (_{LS} *and*).

- (26) Mary listens to the Grateful Dead and she gets depressed. [= If/When Mary listens to the Grateful Dead, she gets depressed.]

In Culicover and Jackendoff 1997, elaborating on Culicover 1972,⁷ we showed that the most plausible account of left-subordinating *and* is as a syntactically ordinary conjunction that maps into a subordinate interpretation of C₁, that is, as a syntax-semantics mismatch. Some of our evidence will be reviewed in sections 6 and 7, in support of a parallel analysis of CCs. For present purposes, the point is that a conditional interpretation need not entail syntactic subordination; hence, the semantics of CCs does not constitute immediate grounds for ruling out the paratactic hypothesis.

A further consideration concerns the ordering of the clauses. Typically, a subordinate clause can occur either before or after its main clause.

- (27) a. If/When Mary listens to the Grateful Dead, she gets depressed.
 b. Mary gets depressed if/when she listens to the Grateful Dead.

(An exception, of course, is the *if-then* construction, which cannot be reordered into a *then-if*

⁷ Lakoff (1986) and others cited in Culicover and Jackendoff 1997 have observed similar facts, though in most cases in connection with *right*-subordinating *and*, as in *How many counterexamples can the CSC sustain t and still be believed?*

construction.) By contrast, neither a CC, nor a CC', nor either of the constructions in (24)–(26) can be reordered. In the cases where reordering is grammatical, the conditional meaning is reversed.

- (28) a. The angrier Mary got, the more she looked at pictures. [\neq The more she looked at pictures, the angrier Mary got.]
 b. *The more pictures Mary looked at, she got angrier and angrier.
 c. *Mary gets depressed↓, she listens to the Grateful Dead↑.
 d. Mary gets depressed and she listens to the Grateful Dead. [\neq (26)]

Thus, the CC and CC' constructions parallel these other nonsubordinating constructions, rather than standard constructions with adjunct subordinate clauses. (The behavior of *if-then*, however, keeps this from being an altogether decisive argument in favor of the paratactic hypothesis.)

We thus find ourselves in an interesting tension. On the one hand, all these constructions have approximately the semantics of the conditional. On the other hand, although the conditional itself ((1b), (14b)) syntactically subordinates the *if*-clause, these other constructions (the CC and CC' constructions and (24)–(26)) show no overt syntactic subordination. If semantic parallelism were always to require syntactic parallelism, we would be forced to assimilate all these latter constructions to subordinating syntax, at whatever price in the naturalness of syntactic derivations. Alternatively, we could accept complexity in the syntax-semantics interface, with the advantage of rendering the syntactic component far simpler. So the basic issue is in which component—the syntax itself or the syntax-semantics interface—the necessary complexity is to be localized. Our position is that there is no cut-and-dried answer; the question must be decided on a case-by-case basis.⁸

The paratactic hypothesis, of course, has one particularly pointed consequence for this issue. In the previous section, subjunctive, tag, question, and imperative formation showed that C_2 in CCs has main clause force and that C_1 is subordinate. If, as the paratactic hypothesis proposes, this subordination occurs only in semantics, we have to conclude that the notion ‘‘main clause force’’ is based on the semantics of the CC construction, not on its syntax; by extension the same would have to be the case in ordinary sentences as well. We are not yet ready to reject this conclusion, unusual though it may seem. For the moment we will leave the issue in tension, keeping both the left-subordinate and the paratactic hypotheses alive.

We now turn away temporarily from the choice between these two and address the balance between idiosyncrasy and regularity *within* C_1 and C_2 .

5 The Internal Structure of CC-Clauses

5.1 The Subject Requirement and Be-Omission in CC-Clauses

We noted in section 3 that a subject is obligatory in C_2 of imperative and gerundive CCs ((19) and (21)). By contrast, the subject can be omitted in C_2 of an imperative or gerundive CC' ((19e)

⁸ This tension is of course precisely what animated the generative semantics–interpretive semantics disputes of the late 1960s (Jackendoff 1972).

and (22)). We ascribed this to the fact that C_2 in a CC is a CC-clause, but C_2 in a CC' is a normal clause. Hence, part of the syntactic characterization of a CC-clause must be the obligatoriness of its subject.

McCawley (1988b) points out another curious aspect of CC-clauses: the verb *be* can be omitted under certain conditions. (29a) illustrates with both clauses of the CC construction. (29b) shows that *be*-omission is a characteristic of CC-clauses rather than of the CC construction as a whole: in a CC' *be* can be omitted in the CC-clause C_1 but not in the normal clause C_2 . McCawley further observes that the subject of the clause must have a nonspecific (generic or variable) interpretation. (29c), with a specific subject, illustrates: *be* cannot be omitted.

- (29) a. The higher the stakes (are), the lower his expectations (are).
 b. $\left. \begin{array}{l} \text{His expectations are lower} \\ \text{*His expectations lower} \end{array} \right\}$, the higher the stakes.
 c. $\left. \begin{array}{l} \text{The more obnoxious Fred is} \\ \text{*The more obnoxious Fred} \end{array} \right\}$, the less attention you should pay to him.

Since CC-clauses look a bit like comparatives, one might think this omission is related to that in ordinary comparatives such as (30a). However, (30b) shows that *be*-omission in ordinary comparatives does not depend on whether the subject is specific or not.

- (30) a. His expectations are always lower than mine (are).
 b. John was lots more obnoxious than Fred (was).

Hence, *be*-omission in CCs is probably not related directly to that in comparatives. Rather, the specificity restriction in CC-clauses appears more closely related to a parallel restriction in the complements of *no matter* and *wh-ever*, discussed by Culicover (1999). Note that these, like CC-clauses, have fronted predicates.

- (31) a. You should always lock your door, no matter how fancy the hotel (might be).
 b. I don't plan to lock the door, no matter how fancy this hotel *(is).
 c. I'm going out, whatever the weather.
 d. I'm going out, wherever that hurricane *(might be).

Although we suspect this omission is related to the semantics as well as the syntax of these constructions, we have no explanation to offer at this point. We note, however, that the constructions in (31) are very close in meaning to the CC construction: whereas the latter asserts a dependency of C_2 on C_1 , the former assert that the main clause has *no* dependency on the situation described in the free relative.

Again, however this *be*-omission is characterized, it appears to be a feature of CC-clauses, in this case shared with *no matter* and adverbial *wh-ever* clauses.

5.2 The Usual Constraints on Long-Distance Dependency

CC-clauses appear to have an internal structure involving a long-distance dependency between the comparative phrase at the front and a gap within the CP. Indeed, the gaps in the two clauses

show the typical constraints on long-distance dependencies. (32)–(36) illustrate with gaps in C_1 ; (37)–(41) illustrate with gaps in C_2 . We will assume for the moment that the comparative phrase is the antecedent of the trace.

Constraints in C_1

- (32) [The more counterexamples]_i Mary says that Bill has helped Fred to discover t_i , the less I believe her. [unbounded dependency in C_1]
- (33) a. *[The more food]_i Mary knows a man that eats t_i , the poorer she gets. [Complex NP Constraint (CNPC)]
 b. ?[The more great books]_i he makes the claim [that he has read t_i], the more suspicious I become of him. [CNPC]
 c. *[The fatter]_i he goes to a doctor when he gets t_i , the more he eats. [Condition on Extraction Domain (CED), Empty Category Principle (ECP)]
 d. *[The fatter]_i that [that he gets t_i] bothers him, the more he eats. [Subject Condition, ECP]
 e. ?[The more books]_i I ask whether he'll read t_i , the angrier he gets. [weak *wh*-island]
 f. *[The more books]_i I ask to whom_i, he will give t_j t_i , the more he reads. [strong *wh*-island]
 g. *[The more people]_i I ask what_j he will give t_j to t_i , the more he reads. [strong *wh*-island]
- (34) a. [The more carefully]_i he words the letter t_i , the safer he'll be.
 b. *[The more carefully]_i he knows a man that worded the letter t_i , the safer he'll be. [CNPC, ECP]
- (35) a. The more geniuses John meets, the angrier he gets.
 b. *[The more]_i John meets [t_i geniuses], the angrier he gets. [Left Branch Condition]
- (36) a. [The more people]_i you say t_i will buy tickets, the happier I'll be.
 b. *[The more people]_i you say that t_i will buy tickets, the happier I'll be. [*that*-t effect]
 c. [The more people]_i you say that right after the show opens t_i will buy tickets, the happier I'll be. [adverb effect]

Constraints in C_2

- (37) The more I talk to Joe, [the less about linguistics]_i I am inclined to think Sally has taught him to appreciate t_i . [unbounded dependency in C_2]
- (38) a. *The more he eats, [the poorer]_i he knows a woman that gets t_i . [CNPC]
 b. ?The more he eats, [the poorer]_i he makes the claim that he gets t_i . [CNPC]
 c. *The more he eats, [the fatter]_i he goes to a doctor when he gets t_i . [CED, ECP]
 d. *The more he eats, [the fatter]_i that [that he gets t_i] really bothers me. [Subject Condition, ECP]
 e. ?The more books he buys, [the more books]_i I wonder whether he'll read t_i . [weak *wh*-island]

- f. *The more he reads, [the more books]_i I wonder to whom_j he will give t_i t_j. [strong *wh*-island]
- g. *The more he reads, [the more people]_i I wonder what_j he will give t_j to t_i. [strong *wh*-island]
- (39) a. The sooner you call, [the more carefully]_i I will word the letter t_i.
 b. *The sooner you call, [the more carefully]_i I know a man that will word the letter t_i. [CNPC, ECP]
- (40) a. The richer John gets, the more geniuses John meets.
 b. *The richer he gets, [the more]_i John meets [t_i geniuses]. [Left Branch Condition]
- (41) a. The more articles he reads, [the fewer people]_i he thinks (*that) t_i will go into linguistics. [*that*-t effect]
 b. The more articles he reads, [the fewer people]_i he thinks (that) under the current circumstances t_i will go into linguistics. [adverb effect]

In CCs the gap can also be in subject position in C₁ (42a) or C₂ (42b).

- (42) a. The more people that t arrive, the louder (that) it gets.
 b. The more people (that) you give beer to, the more people that t get sick.

In such cases some speakers cannot omit *that* after the comparative phrase. Consider C₁ in (43a) and C₂ in (43b).

- (43) a. The more people arrive, the louder (that) it gets.
 b. The more people (that) you give beer to, the more people get sick.

In (42a) what is compared is the number of people that arrive. Some speakers (including RJ) find (43a) ambiguous between this interpretation and one in which what is compared is the extent to which people arrive (as in *the more that people arrive . . . t . . .*). Other speakers (including PWC) find that (43a) cannot paraphrase (42a), and only the latter interpretation is possible. For these speakers, *that* cannot be absent before a covert subject, just as in a standard relative clause.⁹

⁹The two clauses in CCs also show exactly the reconstruction and antireconstruction effects of relative clauses. Compare the relative clause examples in (i) with the CC examples in (ii).

- (i) a. [The pictures that John_i owns]_j that he_i likes t_j are in the living room.
 b. (*)[The pictures of John_i]_j that he_i likes t_j are in the living room.
- (ii) a. [The more pictures that John_i buys]_j that he_i likes t_j, the more arrogant he becomes.
 b. * [The more pictures of John_i]_j that he_i buys t_j, the more arrogant he becomes.
 c. [The more pictures of himself_i]_j that John_i buys t_j, the more arrogant he becomes.

It is well known that adjuncts moved into an \bar{A} -position do not reconstruct. Hence, in (ia), for example, *he_i* and *John_i* can be coreferential, even though *he_i* c-commands the trace of the constituent containing *John_i* and would otherwise be expected to produce a Condition C violation. On the other hand, (for some speakers—e.g., PWC but not RJ) an argument does reconstruct, so that (ib) is in fact a Condition C violation. The same pattern appears in the CC construction, which is precisely what we would expect if this were an instance of \bar{A} -movement.

The CC construction also licenses parasitic gaps.

- (iii) a. The more people he talks to t without inviting pg to the party, the more embarrassed he becomes.
 b. The more timid he feels, the more people he interviews t without asking questions of pg.

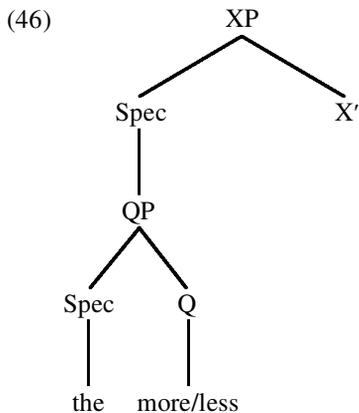
- (44) a. The man that arrived on the 10:53 was my brother.
 b. *The man arrived on the 10:53 was my brother.

5.3 The Upper End of the Long-Distance Dependency

Consider the structure of the comparative phrase. *More* and *less* are, as usual, quantifiers; they can occur in the specifier of NP (45a), AP (45b), AdvP (45c), or PP (45d), or alone as an adverbial modifier of VP (45e) (Jackendoff 1977:chap. 6). (*More* is of course subject to the usual morphological alternations of the comparative, yielding forms such as *bigger*, *fatter*, and the suppletions *better* and *worse*.) *The* appears to be a specifier of *more* and *less*, in alternation with other specifiers such as *much*, *far*, and *a lot*. (Jespersen (1949:509–512) points out that it bears no diachronic relation to the definite article. There are a few related relics in the language, such as *none/not much the wiser*, *the better to see you with*, and *all the more*; we return to the last of these in section 5.4.)

- (45) a. (much/far/a lot) more/less custard
 b. (much/far/a lot) more/less beautiful
 c. (much/far/a lot) more/less quickly
 d. (much/a lot) farther [= more far] down the road
 e. sleep (much/far/a lot) more/less

So the overall structure of the comparative phrase is roughly (46) (except for (45e), where the topmost XP, Spec, and X' are absent).



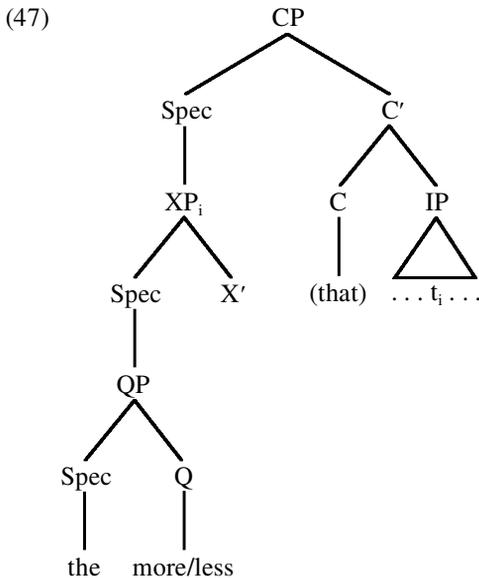
Curiously, the CC construction does not license parasitic gaps where the parasitic gap is contained within a relative clause that precedes the true gap.

- (iv) a. *The more people everyone who likes pg pays attention to, the happier we all are.
 b. *The later it gets, the more people everyone who likes pg pays attention to t.

We have no explanation for this fact.

As with fronted *wh*-phrases, fronted comparative phrases carry along everything of which *the more* is a left branch.

What is the position of the comparative phrase? The structures in (32)–(41) assume that there is a direct dependency between the comparative phrase and the trace position. In the usual movement theory of long-distance dependencies, this amounts to saying that the comparative phrase moves to [Spec, CP]. (In a nonmovement theory such as Head-Driven Phrase Structure Grammar (HPSG), the comparative phrase is generated in [Spec, CP] and a connection is formed between it and the empty category.) The most plausible structure is thus (47), paralleling for instance indirect questions.



5.4 Behavior of the more Compared to That of Other Specifiers of QP

For a further basis of comparison, let us examine yet another construction semantically related to the CC construction, illustrated in (48), in which *all the* appears as a specifier of *more* or *-er*, parallel to *much*, *far*, and *a lot*. Crucially for the parallel, *all the* can appear as specifier in C₂ of a CC' (48c); *much*, *far*, and *a lot* are less felicitous here (48d).

- (48) a. Whenever Bill smokes, Susan hates him all the more/much more/far more/a lot more.
 b. Once Janet left, Fred became all the crazier/much crazier/far crazier/a lot crazier.
 c. Fred became all the crazier, the more often Janet left.
 d. Fred became *much crazier/?far crazier/*a lot crazier, the more often Janet left.

Unlike *much*, *far*, and *a lot*, *all the more* can front. When it fronts, it brings along in the usual way the entire projection of which it is the specifier, and it requires subject-auxiliary inversion

(49). Coincidentally, such inversion occurs marginally in CCs as well, in C_2 but not C_1 (50).

- (49) a. i. When Bill smokes, all the more/*much more does Susan hate him.
 ii. Once Janet left, all the crazier/*much crazier did Fred become.
 b. i. *When Bill smokes, all the more Susan hates him. [no inversion]
 ii. *Once Janet left, all the crazier became Fred.
- (50) a. ?The more Bill smokes, the more does Susan hate him.
 b. ?The more often Janet leaves, the angrier does Fred become.
 c. i. *The more does Bill smoke, the more Susan hates him. [inversion in C_1]
 ii. *The more does Bill smoke, the more does Susan hate him.

The degree specifier *so* also triggers inversion, but *how* (in indirect questions) does not.

- (51) a. So much did you eat that . . .
 b. So fast did you run that . . .
 c. So intelligent a dog did you buy that . . .
- (52) a. I know how much you ate.
 b. I know how fast you ran.
 c. I know how intelligent a dog you bought.

There are still more distinctions. First, as seen in (48), *all the more* need not appear in clause-initial position; the same is true of *so* (53). By contrast, *the more* must appear in clause-initial position, like *how* (54).

- (53) a. He ate so much that he got sick.
 b. So much did he eat that he got sick.
- (54) a. The more you eat, the more you want.
 b. *You eat the more, the more you want.
 c. *The more you eat, you want the more.
 d. *I wonder you ate how much.

Second, unlike these other specifiers, *the more* in initial position permits the presence of the complementizer *that*, as observed earlier. Under a standard analysis, this is a violation of the Doubly Filled Comp Filter.

- (55) a. The faster that you eat, . . .
 b. *. . . all the faster that do you eat.
 c. *So much faster that did you eat that . . .
 d. *How much faster that you eat . . .

Third, *the more* does not allow pied-piping of a preceding preposition. *All the more* is marginal here, and *so* and *how* are of course acceptable.

- (56) a. *To the more people Bill talks, . . .
 b. ?*. . . to all the more people will Bill talk.

- c. To so many people does Bill talk that . . .
- d. I wonder to how many people Bill talks.

On the whole, then, (47) represents a reasonable structure for CC-clauses: they fall together in many respects with other constructions in which there is movement to the specifier of a complementizer. On the other hand, like the CC construction's unusual IP-over-CP phrase structure, its long-distance dependencies appear to some extent *sui generis*. In particular, one must deal with (a) the obligatory movement of the *the more*-phrase, in contrast with the optional movement of *so* and the nearly synonymous *all the more*; (b) the optional presence of the complementizer *that* in what appears to be a main clause (at least C₂, and, we will argue, C₁); (c) the marginal possibility of inversion in C₂, in contrast with the obligatory inversion with *all the* and *so* and its impossibility in indirect questions and C₁; and (d) the failure of a preposition to pied-pipe with *the more*.

One may of course attempt to use the Minimalist Program technology of functional heads, strong features, and attraction to describe these idiosyncratic properties of the CC construction; but as far as we can see, these properties are not in any interesting sense reducible to those of other constructions in the language. It would be equally natural—in fact, more natural from a theory-neutral perspective—to say simply that these are the properties of the construction, without reference to formal devices that have no particular explanatory role here.

Instead of encoding these peculiarities in terms of the movement of *the more* to C, one might adapt a treatment of long-distance dependencies advanced in differing versions by HPSG (Pollard and Sag 1994, Sag 1997), Construction Grammar (Fillmore and Kay 1993), and Tree-Adjoining Grammar (Frank and Kroch 1995). In these frameworks the grammar contains a general treatment of long-distance dependencies in terms of the relation between a gap (or trace) within a clause and a constituent or operator in front of the clause. This constituent, along with properties of C, subject, and I of the uppermost clause of the construction, are characteristic of the construction; we might call them together the construction's "signature." Different long-distance dependency constructions have different signatures. For example, English tensed relative clauses permit a *wh*-phrase, *that*, or null in their signature; infinitival relative clauses permit a *wh*-PP, a *for* + subject NP, or null in their signature; free relatives permit, among other things, *wh* + *ever* in their signature (*whoever Bill sees*). The signature is regarded as base-generated, not as moved from some underlying syntactic position; hence, it can be described directly in phrase structure terms.¹⁰ Its relation to the trace is regarded as a constraint on surface configuration. Any particular long-distance dependency construction, then, involves (a) specifying a signature, including the structure of the fronted constituent and I-features of the clause to which the signature is adjoined (e.g., tensed, infinitival, inverted), and (b) inheriting all the general properties of long-distance dependencies from the constraints on relation of signatures to traces. Proposals of this sort have been in the literature, even within the GB tradition, for years (e.g., Brame 1978, Koster 1987). Nevertheless,

¹⁰ One might use inheritance hierarchies to capture the similarities among the constructions, as in Fillmore and Kay 1993, Michaelis and Lambrecht 1996, and Sag 1997.

the movement analysis has been predominant within the standard tradition, often introducing empty operators or equivalent devices in order to have something to move.

The problem for a nonmovement theory, of course, is to explain the constraints on possible signatures: the possibilities seem rather stipulative. We know of no attempts to address this problem. On the other hand, we know of no (recent) attempts to address the full variety of signatures within the movement approach either. Thus, the CC construction adds yet another situation on the basis of which the two approaches must be compared.

In any event, we have concluded that the internal structure of CC-clauses presents a number of idiosyncrasies: the obligatoriness of the subject in imperatives and gerundives, the possibility of omitting *be* with a nonspecific subject, and the peculiar configuration in the specifier of the comparative phrase. The last of these is pretty much within the range of variation for constructions in which a specifier motivates fronting of a phrase. On the other hand, CC-clauses help point up how peculiar this range of variation is, a fact neglected in the recent tradition of GB, principles and parameters, and the Minimalist Program.

6 Binding

We return now to a fuller investigation of the relationship between C_1 and C_2 . Sections 3 and 4 presented evidence that C_1 functions as a subordinate clause and C_2 functions as a main clause. This section will provide additional evidence to support such a conclusion.

The evidence concerns binding of reflexives and pronouns. As in other work (Culicover and Jackendoff 1995, 1997), we will argue here that evidence of this sort does not lead to the conclusion that the subordinate clause is syntactically subordinate (i.e., an adjunct), but simply supports the conclusion that it is subordinate at least at the level of semantic interpretation. In the next section we will in fact argue that the *syntactic* evidence shows that the clauses are paratactic and that C_1 is subordinate *only* at the level of semantic/conceptual structure.

McCawley (1988b) points out that in the CC construction, pronouns in C_1 can have antecedents in C_2 , paralleling fronted subordinate clauses (57a–b). Furthermore, in a CC', pronouns in C_2 cannot have antecedents in C_1 , so CC' constructions parallel sentences ending with *if*-clauses (57c–d). ((57) is McCawley's (6).)

- (57) a. The longer he has to wait, the angrier John gets.
 b. If he has to wait, John gets angry.
 c. *He_i gets angry, the longer John_i has to wait.
 d. *He_i gets angry if John_i has to wait.

Next let us consider the binding of a reflexive. As we showed in Culicover and Jackendoff 1997, a reflexive complement of a picture-noun can have an antecedent in another clause if (i) the reflexive is in a clause subordinate to the clause that contains the antecedent, and (ii) the reflexive is logophoric (i.e., (roughly) its antecedent is understood as having an attitude toward the situation described by the clause containing the reflexive). The examples in (58)–(59) show that this relationship holds both in standard conditionals and in sentences with left-subordinating *and*, but does not hold in ordinary conjoined clauses.

- (58) a. *Conditional with logophoric reflexive*
 If another picture of him(self)_i appears in the news, (Susan's afraid that) John_i will really get angry.
- b. *Left-subordinating and with logophoric reflexive*
 Another picture of him(self)_i appears in the news and (Susan's afraid that) John_i will get really angry.
- c. *Ordinary conjunction with logophoric reflexive*
 Another picture of him(*self)_i appeared in the news yesterday, and unfortunately (Susan's afraid that) John_i will really get angry.
- (59) a. *Conditional with nonlogophoric reflexive*
 If another picture of him(*self)_i appears in the news, (Susan suspects) John_i will be arrested.
- b. *Left-subordinating and with nonlogophoric reflexive*
 Another picture of him(*self)_i appears in the news and (Susan suspects) John_i will get arrested.

On the basis of examples such as these we argue that condition (i) must be stated over conceptual structure representations, since it is at this level where the conditional and left-subordinating *and* share the subordinate character of the clause containing the reflexive. Furthermore, inasmuch as logophoricity inherently involves aspects of clause meaning, it is impossible to state condition (ii) in purely syntactic terms.

Similar examples can be constructed for the CC construction, with the reflexive in C₁, the putative subordinate clause.

- (60) *Logophoric reflexive*
- a. The more pictures of him(self)_i (that) appear in the news, the more embarrassed John_i becomes.
- b. The more (frequently) that pictures of him(self)_i appear in the news, the more embarrassed John_i becomes.
- (61) *Nonlogophoric reflexive*
- a. The more pictures of him(*self)_i (that) appear in the news, the more likely John_i is to get arrested.
- b. The more (frequently) that pictures of him(*self)_i appear in the news, the more likely John_i is to get arrested.

(62) shows that a logophoric reflexive cannot appear in C₂, the putative main clause—even though it *follows* its antecedent. (For some reason, it is not easy to find an example that is clearly logophoric; we think (62) is a reasonable case.)

- (62) The more that John_i gets upset by them, the more that stories about him(*self)_i seem to show up on the evening news.

In the CC' construction, where C₂ comes first, still the reflexive can only appear in C₁. Again logophoricity is necessary.

- (63) a. John_i is more embarrassed, the more pictures of him(self)_i (that) appear in the news.
 b. John_i is more embarrassed, the more (frequently) that pictures of him(self)_i appear in the news.
 c. John_i is more likely to get arrested, the more (frequently) that pictures of him(*self)_i appear in the news. [nonlogophoric]
 d. Stories about him(*self)_i seem to show up more on the evening news, the more that John_i gets upset by them. [logophoric reflexive in C₂]

Another type of binding evidence that supports the conclusion that C₁ is a subordinate clause at some level of representation involves the binding of a pronoun by a quantifier. (Examples adapted from Culicover and Jackendoff 1997.)

- (64) a. If you give him_i enough opportunity, every senator_i will succumb to corruption. [conditional]
 b. You give him_i enough opportunity and every senator_i will succumb to corruption. [L_S *and*]
 c. *We gave him_i enough opportunity and, sure enough, every senator_i succumbed to corruption. [ordinary conjunction]

(64b) shows that a coordinate structure whose first conjunct has a subordinate interpretation allows leftward binding by *every*, parallel to preposed *if*-clauses; a semantically coordinate structure such as (64c) does not. In addition, fronted conditionals and left-subordinating *and* allow rightward binding by *any* (but not *every*) out of the first clause.

- (65) a. If you give any/*every senator_i enough opportunity, he_i will succumb to corruption.
 b. You give any/*every senator_i enough opportunity and he_i will succumb to corruption.
 c. *We gave any/every senator_i enough opportunity and, sure enough, he_i succumbed to corruption.

CCs have a similar pattern.

- (66) a. The more lobbyists he_i talks to, the more corrupt every senator_i seems to become.
 b. The more lobbyists wine and dine him_i, the more every senator_i is susceptible to improper influence.
 c. The more time that any/*every senator_i spends with lobbyists, the more likely he_i is to succumb to corruption.

The pattern for the CC' construction (67a–b) parallels that of postposed *if*- or *as*-clauses (67c–d).

- (67) a. Every/Any senator_i becomes more corrupt, the more lobbyists he_i talks to.
 b. *He_i seems to become more corrupt, the more lobbyists every/any senator_i talks to.
 c. Every/Any senator_i seems to become more corrupt, if/as he_i talks to more lobbyists.
 d. *He_i seems to become more corrupt, if/as every/any senator_i talks to more lobbyists.

To sum up this section, we have found confirming evidence from anaphora and binding that

C_1 in a CC is interpreted as a subordinate clause. Under the assumption that binding is determined (entirely) by syntactic configuration, this would constitute a strong argument for the left-subordinate hypothesis of the CC construction. But if binding has (in part) a semantic basis, as we and others have argued (e.g., Culicover and Jackendoff 1995, 1997, Jackendoff 1992, Fauconnier 1985, Kuno 1987, Van Valin and LaPolla 1997), then this evidence does not distinguish the left-subordinate hypothesis from the paratactic hypothesis. In the next section we provide evidence that the subordination is not in syntactic structure.

7 Extraction from Comparative Correlatives

If C_1 is a subordinate clause and C_2 is a main clause, as suggested by the data discussed in sections 3 and (perhaps) 6, we would expect—at best—that extraction from C_2 would be unproblematic whereas extraction from C_1 would produce CED violations. Moreover, since both clauses contain long-distance dependencies parallel to those in indirect questions and free relatives, we might expect *both* clauses to be islands by virtue of their internal structure.

These predictions turn out to be incorrect. Extraction from C_2 is indeed possible; and, more surprisingly, so is extraction from C_1 . In fact, if anything, extraction from C_1 sounds better to our ears, perhaps because it does not involve intervening clausal material. We consider relative clauses in (69), topicalization in (70), and *it*-clefts in (71).¹¹

- (68) The sooner you solve this problem, the more easily you'll satisfy the folks up at corporate headquarters. [“base” sentence]
- (69) a. This is the sort of problem which_i the sooner you solve t_i, the more easily you'll satisfy the folks up at corporate headquarters. [extraction from C_1]
 b. The folks up at corporate headquarters are the sort of people who_i the sooner you solve this problem, the more easily you'll satisfy t_i. [extraction from C_2]
- (70) a. This problem_i, the sooner you solve t_i, the more easily you'll satisfy the folks up at corporate headquarters. [extraction from C_1]
 b. ?The folks up at corporate headquarters_s, the sooner you solve this problem, the more easily you'll satisfy t_i. [extraction from C_2]

¹¹ It is also possible to extract from both clauses simultaneously.

(i) This is the problem that the quicker you solve t, the quicker you'll be able to tell your friends about t.

This might be seen as across-the-board extraction, even though the two clauses are not (superficially) coordinate conjuncts. However, we suspect that one of the extractions is actually a parasitic gap of the sort found when both extractions are independently possible; this situation would be similar to multiple extractions from two complements of the same verb, as in (ii).

(ii) a. Who did you tell t that you would pay a call on t?
 b. Who did you give pictures of t to friends of t?

Given the possibility of examples such as those in (ii), the possibility of parasitic gaps in CCs cannot serve as a diagnostic for structure.

- (71) a. It is this problem_i that the sooner you solve t_i, the more easily you'll satisfy the folks up at corporate headquarters. [extraction from C₁]
 b. ?*It is the folks up at corporate headquarters_i (not the ones here at the regional office) who_i the sooner you solve this problem, the more easily you'll satisfy t_i. [extraction from C₂]

Wh-questions should also be possible with CCs, but the simplest cases are not (72).

- (72) a. *Which problem does the sooner (that) you solve, the more easily you'll satisfy the folks up at corporate headquarters?
 b. *Which problem the sooner (that) you solve, will the more easily you satisfy the folks up at corporate headquarters?

One possible reason for this ungrammaticality is that the clause-initial comparative phrase blocks inversion in main clauses. However, when the *wh*-phrase is a subject, and hence need not cause inversion, a CC is still impossible.

- (73) a. *The harder it rains, the faster who runs?
 b. *The louder who talks, the angrier you get?

This suggests that both (72) and (73) are unacceptable because the comparative phrase occupies the initial position in the clause, preventing the *wh*-phrase from having this status. In (73) the *wh*-phrase is thus in situ, where it cannot get the necessary wide scope interpretation. This analysis is confirmed when we recall from section 3 that direct questions can be posed in C₂ by substituting *how much* for *the*. Notice that in this case inversion is acceptable.

- (74) The harder (that) it rains, how much faster a flow do you see in the river?

Here *how much* evidently serves in both the role of *wh*-phrase and that of *the more*, so that a single initial position suffices.

Embedded *wh*-questions and long *wh*-movement turn out to be possible, evidently because they do not raise these problems.¹²

- (75) a. They failed to tell me which problem the sooner I solve t, the quicker the folks up at corporate headquarters will get off my back. [extraction from C₁]
 b. ??I finally worked up enough courage to ask which people up at corporate headquarters the sooner I solve this problem, the quicker I'll get free of t. [extraction from C₂]
- (76) a. Which problem do you think that the sooner you solve t, the quicker you'll be able to tell the folks up at corporate headquarters to buzz off?

¹² Note, however, that *which problem* in (75) does precede the comparative phrase, a configuration we found impossible in (72). There are at least two possible accounts of this difference. First, it might be that main clauses beginning with *the more* lack a further initial position to which a *wh*-phrase could move, but such a position is available with subordinate *the more*-clauses. Second, it might be that comparative clauses lack a position before *the more* to which I could move when a fronted *wh*-phrase makes inversion necessary. The exact solution is not crucial to our general point.

- b. ??Which folks up at corporate headquarters do you think that the sooner you solve this problem, the quicker you'll be able to tell t to buzz off?

What are we to make of these facts? In Culicover and Jackendoff 1997 we found that extraction is the one place where left-subordinating *and* differs from its conditional paraphrase: although as usual one cannot extract from an adjunct *if*-clause, one can extract from either clause conjoined by left-subordinating *and*.

- (77) a. ??This is a problem that you'll be able to tell the folks up at corporate headquarters to buzz off if you solve t.
 b. ?This is a problem that you solve t and you'll immediately be able to tell the folks up at corporate headquarters to buzz off. [_{LS}*and*: better than (77a)]
 c. ?Those are the folks that you just solve *this* problem and you'll be able to put t on ice. [_{LS}*and*: better than (77a)]

Though admittedly the judgments are delicate, we find that *if*-clause extraction is improved by a resumptive pronoun, but extraction from a left-subordinating conjunction becomes worse.

- (78) a. This is a problem that you'll be able to tell the folks up at corporate headquarters to buzz off if you solve it. [better than (77a)]
 b. ??This is a problem that you solve it and you'll be able to tell the folks up at corporate headquarters to buzz off. [worse than (77b) and (78a)]
 c. ??Those are the folks that you just solve *this* problem and you'll be able to put them on ice. [worse than (77c)]

Our reasoning in Culicover and Jackendoff 1997 was

- that constraints on binding and quantification apply at the level of conceptual structure, where conditionals and left-subordinating coordination are equivalent;
- that the Coordinate Structure Constraint (CSC) applies at the level of conceptual structure, where standard coordinating conjunction is a coordinate structure but left-subordinating coordination is not—hence, the CSC does not apply to left-subordinating coordination;
- that left-subordinating coordination is however still coordination in syntactic structure; and
- that the constraints on extraction from adjuncts apply at the level of syntactic structure, so that they apply to *if*-clauses but not to left-subordinating coordination.

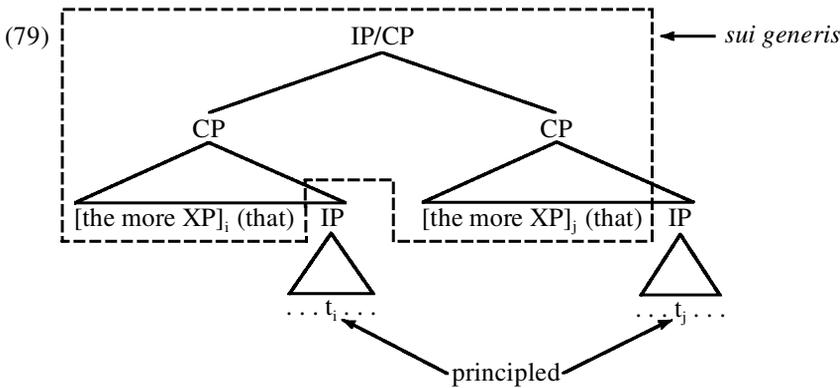
We concluded that this is the only possible way of partialing out the constraints that enables us to account for the facts. The consequence is that left-subordinating conjunction is a case of a syntax-semantics mismatch: coordinate in syntax but subordinating in semantics.

On the basis of the fact that extraction is possible from both C_1 and C_2 in CCs, we arrive at a conclusion parallel to that for left-subordinating conjunction: C_1 is not an adjunct, despite all the evidence for its subordinate semantics. As noted in section 2, there is no evidence that it is a conjunct, either, in the sense of being linked to C_2 by an obligatorily empty conjunction. In

fact, the most plausible hypothesis on these grounds is the paratactic hypothesis: C_1 and C_2 are paratactically linked to form an IP or CP.

Recapitulating, then, the CC construction is *sui generis*, in that its basic paratactic structure does not conform to the general patterns of X-bar theory. The clauses that make up this construction have many of the properties of CP, and the comparative phrase is linked to a trace by a classic long-distance dependency. However, the specifier of the comparative phrase in both clauses is sufficiently idiosyncratic that there is no natural way to subsume it completely under normal *wh*-movement or under the fronting of other degree specifiers such as *all the* and *so*. Finally, C_1 has the interpretation of a subordinate clause, and C_2 has the interpretation of a main clause; but from the perspective of extraction, both clauses have the status of coordinate clauses (i.e., neither is an adjunct or argument of the other).

(79) sums up some details of the construction, distinguishing the *sui generis* parts of the structure from the parts that follow from more general principles.



Two major questions remain. First, it is not clear to us why the CC construction permits extraction from what appears to be a complex construction along the lines of an indirect question or free relative; such constructions are normally islands to extraction. The uniqueness of the construction makes it difficult to find analogous cases elsewhere in the language; we will not pursue the question further here.¹³

¹³ One potential hypothesis is revealed by attempting extraction from a CC'. To our ears, extraction from the normal clause C_2 is perfectly acceptable, as might be expected, but extraction from the CC-clause C_1 turns out to be fairly bad.

(i) You'll beat the competition more easily, the sooner you solve this problem. [“base” sentence]

(ii) a. They failed to tell me who I'll beat t more easily, the sooner I solve this problem. [extraction from C_2]
 b. ??They failed to tell me which problem I'll beat the competition more easily, the sooner I solve t. [extraction from C_1] [worse than (75a–b)]

(iii) a. These are the guys that you'll beat t more easily, the sooner you solve this problem.

b. ??This is the problem that you'll beat the competition more easily, the sooner you solve t. [worse than (69a–b)]

This suggests that in the CC' construction C_2 is syntactically a main clause and C_1 is subordinate, matching the semantics. (However, our judgments are indeterminate with respect to whether a resumptive pronoun improves (iib) and (iiib), as

Second, recall from section 4 that the paratactic analysis led to an unsettling consequence: the notion of “main clause force,” necessary to determine where the subjunctive appears and on which clause imperatives, tags, and *how much* questions can be formed, is not a syntactic matter under this analysis. Rather, “main clause force” distinguishes C_2 from C_1 only in the semantics. We are not certain to what extent this forces reconsideration of many deeply held beliefs about syntactic structure. For now, we must leave such reconsideration, important though it may be, for future research.

8 Consequences for Universal Grammar

The CC construction, then, is another construction that confirms the general perspective of Culicover and Jackendoff 1995, 1997 concerning the relationship between syntactic structure and semantic structure. On our view, syntactic structures may reflect to a greater or lesser degree the semantic structures they express. On the plausible assumption that conceptual structure is universal, universals of conceptual structure will in the default case project invariantly into syntactic universals. Such a default projection may be able to explain, for example, why all languages have nouns and verbs, why verb phrases have more or less the same form across languages, why adjuncts appear external to the arguments of a phrase crosslinguistically, and other apparent but not entirely rigid universals of X-bar theory. The “two dogmas” discussed in section 1 attempt to assimilate all aspects of syntax to this default situation.

However, at the other extreme, the mapping of syntactic structure into conceptual structure is more or less arbitrary. This aspect of syntactic structure is autonomous and unpredictable, insofar as it is *not* a projection of universal conceptual structure. The CC construction is one such case, we argue. Although the internal structure of the construction’s individual clauses follows the standard principles of semantics-syntax mapping, its syntax at the level above the clause and its mapping into conceptual structure is *sui generis* and follows from no general principles, as far as we can tell, as shown in (79).

In this light, let us consider the status of the extraction constraints observed in sections 5 and 7. Section 5 showed that extraction of the comparative phrase *within* the clauses of a CC is

it does (78a).) If these admittedly delicate judgments are correct, the CC construction differs from the CC’ construction because of the total symmetry of its clauses in syntax; just as in a coordinate construction, there is no syntactic evidence that one clause is subordinate to the other. Such an analysis would comport with the treatment of left-subordinating *and*, where there is also no evidence in the syntax for an asymmetry between the conjoined clauses, and where, as we have argued, extraction acts as though both are main clauses rather than conjuncts.

It is also worth noting that when the complementizer *that* is present, the possibilities for extraction are more restricted. Consider the following contrasts:

(iv) This is the kind of rice that the quicker (*that) you cook *t*, the better it tastes.

(v) This is the kind of rice that the more pounds (??that) you cook *t*, the longer it takes.

To our ears, the insertion of *that* with an adjunct extraction from a CC, as in (iv), produces ungrammaticality, whereas it has a weaker effect when an argument is extracted, as in (v). We might speculate that the presence of *that* produces a reanalysis in which the head *quicker* or *more pounds* is treated like the head of a relative clause. The difference in (iv)–(v) brings to mind the classical argument/adjunct asymmetries, and, although it does not resolve the question of extraction raised in the text, it may shed some additional light on it.

subject to the usual constraints; but section 7 showed that extraction from the clauses of a CC to *outside* the CC is not so constrained. On this basis we argued that the clauses of a CC below the level of the adjunction site of the comparative phrase have normal clausal structure, but that the two clauses of a CC are connected paratactically.

The question naturally arises of how a learner could possibly acquire this information. A learner would be unlikely to have enough experience with CCs to be able to identify their crucial properties. It is especially striking that speakers judge C_1 to be syntactically a main clause in the absence of evidence, given that, as shown in sections 3 and 6, this clause receives a subordinate interpretation.

One possibility is that UG permits parataxis as a generally available syntactic option in natural language, making available coordinate clauses that are individually accessible to extractions just as simple clauses are. On the view that constraints on extraction are simply a product of the relative syntactic complexity of the configuration (see Culicover 1999 for one version of this view), nothing more needs to be said about extraction from such clauses. In particular, if a construction exhibits no syntactic subordination, then it will contribute no extraordinary complexity to extraction.¹⁴ In the CC construction the subordinate interpretation of one of the clauses is only a *semantic* fact, and hence irrelevant to syntactic complexity for purposes of extraction.

On this view, what the child must learn is not the possibility of parataxis but the particular semantics to associate with the paratactic CC construction and its counterparts in other languages. It is surely significant that many other languages have similar constructions expressing the comparative correlative reading found in CCs, although specific syntactic properties of the constructions differ from language to language. It is as though universal semantics has a kind of meaning to express—the comparative correlation of two scales of magnitude—for which standard default mapping to syntax does not provide a complete solution; and languages therefore are forced to “cobble together” some kind of mechanism to express it. We have not explored the extent to which such comparable constructions provide evidence for or against paratactic syntax. But in many cases the constructions are superficially unusual in just the way the English construction is, often appearing paratactic and often inheriting properties of comparatives and/or conditionals, as McCawley (1988b) and Beck (1997) have documented. At the moment we are unable to be more explicit about this intuition, or about its consequences for the theory of UG. More crosslinguistic work is called for on this and other idiosyncratic constructions (e.g., left-subordinating *and*, which appears to be widespread).

Let us close by reflecting on the consequences of the CC construction for the “two dogmas” of section 1. The CC construction has proven in numerous respects to behave unlike more familiar “core grammar.” Yet it *is* learned; and its details emerge as surprisingly robust, given the surely minimal exposure any speaker has to such exotica as imperative CCs and extraction from embedded CCs. Although we are far from solving all the puzzles that the CC construction poses, we

¹⁴ We note that this approach runs orthogonal to another thread in the literature (see, e.g., Kroch 1989, Pollard and Sag 1994), in which at least some apparent constraints on extraction are due to the pragmatic oddity of questioning or relativizing in certain presuppositional contexts. However, we see nothing to prevent *both* conclusions from being drawn at once.

have seen clearly the rich interaction of the totally idiosyncratic with the more general that this construction exemplifies. The moral of these results, we believe, is that the ‘‘periphery’’ cannot be neglected in the study of UG, in contrast to what has been assumed in many influential quarters.

We also conclude, as we have in earlier work, that although syntactic form reflects and to an extent depends on semantic representation, syntax and semantics are independent systems of representation with their own nontrivial autonomous principles of organization, principles that remain to be discovered by research in both domains. Under such a view, the structure of the mapping between the two levels must be regarded as nontrivial and constitutes a further important domain of research on UG, again in contrast to more standard approaches.

References

- Beck, Sigrid. 1997. On the semantics of comparative conditionals. *Linguistics and Philosophy* 20:229–232.
- Brame, Michael. 1978. *Base generated syntax*. Seattle, Wash.: Noit Amrofer.
- Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, Mass.: MIT Press.
- Chomsky, Noam. 1972. *Studies on semantics in generative grammar*. The Hague: Mouton.
- Chomsky, Noam. 1986. *Knowledge of language: Its nature, origin, and use*. New York: Praeger.
- Chomsky, Noam. 1995. Categories and transformations. In *The Minimalist Program*, 219–394. Cambridge, Mass.: MIT Press.
- Culicover, Peter W. 1972. OM-sentences. *Foundations of Language* 8:199–236.
- Culicover, Peter W. 1999. *Syntactic nuts: Hard cases in syntactic theory*. Oxford: Oxford University Press.
- Culicover, Peter W., and Ray Jackendoff. 1995. *Something else* for the binding theory. *Linguistic Inquiry* 26:249–275.
- Culicover, Peter W., and Ray Jackendoff. 1997. Semantic subordination despite syntactic coordination. *Linguistic Inquiry* 28:195–217.
- Fauconnier, Gilles. 1985. *Mental spaces*. Cambridge, Mass.: MIT Press.
- Fillmore, Charles J., and Paul Kay. 1993. *Construction Grammar coursebook*. Berkeley, Calif.: Copy Central.
- Fillmore, Charles J., Paul Kay, and Mary Catherine O’Connor. 1988. Regularity and idiomaticity in grammatical constructions: The case of *let alone*. *Language* 64:501–538.
- Frank, Robert, and Anthony Kroch. 1995. Generalized transformations and the theory of grammar. *Studia Linguistica* 49:103–151.
- Jackendoff, Ray. 1972. *Semantic interpretation in generative grammar*. Cambridge, Mass.: MIT Press.
- Jackendoff, Ray. 1977. \bar{X} syntax: A study of phrase structure. Cambridge, Mass.: MIT Press.
- Jackendoff, Ray. 1992. Mme. Tussaud meets the binding theory. *Natural Language & Linguistic Theory* 10:1–31.
- Jackendoff, Ray. 1997. *The architecture of the language faculty*. Cambridge, Mass.: MIT Press.
- Jespersen, Otto. 1949. *A Modern English grammar on historical principles*. Part VII, *Syntax*. London: George Allen & Unwin.
- Katz, Jerrold J., and Paul M. Postal. 1964. *An integrated theory of linguistic descriptions*. Cambridge, Mass.: MIT Press.
- Kay, Paul, and Charles Fillmore. 1999. Grammatical constructions and linguistic generalizations: The *What’s X doing Y?* construction. *Language* 75:1–33.
- Koster, Jan. 1987. *Domains and dynasties: The radical autonomy of syntax*. Dordrecht: Foris.
- Kroch, Anthony. 1989. Amount quantification, referentiality, and long *wh*-movement. Ms., University of Pennsylvania, Philadelphia.
- Kuno, Susumu. 1987. *Functional syntax*. Chicago: University of Chicago Press.
- Lakoff, George. 1986. Frame semantic control of the Coordinate Structure Constraint. In *CLS* 22. Part 2,

- The Parasession on Pragmatics and Grammatical Theory*, 152–167. Chicago Linguistic Society, University of Chicago, Chicago, Ill.
- May, Robert. 1985. *Logical Form*. Cambridge, Mass.: MIT Press.
- McCawley, James D. 1988a. Comments on Noam A. Chomsky, “Language and problems of knowledge.” In *University of Chicago working papers in linguistics 4*, 148–156. Department of Linguistics, University of Chicago, Chicago, Ill.
- McCawley, James D. 1988b. The comparative conditional construction in English, German, and Chinese. In *Proceedings of the Fourteenth Annual Meeting of the Berkeley Linguistics Society*, 176–187. Berkeley Linguistics Society, University of California, Berkeley.
- McCawley, James D. 1998. *The syntactic phenomena of English*. 2nd ed. Chicago: University of Chicago Press.
- Michaelis, Laura, and Knud Lambrecht. 1996. Toward a construction-based theory of language function: The case of nominal extraposition. *Language 72*:215–247.
- Pesetsky, David. 1989. Language-particular processes and the Earliness Principle. Ms., MIT, Cambridge, Mass.
- Pollard, Carl, and Ivan Sag. 1994. *Head-Driven Phrase Structure Grammar*. Chicago: University of Chicago Press.
- Ross, John Robert. 1967. Constraints on variables in syntax. Doctoral dissertation, MIT, Cambridge, Mass.
- Sag, Ivan. 1997. English relative clause constructions. *Journal of Linguistics 33*:431–483.
- Van Valin, Robert D., Jr., and Randy J. LaPolla. 1997. *Syntax: Structure, meaning, and function*. Cambridge: Cambridge University Press.

(Culicover)

Department of Linguistics and Center for Cognitive Science
The Ohio State University
208 Ohio Stadium East
1961 Tuttle Park Place
Columbus, Ohio 43210-1102
culicover.1@osu.edu

(Jackendoff)

Program in Linguistics and Cognitive Science
Volen Center
Brandeis University
Waltham, Massachusetts 02454
jackendoff@brandeis.edu