Correlational Comparison in English

Dissertation

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Elizabeth Allyn Smith

Graduate Program in Linguistics

The Ohio State University

2010

Dissertation Committee:

Carl Pollard, Advisor

Craige Roberts, Advisor

Robert Levine
Abstract

This dissertation proposes a novel analysis of the syntax and semantics of Comparative Correlative sentences in English such as *the bigger they are, the harder they fall* or *the faster we drive, the sooner we'll get there*. The analysis is cast in a framework that distinguishes between argument structure and word order, called Pheno-Tecto-Differentiated Categorial Grammar. The analysis fares better than past analyses at meeting the benchmarks set out by previous authors, and it also accounts for the new benchmarks presented here. These new benchmarks include a demonstration of the inherently proportional quantificational force in the *the* that begins a Comparative Correlative and new data showing that these kinds of sentences cannot appear with a full range of adverbs of quantification, unlike conditionals.
Dedication

Dedicated

to Albert and Marilyn Dorton
Acknowledgments

Despite the fact that talking to people who hold PhDs has become commonplace in my life, my thrill at the thought of earning one myself is undiminished. In achieving this goal, no one has been more helpful than the faculty of the Linguistics Department at The Ohio State University. My primary advisors, Carl Pollard and Craige Roberts have helped to shape me as a scholar in ways I'm sure I have yet to even notice. I am so thankful for the incredible amount of time you each have spent meeting with me over the years, and particularly during my dissertation writing and research. I have felt supported and cared for as a person as well as a student, which was exactly what I needed in order to succeed in this endeavor (as I suspect you very cleverly figured out).

I would also like to thank David Dowty for working with me closely up to and even after his retirement. His courses and guidance not only led to this dissertation topic but also honed my ear for interesting linguistic phenomena. Bob Levine and Peter Culicover also deserve special thanks for adding balance to my committee, preventing a myopic, categorial grammar-centered view of natural language grammars that I might otherwise not have reflected on. They have both provided valuable insights and suggestions that have informed the way I have approached many parts of this thesis. Bob, in particular, has helped me in a number of other ways, including everything from
writing an abstract to picking a good bottle of wine.

Though my thesis does not extend beyond the realm of syntax, semantics and pragmatics, my interests certainly do, thanks largely to Mary Beckman. From her I learned more than phonetics and laboratory phonology (though certainly those were the seeds for most of what came later); I learned to search corpora and become comfortable with the variability in real speech, what linguistics might be like if most work weren't done on English, and what interdisciplinary work could look like. From the beginning, Kathleen Currie Hall was my partner in crime (okay, fine, research) in following through on thinking outside the box, and our collaboration continues to be the ideal I strive for in all collaborative endeavors. I am consistently amazed at how well we complement one another, and I am thankful that our collaboration has led to such a deep friendship.

There are a number of other faculty who have helped me with this thesis whether they know it or not, including Dave Odden, who taught me how to write clearly about linguistics (though if I haven't accomplished that here, it's not his fault). Chris Brew, Beth Hume, and Brian Joseph have been encouraging since the beginning of my graduate career, always sparing time to talk or answer questions. I feel especially lucky to have been at Ohio State during a time when so many people joined the faculty, especially at earlier stages in their careers. Kathryn Campbell-Kibler, Cynthia Clopper, Judith Tonhauser and Mike White have patiently answered all of my questions about the 'next step' and the things that lead up to it: dissertating and job hunting. I sometimes feel that
they are the role-models that I am trying to emulate in a younger sibling-like way. All of these faculty have truly treated me like a junior colleague, which is what I am most thankful for.

I am also indebted to a number of other professors outside of Ohio State, especially those at the University of Massachusetts at Amherst where I was a visiting scholar and at Brown University where I was an undergraduate. Pauline Jacobson, Rajesh Bhatt, Barbara Partee, and Angelika Kratzer have all aided this dissertation in various ways, and I am additionally thankful to Mark Johnson for pushing me into linguistics. I also appreciate the mathematics education I received at Brown, especially the mentorship of Jeff Hoffstein and Jill Pipher. I am grateful to the National Science Foundation, the OSU Dean’s Distinguished University Fellowship, and the Royce Fellowship at Brown University for funding my growth as a scholar.

Graduate school has included many of the happiest years of my life, partially thanks to those mentioned above, but also thanks to my peers and other colleagues as well as friends and family. Through linguistics (broadly construed), I'm so glad to have gotten to know Patricia Amaral, Jan Anderson, Joanna Anderson, Leah Bateman, John Beavers, Anouschka Bergmann, Maria Biezma-Garrido, Steve Boxwell, Adriane Boyd, Katie Carmichael, Balder ten Cate, Wes Collins, Amy Rose Deal, Markus Dickinson, Hope Dawson, Michelle Dionisio, Robin Dodsworth, David Durian, Salvatore Florio, Jirka Hana, Jim Harmon, Jane Harper, Jesse Harris, Ilana Heintz, Jeff Holliday, DJ
Hovermale, Matt Husbands, Alison Duncan Kerr, Rachel Klippenstein, Eunjong Kong, Barak Krakauer, Ruth Kramer, Yusuke Kubota, Jungmee Lee, Lisa Levinson, Lia Mansfield, Scott Martin, Laurie Maynell, Julie McGory, Grant McGuire, Paula Menendez-Benito, Vanessa Metcalf, Jeff Mielke, Vedrana Mihalicek, Claudia Morettini, Deborah Morton, Cathy Muller, Ben Munson, Crystal Nakatsu, Marie Nilsenova, Julia Papke, Nicolai Pharao, Pat Reidy, Hannah Rohde, Mary Rose, Sharon Ross, Aynat Rubinstein, Anton Rytting, Osamu Sawada, Kevin Scharp, Florian Schwarz, Chung-chieh (Ken) Shan, Andrea Sims, Anastasia Smirnova, Stephanie Solt, Tamina Stephenson, Tom Stewart, Kristen Syrett, Anne Michelle Tessier, Giorgos Tserdanelis, Na'im Tyson, Laura Wagner, Pauline Welby, Abby Walker, Christin Wilson, Peggy Wong, Chris Worth, and Youri Zabbal.

I give thanks for friends and colleagues outside of linguistics who have supported me through the writing of this thesis and in various times and ways throughout graduate school, especially Leah Anderson, Amy Aspey, Ray Ball, Jackie Beam, Michael Beam, Amy Bishop, Gary, Carol and Jason Brand, Matt Brinkman, Philip Brooks, Patricia Cunningham Jr., Joseph Ferguson, Natalie Hamilton, Brian Hauser, Michael Heintz, Emily Hickman, Lisa, Kevin, Ella and Max Holtsberry, Brian Jones, Laura Jones, Joanna Kelley, Danielle Langfield, Chris Lindemann, Chris Manion, Colleen Ogle, Amanda Rosen, Liza Toher, Nicole VanderDoes, Tara and Valerie Van Ho, and Christina Xydias. In addition, the love of my family has made a big difference, including my great-
grandmother Minerva Halbert, and all four of my grandparents: Eugene and Suzette Smith and Albert and Marilyn Dorton. I love the way my grandfather Smith dotes on me and the pride in his voice when he introduces me to his friends. I loved my grandfather Dorton's frequent compliments and the way he was able to make everyone feel how much he cared for them every time he was with them. I loved my grandmother Dorton's way of keeping you honest and her clean soap smell. I loved the way she hugged you longer as she got older and always waved from the screened door until your car was out of sight.

I've been lucky to be so close to much of my family while in Columbus. I've enjoyed easy-going dinners and nights on the porch with Deanna and Jim Garrett. I've honed my euchre skills with Vicki and Brad Stebel. I've enjoyed overlapping at OSU with Sarah Rudolph and seeing Jimmy, Mary and Vic Rudolph, Teri O'Brien, Lori Hoyle, and Ron, Chuck, and Rick Sheets and their families at bonfires, holidays, and other celebrations. I have especially enjoyed my time with my great aunts and uncles, having lunch or just stopping by to chat with Dick Sheets and Jim and Virginia Rudolph (I always did love to visit 'aunt Ginny with the two doggies').

In addition, lots of other family has been close enough for visits or made the effort anyway, which has prevented graduate school from ever feeling isolating (well, aside from the actual thesis writing). I've been happy to see Cheri and Daniel Brunato, Rick, Jenny and Clayton Kinas, Jen and Josh Braver, Matt Garrett, Mel Davis, Ann, Eric, Andy, Amber and Logan Brown, and Colleen Jennings. I hope that I haven't inadvertently
forgotten anyone in these acknowledgements, and I apologize if I have. But, of course, no acknowledgement would be complete without thanks for my immediate family with whom I've been through thick and thin my whole life.

Thanks to my brother, Alexander Wemyss, for being the best little brother I could have asked for. Your childhood admiration inspired me to try to accomplish things that were worthy of admiring, and I hope that this is something you can be proud of. I'm incredibly proud of you. Thanks to my dad, Randolph Halbert Smith, for his sense of spatial reasoning and problem solving. I use them every day, so they're arguably the most important facet of my success in creating linguistic analyses. Thanks also for making me adaptable and instilling a sense of wanderlust from our early trips together. And finally, thanks to my mom, Pamela Dorton Smith. You're my rock, and I am in awe of how much you've done for me my whole life to get me to this point. I never suffered from a lack of time or attention, and I can't even imagine how much effort you put in (in reading to me, alone!). Most of all, I'm thankful that you always told me that I could do whatever I wanted with my life, and that's exactly what I've done.
Vita

2003 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . Sc.B. in Mathematics, Brown University

2003 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . A.B. in Linguistics, Brown University

2009 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . M.A. in Linguistics, The Ohio State University

Publications


Fields of Study

Major Field: Linguistics
# Table of Contents

Abstract..............................................................................................................................ii
Dedication.........................................................................................................................iii
Acknowledgments............................................................................................................iv
Vita.....................................................................................................................................x
List of Tables......................................................................................................................xiv

Chapter 1 : Introduction....................................................................................................1

1.1 An Historical Note......................................................................................................4
1.2 Constructions, Lexicalism and Compositionality.....................................................9
1.3 Dissertation Layout.....................................................................................................14

Chapter 2 : Survey of Syntactic Desiderata for an Analysis of CCs..............................15

2.1 Earliest Analyses of the Syntax of CCs.................................................................16
2.1.1 Ross 1967...........................................................................................................16
2.1.2 Thiersch 1982......................................................................................................19
2.1.3 Fillmore 1987, Fillmore, Kay, and O'Conner 1988.........................................24
2.1.4 McCawley 1988...............................................................................................29
2.2 Beck 1997...............................................................................................................38
2.3 Culicover and Jackendoff 1999..............................................................................42
2.3.1 Macrostructure..................................................................................................44
2.3.2 Structure of each CC clause...............................................................................52
2.3.3 Properties of the as a comparative specifier.....................................................57
2.3.4 Discussion..........................................................................................................59
3.4.2 Discussion.................................................................179
3.5 Conclusion.....................................................................183
3.6 Semantic Benchmarks..................................................184
Chapter 4 : Framework and Analysis........................................186
  4.1 Situating PTDCG with respect to other frameworks............186
  4.2 Pheno/Tecto Differentiated Categorial Grammar...............194
    4.2.1 PTDCG Analysis of Simple Sentences....................202
    4.2.2 PTDCG Analysis of Quantifier Scope Possibilities......204
    4.2.3 PTDCG Analysis of Comparatives and Parasitic Scope...209
    4.2.4 PTDCG Analysis of Unbounded Dependency..............226
  4.3 PTDCG Analysis of the CC and RCC.............................230
  4.4 Conclusion...................................................................254
Chapter 5 : Discussion and Conclusion....................................256
  5.1 Benchmarks Redux......................................................256
  5.2 Future Research.........................................................277
  5.3 General Conclusion...................................................288
Bibliography........................................................................290
List of Tables

Table 2.1: Extraction patterns of bi-clausal constructions........................................63
Table 3.1: Model for (24)..................................................................................................98
Table 3.2: Alternative Model for (24)................................................................................99
Table 3.3: Summary of entailment and polarity patterns across similar constructions...102
Table 3.4: CC search term “the more the”; Conditional search term “if the” (eliminating
“not always”, etc.) ........................................................................................................140
Chapter 1: Introduction

This dissertation is devoted to the description and analysis of Comparative Correlative (CC) sentences in English, such as those in (1)-(3).

(1) The bigger they are, the harder they fall.
(2) The more specific a feature description, the fewer feature structures satisfy it.
(3) The faster we get our groceries, the sooner we can get to the lake.

These examples all have the same the...,the... form, which is what marks them as being CCs. Chapters 2 and 3 are devoted to a detailed examination of the syntactic and semantic properties that CCs share, but this the...,the... template is the most basic identifier. Though CCs such as (1) or others like the more, the merrier are idioms in English, the examples in (2) and (3) demonstrate that this is a productive, largely non-idiomatic construction that spans all registers of speech from formal (example (2) is a quote from an academic lecture) to informal (example (3) appeared in a Wal-Mart commercial, clearly intended for a broad audience). CCs are often used to express natural relationships, such as the one between speed and time in (3), or other cause-effect relationships, as in the other examples. It is nevertheless possible to use a CC to merely note a correlation rather than a causation, as Beck 1997 noticed (discussed in chapter 3). This is the sense in which CCs are correlational, as the title of this thesis indicates. They are also comparative sentences due to the necessary presence of a marker of comparison such as more or -er in each clause.
Throughout this dissertation, I call the constituents headed by the two *the's* 'clauses', despite the fact that the examples in (4), below, could be viewed as clauses having a VP (etc.) elided, or alternatively as CCs formed directly from constituents other than sentences. I will leave these as a special topic within CCs and not treat them in detail in the dissertation, though I expect that the analysis presented here extends to these examples given appropriate theories of ellipsis and anaphora.

(4)  
   a. The nicer the man, the better the dog. (no verb phrase)  
   b. The bigger, the better. (no verb phrase or subject noun phrase)  
   c. The more, the merrier. (no verb phrase, subject NP or, in the first clause, adjective)

Even within the obviously clausal examples I will analyze, there is a lot of syntactic variation. The comparative morpheme occurs with a variety of (preposed) syntactic categories in the CC, as in (5).

(5)  
   a. The more books I read, the more facts I know. (noun phrase)  
   b. The older you are, the taller you are. (adjective)  
   c. The more quickly we run, the more likely we are to win the race. (adverb)  
   d. The more I read books, the more I like them. (sentence)  
   e. The more out of breath I am, the harder it will be to finish the race. (prepositional phrase)

There are also other kinds of sentences in English that express correlations, such as (6)-(9). The sentence in (6) is an example of a Reverse CC (RCC), which, along with the CC, is analyzed in this dissertation. The analyses of sentences like (7)-(9) is left for future work.

(6) We'll get to the lake sooner, the faster we get groceries.  
(7) As you eat more, you get fatter.
(8) As I build more muscle, I run faster and faster.
(9) If he steals more and more money, he's more and more likely to end up swimming with the fishes.

I propose that, though the need for a means of linguistically expressing correlated increases, which is what all of these sentence meanings have in common, may be an expressive necessity or universal, examining CCs cross-linguistically or across sentence types in English reveals the subtle ways in which these meanings differ and the ways in which the syntactic properties or other differences between languages/constructions affect those meaning differences. This is reminiscent of the approach Chomsky takes to passivization in *Lectures on Government and Binding* (1981):

The category that is commonly called "passive" may not constitute a natural class, either within or across languages. Syntactic passives are unlike lexical passives, and in languages as closely related to English as German, syntactic passives seem to behave differently and may involve a rather different rule structure. In other languages, what might be translated as passive in English has still different properties... In short, it is not obvious that the notion "passive" refers to a unitary phenomenon, still less one that can serve as a foundation stone or even guiding intuition for a theory of syntax. It may be a useful descriptive category, and one can imagine functional explanations for the prevalence of some such device. But the range of phenomena that fall within this category in some sense appear to be rather heterogeneous in character." (120-121)

By showing the non-uniform behavior of the CC as compared to the conditional, sentences like (7), and even the RCC, I similarly conclude that it is a purely descriptive category, not a syntactic universal.

I will provide explicit formal analyses of the CC and RCC based in a kind of Categorial Grammar (CG) called Pheno/Tecto-Distinguished CG (PTDCG). The semantic analysis follows the tradition of Montague (1973), though chapter 3 discusses
the relationship between this and a dynamic analysis. What follows are my theoretical desiderata used in constructing the analyses in this dissertation. I prefer analyses that:

- Make clear, testable predictions (i.e. are falsifiable)
- Do not make the assumption that all languages are structurally similar (or lexically parallel: the CG equivalent)
- Are based in rigorous formal foundations (such as logic)
- Are as compositional as possible, in the sense of Dowty (2007), including preferences for a simple syntax-semantics interface with no mediating grammatical levels, a one-to-one correspondence between words and meanings where sentences' meanings are built from the combination of individual word/morpheme meanings, and rejections of empty categories and reordering, to the extent possible ('surface faithfulness').

I succeed in most of these goals, though PTDCG does not respect surface faithfulness in that the order in which words and phrases are combined does not reflect their order in a given sentence. The analyses show that English CCs and RCCs are epiphenomena of unique combinations of natural language devices (comparison, generalized quantification, anaphora, long-distance dependencies, etc.) independently motivated within English that, in concert, achieve a particular class of correlational interpretations.

1.1 An Historical Note

It is a curious fact about English CCs that the word that introduces each of its main constituents is homophonous with the definite article in English. Cross-linguistically, many languages use a relative pronoun like how much to introduce the first clause and a demonstrative pronoun like that much to introduce the second clause in a CC-like sentence, i.e. one with a correlational meaning (den Dikken 2005). Still others use a subordinating morpheme that is found in conditional sentences of those languages. Though I leave the analysis of other languages' CCs for future research, I mention this
fact here because it parallels what the *Oxford English Dictionary* says about the *the's* that occur in the CC (p. 260):

**The**, adv. [OE *pe*, originally locative or instrumental case of the demonstrative and relative pron. *se, seo; paet*. In OE. interchanging with *py*: see THY adv.]

1. [...]  
2. *The...the...*: by how much... by so much; in what degree... in that degree...: denoting proportional dependence between the notions expressed by the two clauses, each having *the* + a comparative; one *the* being demonstrative, and the other relative.

This is significantly different from the lexical entry for the definite article *the* in English, which the *OED* says was originally the nominative case of *se/seo* as compared to this adverbial *the*, which originated in the locative or instrumental cases of that demonstrative. This diachronic information can inform the synchronic analysis given here insofar as it supports the idea that the *the's* we find in the CC do not need to be unified with the definite article. Of course, this does not preclude their semantic definiteness as that is not limited to the definite article, but it does help to explain why their syntactic distribution is so different from the definite article *the*.

Interestingly, the *OED* entry does not indicate that the two *the's* in the CC were derived from different sources as compared to one another, but its conclusion that one is a relative pronoun and the other a demonstrative is clear. Jespersen 1949, 1961 and Quirk et al. 1972 say a little more. Jespersen states that the first *the* derives from the instrumental form of the Old English *that* but that the second *the* derives from a relative pronoun. Den Dikken 2005 contains more discussion of the historic properties of *the*, but the point here is merely that these merit lexical entries that are independent of one another and of the English definite article.  

5
There is a similarity between the first the in the CC (introducing the subordinate clause, henceforth the1), and the second the in the CC (introducing the matrix clause, henceforth the2) in that the OED considers both 'adverbial' the's. The OED then goes on to illustrate other uses of adverbial the outside the CC: pronouns. By labeling the1 a relative and the2 a demonstrative, the dictionary suggests a parallel between these examples and the correlatives we saw above, but it also explicitly draws a parallel between the2 and the other instances of adverbial the cited for English:¹

(10) Your fav’rite horse Will never look one hair the worse. -Cowper
(11) And if others do not follow their example,—the more fools they. -Ruskin
(12) If you sow them now, they will come up the sooner.
(13) He has had his holiday, and looks the better.

While the example in (13) could just as easily be part of a contemporary English conversation as the 18th and 19th century English they were drawn from, most of the examples in (10)-(13) are markedly from an older era. However, some have survived through popular stories, as in (14), and they continue to be spoken and written, if infrequently, at least productively, as shown in (15) (a created example) and (16)-(18), which I have heard or read in a book published in the past three years.

(14) A: Grandmother, what big eyes you have!
    B: The better to see you with, my dear.
(15) Mary is coming? All the better.
(16) ...critics hearts, who beat the faster knowing...
(17) All of us would be the poorer without their devout and disciplined work.
(18) So that...his pupil might learn the better from him, shape himself more finely in his uncle’s image. (The Emperor’s Children by Claire Messud, p.194)

¹ I do not posit that these the's are identical however, due to possibilities such as the hypothesis that the2 is a bound anaphor in the CC, while adverbial the could be discourse anaphoric. A detailed comparison of these items is beyond the scope of this study and will be studied in future work.
I leave the analysis of these sentences for future work, but ideally, one of the adverbial the's in the CC could be used in the composition of such examples.

Turning now to a bit of history of the analysis of the CC, CCs were first noted by Ross 1967 and named *Comparative Conditionals* by McCawley 1988 who noted many similarities and a few dissimilarities between CCs and their simple conditional counterparts (i.e. *If they are bigger, then they fall harder*). A formal semantic analysis of CCs later appeared in Beck 1997, who also considered them to be conditional sentences and used McCawley's term to describe them. CCs were re-named *Comparative Correlatives* by Culicover and Jackendoff 1999, who begat a storm of syntactic interest when they claimed that these sentences were on the ‘periphery’ of grammar and could not be accounted for within Minimalism or any other contemporary syntactic framework without the addition of ad hoc rules or lexical entries. Responses to C&J by Borsley 2004a,b for HPSG and DenDikken 2005 for Minimalism dispute this claim.

It seems that Culicover and Jackendoff traded *conditional* for *correlative* in the name of the construction on the basis of their argument that neither of the CC’s two main constituents is *syntactically* subordinate to the other, unlike the conditional. The newer term *correlative* was chosen to express a syntactic similarity between CCs and *co-relative* (also *corelative, correlative*) clauses (Downing, 1973). Co-relative clauses (a class of relative clauses), are productive in many languages, including Bosnian/Croatian/Serbian:
(19) \[Kamo \ Sup \ [tamo \ novinari, \ policija, \ ljudi \ iz \ [where \ vulture-nomsg][there \ reporter-nompl \ police-nomsg \ people-nompl \ from \ Helsinki-gensg] \]

Lit: ‘Where [there is] a vulture, there [there are] reporters, police, people from Helsinki.’
‘There are reporters, police & people from Helsinki where there are vultures.’

Here, the first clause kamo sup is the correlative clause, which is linked to the demonstrative tamo in the main clause. Correlatives are typically possible with the full range of wh-words, and when the where and there of (19) is replaced with how much and that much, correlative comparatives result, as in the Hindi CC in (20) below. Thus, CCs are a subspecies of co-relative in many languages.

(20) \[Jiitnii \ der \ ho-tii \ gayii \ ][utnii(-hii) \ ThanD \ baRh-tii \ gayii \ ].
[how-much \ late \ be-HAB \ go-PF][that-much(-only) \ cold \ increase-HAB \ go-PF].

Lit: ‘How much late it becomes, that much the cold increases’
‘The later it got, the colder it became.’

Unfortunately, co-relative is now written as correlative (cf. Srivastav, 1991), which leads to confusion in this case because one might think that correlative meant 'having a correlational meaning' or 'having the same syntactic structure as a co-relative clause.'

Because it is consistently used in the CC literature to mean the latter, I will continue to use the term correlative when talking about this kind of syntactic clause. I remain agnostic about whether correlative clauses have the same syntactic structure as CCs, which is why I prefer the term Correlational Comparative to Comparative Correlative. Since both have the same CC acronym, this won't make much of a difference.
English historically had correlative clauses, but they have disappeared through time. Some headless relative adverbials in English have the older correlative sense, however, as in the locative adverbial in (21) and the instrumental adverbial in (22).

(21) Wherever Mary goes, John goes (there).
(22) However John cuts the cake, Mary cuts it that way, too.

Additionally, some authors have argued that conditionals are also syntactically correlative structures (Bhatt and Pancheva 2006, inter alia), obfuscating the need for the original change from Comparative Conditional to Comparative Correlative, but I will not discuss that connection further in this thesis. In fact, I will show that regardless of their similarity or difference in syntactic structure, CCs are semantically different from conditionals and other kinds of correlatives such as the headless relative adverbials above. In chapter 3, I argue that the relationship between what is denoted by the correlative and retrieved by the demonstrative in a regular correlative sentence is one of co-reference and not correlation of increases as in the CC. I also show that, contra past proposals, quantification in the CC does not exhibit the same pattern as quantification in the conditional. The CC's unique pattern is then argued to come from its unique lexical item: the1.

1.2 Constructions, Lexicalism and Compositionality

Fillmore et al. 1988 explicitly state that they do not believe that the CC exhibits properties that are fully predictable from independently known properties of its lexical makeup and its grammatical structure. This dissertation shows that even though it will require special lexical items for the (motivated by the fact that they aren't historically the
same *the* as the definite article anyway) and a special syntactic category *CC* for the type of the *the* phrase, it is possible to give an analysis of the CC that is fully predictable from independently known properties, lexically and grammatically. The real question is: how ad hoc are those rules and lexical items? I hope to show that neither are ad hoc.

Compositionalist analyses are extremely difficult to characterize, but one thing that they all presuppose is that syntactic structures and semantic interpretations in natural language are compositional and that analyses should reflect that fact. Dowty (2007) tries to get at the crux of the problem, saying that the question is not *whether* language is compositional, but *how* it is compositional. Dowty shows that Frege's original definition of compositionality is vague (assuming unified notions of 'syntactic mode of combination', 'meaning', etc. where none exist), and, in its weakest form, incontrovertibly true given the recursive possibility of language to generate an infinite number of sentences. Dowty proposes that we redefine natural language compositionality as “whatever strategies and principles natural language(s) actually do employ to construct the meaning of sentences (and other complex expressions), on the basis of whatever syntactic properties and contextual information they do in fact turn out to depend on” (Dowty, section 2). Though I do not go into all the relevant issues here, let us assume that the most rigorous hypothesis about compositionality is that:

1. All syntactic constructions having inputs of the same syntactic categories are interpreted by the same compositional semantic rule (Dowty, section 3.6).
2. There is a one-to-one correspondence between syntactic rules and semantic interpretation rules.

This pair of requirements is too restrictive for describing natural language on its face.
We need at least some way to account for wide scope of in situ operators, such as Cooper Storage or Quantifying In, but any way of doing that violates #2. As for #1, subscribing to this rule was a hallmark of Montague Grammar, where names, bound pronouns and quantified NPs were interpreted in the same way via the same rules; the same was true of extensional/intensional predicates (Montague 1973). Nevertheless, the value of that aspect of Montague's approach is disputed in contemporary semantic theory, which I will not discuss in any detail here other than to say that I am sympathetic with the Montogovian approach. Despite the potential need to violate these two principles to account for a full range of natural language structures and interpretations, they are a good starting place for stating explicitly what the minimum set of additions are that are needed to capture natural language's syntax-semantics interface. Given that my goal is to create analyses of natural language phenomena using the minimal technology necessary for the treatment of natural language as a whole (which I am here doing for the CC), this technique will serve as a means of comparing which systems are more economical than others.

Related to economy, Dowty has additional principles of 'transparency' and 'economy' that spell out the efficiency and minimality of syntax/semantics/their interface, but the idea behind these (in combination with the other desiderata mentioned above) is that analyses should try to maximize economy of the syntactic system, economy of the semantic system, and economy of the syntactic-semantic interface simultaneously. Often, syntax-semantic analyses economize one at the expense of another or two others. In some ways, this is inevitable, but the challenge is to balance the complexity of the three rather than having one be economical and the others extremely complex. This
relates to the stated goal of achieving, under the name 'compositionality', the most economical/restrictive system possible with respect to these three components of linguistic theories. The theory presented here is maximally economical in terms of having only two rules, each of which is a one-to-one mapping between syntactic and semantic components. The economy of its lexicon is less clear. For example, *the1* in the CC does not have the same lexical entry as the *the* in the RCC. In terms of transparency, the more 'levels' of syntax and semantics that are needed in a given theory, the less transparent it becomes. My theory has a semantic component and two syntactic components, one of which also handles some aspects of sound structure. This is less transparent than a so-called 'directly compositional' system (Jacobson 2002) where there is a single syntactic component and the same simple interface between it and the semantics that my framework shares, but it is more transparent than a system that crucially involves 3 or more syntactic levels such as Deep Structure (DS), Surface Structure (SS), Phonological Form (PF) and Logical Form (LF) all working together along with the actual semantic interpretation to achieve the analysis of a single sentence.

As alluded to above, my analysis is a lexicalist analysis. *Lexicalism* or *Radical Lexicalism* (Karttunen 1989) names an analysis or general approach in which every part of the sentence-level (or multi-utterance) interpretation is contributed by (a part of) the meaning of a lexical item. In other words, to quote Dowty (2007), “all 'construction-specific' compositional meanings...must be analyzed as packed into the meaning of some lexical item(s) in the construction: there can be no 'constructional meaning' specific to an individual construction” (section 3.7). For example, in Type Logical Grammar (TLG), functional application and functional abstraction are the only compositional possibilities
for interpretation, as they are the semantic counterparts to the syntactic rules of slash elimination and slash introduction, respectively (Carpenter 1997). Thus, any meaning besides function application and abstraction must be the semantic contribution of a lexical entry. This is what makes analyses within TLG lexicalist. As is clear from Dowty's description of lexicalism, the alternative extreme would be the strongest possible form of constructionalism in which all sentence-level interpretations were given by construction-specific rules. This kind of approach is so strong a violation of compositionality that it is not what is promoted under the title of contemporary 'constructionalist' frameworks such as Construction Grammar. Instead, Construction Grammar differs from TLG mostly in terms of the locus of an arbitrary meaning for a phrase or clause; in TLG, that meaning would be attached to the lexical specification of some aspect of the sound string that occurs in all and only those kinds of phrase or clauses, while in Construction Grammar, it would be contributed to a clausal or phrasal frame, which is like a lexical item but for bigger chunks of language. Then, instead of combining words together with a few rules as in TLG, you would begin with a sentence frame and use lexical items to fill in the blanks in that frame to arrive at a sentential meaning. One could similarly compare Construction Grammar to other frameworks such as HPSG, noting that constructions are basically just more complex phrase structure rules. My dissertation won't have anything to say about which of these is a more useful way to approach natural language, but it will undercut Fillmore's argument that the CC is evidence for one position (the constructional position) over the other. He says:

...when constructions are interpreted as the products of maximally general rules, no place remains in the grammar for spelling out the non-predictable
semantics and pragmatics that is frequently conventionally associated with particular constructions such as those we will describe [the the X-er the Y-er construction, or the let alone construction] (Fillmore et al. 1988: 503)

I show in this dissertation that all of the supposedly non-predictable aspects of the semantics and pragmatics of the CC that he is alluding to are predictable. I then develop an analysis that predicts them within a system that is both lexicalist and on the 'very compositional' end of compositionality.

1.3. Dissertation Layout

This dissertation is structured as follows. Chapters 2 and 3 both include a presentation of data and a discussion of past work on the CC. Chapter 2 focuses on the syntactic patterns of the CC and past syntactic analyses, while chapter 3 focuses on the semantic patterns and analyses. Throughout, empirical patterns that need to be captured in an analysis of CCs are set apart as benchmarks and repeated at the end of each of these chapters. Chapter 4 then develops both the framework (PTDCG) and the analysis of CCs and RCCs within that framework as well as examples of other kinds of sentences in natural language that bear on the analysis of CCs, such as sentences with unbounded dependencies, comparatives, etc. Chapter 5 then evaluates the analysis of chapter 4 with respect to the benchmarks set out in chapters 2 and 3 as well as discussing future research.
Chapter 2: Survey of Syntactic Desiderata for an Analysis of CCs

In this chapter, I animate the debate about the syntactic structure both of individual CC clauses and their relationship to one another. Anytime a set of data is presented that shows a pattern that any theory would need to capture, I introduce this pattern as a numbered benchmark (B1), (B2), etc. so that they stand out from regular example numbers. The chapter concludes with a list of these benchmarks repeated together in one location. Almost all of these benchmarks are culled from previous literature on the syntax of the CC, which I move through chronologically in this chapter. The fact that most authors have incorporated desiderata from work before their own influenced my choice to go chronologically in that I can more easily discuss which past benchmarks a theory meets if they have already been introduced. Though this chapter is devoted to syntax, there are many cases a particular pattern claimed to have a syntactic explanation by one author is claimed to be semantic or pragmatic by another. Nevertheless, all patterns that bear on the syntactic representation of CCs and their subparts will be discussed here, which includes these cases insofar as eliminating certain benchmarks from the set of things that the syntax needs to account for could lead to different syntactic analyses.
2.1 Earliest Analyses of the Syntax of CCs

Though earlier works by descriptive linguists such as Jespersen 1949, 1961 and Quirk et al. 1972 include mention of the CC and some information about its properties (discussed as relevant below), I review here only those works that point in the direction of some particular theory or other of CCs. Thus, the earliest works on the syntax of the CC for the purposes of this thesis are Ross 1967, Thiersch 1982, Fillmore 1987, Fillmore, Kay, and O'Connor 1988, and McCawley 1988, spanning both Generative grammar and Construction grammar traditions.

2.1.1 Ross 1967

Comparative Correlatives were first noted in the syntactic literature by Ross 1967:406-407, though he says little about them. His reason for mentioning them is that he is interested in natural language phenomena whose grammatical rules potentially require the use of a variable in the syntax. He says that regardless of what rules the CC in (1) requires, sentences of this kind are sure to require variables. Specifically, he assumes that CCs contain long-distance dependencies, which are the focus of Ross' thesis. An example of a sentence with a long-distance dependency appears in (2).

(1) The more you read, the more you'll know.
(2) What book did John buy?

The constituent *what book* in (2) depends upon the verb *buy* (thus the 'dependency'), but rather than appearing in the object position (in which case it would be adjacent to the verb, i.e. *local*), it appears at the front of the sentence, i.e. *long-distance*. The displaced
constituent is the *filler*, and the spot in which it would be locally interpreted is the *gap*.

Though (2) is an example of a dependency in a single clause, dependencies of this kind are not limited to one clause, as (3) demonstrates (since *Kim wanted to date* is an embedded clause, while *woman* is in the matrix clause), which is the reason that these sentences are often called *Unbounded Dependency Constructions* (UDCs).

(3) Which woman, did you think that Kim wanted to date $t_i$?

The following is a list of some well-known UDCs (though they are exhibited here with single-clause examples), to which I add CCs, as in the example in (5) with the dependencies marked.

(4) a. Kim$_1$, Sandy loves $t_i$. (Topicalization)
   b. Who$_i$ does Sandy love $t_i$? (Wh-Question)
   c. This is the politician who$_i$ Sandy loves $t_i$ (Wh-Relative Clause)
   d. It’s Kim who$_i$ Sandy loves $t_i$. (It-cleft)
   e. What$_i$ Kim loves $t_i$ is Sandy. (Pseudocleft)
   f. I bought it O$_i$ [for Sandy to eat $t_i$]. (Purpose Infinitive)
   g. Sandy is hard O$_i$ [to love $t_i$]. (Tough Movement)
   h. This is the politician O$_i$ [Sandy loves $t_i$]. (Non-WH-Relative Clause)

(5) The faster, we drive $t_i$, the sooner, we'll get there $t_j$.

One of the ways to identify a UDC (and one of the properties that makes them a natural class) is that they (ostensibly) all obey some of the same syntactic restrictions. One such restriction is that filler-gap constructions with the gap inside a complex NP are ungrammatical, where a complex NP is one containing at least a noun and a modifying

---

2 From this point on, I will sometimes mark the gap site with a trace $t$ marked with an index to show what its filler is in example sentences, but this should not be taken as an indication that these indices are part of the ontology of the grammar.

3 UDC examples due to Pollard and Sag (1994); 'O' is an inaudible operator binding a trace.
clause (Ross 1967) For example, in (6a), the NP is not complex and the sentence is grammatical, whereas in (6b), the NP *people who like what beverage* is complex, leading to the ungrammaticality of the sentence. In (7), we see the same pattern with the CC.

(6) a. What beverage, can't Leslie drink t_i?
   b. *What beverage, can't Leslie stand people who like t_i?

(7) a. The more cookies, we eat t_i, the more weight, we gain t_i.
   b. *The more cookies, we hate people who like t_i, the less funny we find Cookie Monster.
   c. *The more cookies, we eat t_i, the more weight, we hate people who lose t_i.
   d. *The more cookies, we hate people who like t_i, the more weight, we hate people who gain t_i.

Further, as expected given the dependencies marked in both clauses of the CC in (5) and the constraint violations in (7b-d), either clause's long distance dependency can be unbounded in the CC, as further illustrated in (8).

(8) a. The later, Rob believes Tim expects the party to start t_i, the earlier, he will arrive t_j.
   b. The later, the party starts t_i, the earlier, Rob will try to persuade Tim to arrive t_j.

Thus, we have some evidence that the CC is a UDC. Though Ross doesn't explicitly make this argument himself, he suggests it, and he also outlines many of the other constraints that define UDCs as a class. For further evidence that CCs pattern like UDCs with respect to other constraints, see Culicover and Jackendoff (1999, 2005). This leads to our first benchmark, (B1).

---

4 Of course, like most 'island constraints', the Complex NP Constraint has known violations, such as (i), but we will consider the reasons for constraint violations in future sections.

(i) Which drug, did you hear rumors that they sold t_i?
(B1) *CC as UDC*: Each CC clause contains an unbounded dependency and is therefore subject to island constraints such as the Complex Noun Phrase Constraint, the Empty Category Principle, and the Strong and Weak *wh*-island constraints.\(^5\)

There are certain exceptions where CCs are thought not to pattern with other UDCs, but we will return to those in section 2.2.

### 2.1.2 Thiersch 1982

Thiersch 1982 characterizes the CC, which he calls the 'double comparative,' as having the pattern: \( S \rightarrow R + R \) where \( R \rightarrow \text{The} + \{\text{Adj, Adv}\} -er + \{\emptyset, \text{NP, S}\}.\(^6\) This points to several important benchmarks. First, that CCs are sentences, and second, that a *the +-*er marked constituent cannot occur in isolation as a sentence. Example (9) shows that CCs, like other sentences, can appear as complements of verbs that select for sentential complements and that a subordinate clause (as a non-sentence example) cannot.

(9) a. Robin said that Leslie went to the bank.
   b. *Robin said that if Leslie went to the bank.
   c. Robin said that the more money Leslie puts in the bank, the more she'll have later.

(B2) *CC as S*: CCs are of the same syntactic category as other sentences of English.

(10)*The more she runs.

(11)

---

5 Again, this is only true insofar as any of these island constraints or other principles are correct generalizations about natural language. Nothing about the account developed in this thesis actually predicts the facts that these are supposed to account for.

6 Note that this is not an accurate characterization of the syntactic pattern of CCs. Even assuming that he takes *-er* to be interchangeable with *more* and *less*, and that he can account for how *more* and *less* end up preceding the Adj/Adv, there are still a couple of problems with this description. First, it is possible for what follows the *-er* to be something other than an NP or S (as shown in Ch. 1). It could be a PP, for example, as in *The further up the ladder, the closer to the top*. Second, it is not necessary for each clause to contain an adjective or adverb in addition to the comparative morpheme. In one of the best-known CCs, *the more, the merrier*, the first *the* is followed only by *more*, showing the possibility of neither adjective nor adverb.
CC clauses don't stand alone: whenever one the more constituent occurs, a second the more constituent must follow to form a sentence.

Third, the characterization shows that each clause in the CC must begin with the, as the ungrammaticality of (11) verifies, and that no other determiner may be substituted, as (12) shows.\(^7\),\(^8\)

(11) a. *The more I read, I know more.
    b. *The more I read, more I know.

(12) *A more I read, a more I know.

(B4) Mandatory The: The is required phrase-initially in both main CC constituents.

Finally, these two rules presume that both clauses must have the comparative fronted, making (13) infelicitous.

(13) a. *The I read more, the more I know.
    b. *The more I read, the I know more.

(B5) Mandatory Fronting: the comparative (and whatever part of the filler it brings with it) cannot appear in situ in either CC main constituent.

I will discuss the semantics that Thiersch attributes to the CC in the next chapter, but one point about his meaning for the CC is crucial here as well. He draws a parallel between the meaning of a simple comparative sentence and the meaning of the CC. In the former, the sentence-level meaning is that one amount exceeds another. In the latter, according to Thiersch, the sentence-level meaning is that one scale is proportional to another. Thus,

\(^7\) I have only included examples showing that if the first constituent starts with the, the second must. That is because the RCC, obligatorily does not have a the in its first clause. I return to this in later benchmarks.

\(^8\) Thiersch is the first in the literature to cite historical evidence (from Jespersen 1949) that the in the CC is not a Determiner and is not the same as the definite article in English.
the thrust of his paper is to try to extend an analysis of comparatives (primarily the analyses of Bresnan 1973 and Chomsky 1977) to analyze the CC. He concludes that this is more arduous than it first appeared.

Thiersch wants to show that each CC clause is a sentence with a fronted element that is a result of *wh*-movement. These sentences are then conjoined to create the CC. Following Chomsky's 1977 re-casting of Bresnan 1973, he assumes that comparatives also involve *wh*-movement. This means that in a sentence like *John is taller than Bill*, *than Bill* is essentially recast as a *wh*-relative, namely *than how tall Bill is*. Some evidence that this is potentially on the right track would be the adherence of the gaps in the CC to the Complex Noun Phrase Constraint (among others), as demonstrated in 2.1.1.

The problem, then, comes in the difference between simple comparative sentences and the CC with respect to their similarities with coordinate and subordinate structures. First, Thiersch notes similarities between the CC and other subordinate-matrix relationships with respect to three main criteria. The first relates to the difference between the CC and the RCC, which he is the first to note. Specifically, whereas *the* is obligatory in both clauses of the CC, it is obligatorily absent in the first clause of the RCC, as in (14). Next, the grammaticality of backward pronominalization in the CC matches that of other subordinate-matrix constructions, as in (15). Finally, assuming, as Thiersch does, that gapping applies between coordinate but not subordinate structures, CCs pattern with the subordinate structures in not permitting gapping, as in (16) (Thiersch's (56) and (61)).
(14) a. I know more, the more I read.
   b. *The I know more, the more I read.
   c. I'll cut my hair, if the price is right.
   d. *Then I'll cut my hair, if the price is right.

(B6) *No the in the 1st clause of the RCC: In the RCC, no the can appear initially in the first main constituent.

(15) a. The longer John, has to wait, the angrier he, gets.
   b. #He, gets angrier, the longer John, has to wait.
   c. #If John, has to wait longer, he, gets angrier.
   d. *He, gets angrier, if John, has to wait longer.

(B7) Backward Pronominalization: Pronominalization is possible in either direction in the CC (between clauses).

(16) a. *Fred will feed the dog, if Helen __ the cat.
   b. *John sold more books to Helen, the less (books) Fred __ to Eileen.

(B8) Subordinate Gapping: The CC, like other subordinate-matrix structures, does not allow gapping between its clauses.

Given that Thiersch is giving a parallel analysis to sentences like I play basketball more than I play baseball and I play basketball more, the more I play baseball, he needs for comparative sentences to share these characteristics. The only one Thiersch discusses is the gapping, but it is already damning enough by itself (and the reader is can verify that it is also the case that backward pronominalization is not possible with the comparative examples). The examples in (17) (his (55) and (60)) demonstrate the difference in grammaticality in comparison to (16).

(17) a. Fred fed the dog, and Helen __ the cat.
   b. John sold more books to Helen than Fred __ to Eileen.

Example (17a) shows that coordinate structures generally allow gapping across clauses, unlike the example of the subordinate clause in (16a), and (17b) shows that comparative
sentences pattern like coordinate structures despite the fact that CCs pattern like subordinate structures. Thiersch tries to salvage his parallel analysis by introducing ideas from Williams 1977. These are specific to issues of the revised Extended Standard Theory and include limiting empty nodes to lexical (non-phrasal) nodes and having empty nodes appearing in the surface form interpreted via a level of 'Logical Form' (LF) rather than having those nodes be lexically filled at some step of the derivation and then later deleted by another syntactic process. I don't go into detail about this analysis or how Thiersch implements it because, though it does predict differences in terms of gapping and backward pronominalization between CCs and comparatives, it fails to account for a single CC sentence by not meeting benchmarks (B4: Mandatory the) and (B5: Mandatory Fronting). In other words, his syntactic theory overgenerates CCs without the in their matrix clause and without fronting in that clause.

A further problem for the parallel analysis of comparative sentences and CCs is that comparatives aren't like subordinate-matrix clauses in that the clauses can't be reversed with the morpheme that identifies the second clause (in this case, than) omitted. This is clear from (18) ((18a) is Thiersch's (51), and both it and (18b) are intended as versions of John ate more cookies than Mary did). As we saw above in (14) and (B6: No the in the 1st clause of the RCC), the morpheme that introduces the matrix clause (such as then or the) is necessarily omitted in the reversed versions of the CC and conditional, signaling another dissimilarity between the CC (and conditional) on one hand and the comparative sentence on the other.

(18) a. *Than Mary did, John ate more cookies.
    b. *Mary did, John ate more cookies.
In summary, what we have seen so far in this chapter is that Thiersch fleshes out and gives evidence for the earlier idea of Ross 1967 that each CC clause involves a long-distance \(wh\)-type dependency. However, his theory in which CCs parallel regular comparative sentences cannot capture several empirical generalizations about the CC.

### 2.1.3 Fillmore 1987, Fillmore, Kay, and O'Conner 1988

Whereas Thiersch's aim was to try to fit the analysis of the CC into the analysis of another type of English sentence, Fillmore 1987 and Fillmore et al. 1988 have a different theoretical axe to grind. They want to show that, as Fillmore 1987 puts it, “any successful description of these [CC] phenomena require us to assemble for this single construction an organized body of facts which cannot simply be taken for granted as following from other facts independently knowable from the grammar of English” (1987:166). In other words, Fillmore predicts that any attempt such as Thiersch's will inevitably fail on the basis that it does not recognize the sui generis ("constructional") nature of the CC.

Fillmore 1987 calls the CC the BCHF construction ("Bigger you Come, Harder you Fall"). He calls the two constituent halves of the CC the L and R constituents depending on whether they are to the left or right. L and R are then subdivided into X and Y, where X is the constituent with the compared phrase and Y is the constituent with the gapped clause, for each. Depending upon the syntactic category of the X phrase (excluding its the), let's call it A, the category of the Y phrase (excluding its the) will be S/A, indicating that the category of the gap matches the category of the fronted constituent (and that if
the gapped phrase had had its filler in situ, it would be a sentence). In Fillmore's dialect, *that* can appear only in L (his (5)-(6)):

(19) a. The bigger that it gets, the heavier it becomes.
   b. *The bigger it gets, the heavier that it becomes.

Authors since Fillmore have assumed that *that* is possible in either clause, as will I. The example in (20) illustrates this pattern. I included (20c) to show that examples with multiple clauses containing *that* are felicitous and grammatical and may even sound more natural to some speakers than examples with an asymmetry in *that* inclusion.

(20) a. The wiser that the people around us are, the more we will learn.
   b. The wiser the people around us are, the more that we will learn.
   c. The wiser that the people around us are, the more that we will learn.

(B9) That in the CC: *That* can optionally appear before the gapped clause in a CC clause.

The examples in (19) and (20) illustrate *that* in CC clauses with adjectival, adverbial or comparative gaps, but *that* can appear before any kind of gapped constituent that would be grammatical without it. Additional examples appear in (21), each with a *that* in L.

(21) a. The more cookies that she eats, the more weight she'll gain.
   b. The more people that arrive, the louder it gets.

---

9 Google searches lend credence to both perspectives in that, in CC examples with *that*, *that* is much more likely to occur in L than R (so Fillmore's judgment represents a trend), but there are certainly plenty of naturally-occurring examples where *that* occurs in R as well, such as (i), supporting the resulting generalization in (B9) above.

(i) The older you get, the more that coffee helps your brain.
   (http://scienceblogs.com/grrlscientist/2007/08/javajive_the_older_you_get_the.php)
This last example, (21b), is particularly interesting because though (B9) says that *that* appears optionally, Culicover and Jackendoff 1999 note that in certain cases, namely when a CC clause's gapped clause contains a subject gap, some speakers such as Culicover (but not Jackendoff) require a *that*. For the remaining speakers, these are exactly the cases in which an ambiguity of interpretation arises in the CC without a *that* to disambiguate. For example, for Jackendoff and me, (22a) is ambiguous between (22b) and (22c) (and for this reason, I will leave (B9) unaltered).

(22) a. The more people arrive, the louder it will get.
   b. The more that people arrive, the louder it will get.
   c. The more people that arrive, the louder it will get.

Fillmore notes another behavior of R (but not L), viz. the ability to invert the subject and auxiliary of the gapped sentence, as in (23) (his (8)). He notes that this is an obsolete form in contemporary English but includes it to show an historical difference.

(23) The older she got, the less likely was she to agree with me on this question.

Though Culicover & Jackendoff, Borsley and others return to this fact in their analyses of the CC, I will not include it as a benchmark here since it is, to my knowledge, no longer productive in contemporary English. In any case, it would not be a problem for my analysis in Chapter 4, which could be modified to take inverted complements in order to model the way language was spoken at that time.

These are all of the CC's syntactic patterns first noticed by Fillmore 1987. As stated at the beginning of this section, his goal is to show that there are idiosyncrasies that are part of the CC that preclude a compositional account with the same set of rules.
used for other sentences of English. The result is that he does not propose his own theory of CCs, only that one in the generative grammar tradition would be impossible. Much of the article is devoted to describing particularities of the conditional construction and drawing a partial parallel between the conditional and the CC. He never returns to analyze the CC, but one could presumably use the conditional construction as a model for the analysis of the CC construction within a constructionist framework. However, in this case, that would not be wise given that Fillmore ends up detailing several counterexamples to his analysis of conditionals that were brought to his attention by members of the audience during his talk (Fillmore 1987 is a conference proceedings paper). These counterexamples largely have to do with semantic phenomena such as particular readings of conditionals due to tense and modality. Nevertheless, Fillmore makes two important contributions to the literature on the CC (in addition to noticing some new data patterns). He is the first to suggest analyzing conditionals and CCs along the same lines (Thiersch noted some similarities between them but clearly wanted to analyze the CC like the comparative rather than the conditional). He is also the first to assert that the CC cannot be given an analysis using the same grammatical rules as other constructions or sentence types.

If Fillmore 1987 was primarily an article about conditionals, Fillmore, Kay, and O’Connor 1988 is primarily an article about idioms. They mention the CC in illustrating a distinction they are trying to draw between substantive or lexically filled idioms on the one hand and formal or lexically open idioms on the other. Those that are substantive cannot have other words substituted without losing the entire idiomatic meaning, such as kick the bucket, where substituting a different verb or noun will lose the denotation of
death. They define formal idioms as those “syntactic patterns dedicated to semantic and pragmatic purposes not knowable from their form alone” (1988:505). Fillmore et al. consider there to be a cline between idioms that are lexically closed v. open, and they use the CC as an illustration of how a formal idiom can facilitate substantive idioms, showing that while you can utter CCs that have never been heard before and understand their meaning, there are also CCs that are fixed, such as *the bigger they come, the harder they fall*.

Even though they are claiming that the CC is an idiom, they recognize that the set of CCs is not listable, so there is presumably a basic syntactic pattern to the construction. They thus give it the following syntactic structure:

(24)

![Diagram](image)

The question marks demonstrate where they think the most difficult questions in the CC are. The three main questions (for them, on this basis) are: is *the* in the CC a determiner, what are the categories of the two main CC constituents, and what rule combines them to form a sentence?
In response to the first question, they point out the idiosyncrasy of this use of *the*, citing its different historical origin as authors before them have. They mention that this *the* could be the one that also appears in the superlative (*the most*), but they maintain that the similarities between superlatives and CCs aren't such that the analysis of one could be based on the analysis of the other. In response to the second question, they say nothing. In response to the third question, they mention a reviewer's comment that the CC could be related to other sentences with parallel structure, such as *cold hands, warm heart* but say that such a parallel would only contribute the implication between clauses, not any of the other idiomatic aspects of the CC's meaning. These notes are minimal and do not shed much light on what the answers might be, but Fillmore et al. are more interested in the analysis of another construction (*let alone*) and so this concludes their remarks about the CC. This section has represented the earliest theoretical arguments for considering CCs sui generis idioms that are productive but require their own lexical items and rules.

2.1.4 McCawley 1988

Of the early works, McCawley 1988 has been the most influential in the subsequent literature on CCs. He was the first to look at languages other than English, including Mandarin Chinese and German. He was also the first to try to systematically discuss the similarities and differences between CCs and multiple other linguistic phenomena, focusing on comparatives and conditionals. He coined the term *Comparative Conditional*, which was used until 1999 when Culicover and Jackendoff re-coined the construction the *Comparative Correlative*. Fillmore 1987 is his point of departure, and he aims to defend the kind of analysis suggested but not implemented there. McCawley
claims that constructions that are composed on the basis of other constructions inherit their sub-constructions' behaviors unless those behaviors conflict with rules that are specific to the super-construction that one is trying to analyze.

He begins by discussing the patterns that CCs share with conditionals and comparatives. For conditional parallels, he mentions many of the ones noted by previous authors, including backward pronominalization. Then, in a footnote, he is the first to notice that extraction is better from the subordinate clause of a CC than some other kinds of subordinate clauses, as in (25) (McCawley's judgments intact from his fn. 1). I will not include a benchmark for this here because I will return to discuss it in greater detail in the next section. The point is that, like the conditional, the CC rates one question mark, and McCawley clearly considers it an instance of a subordinate-matrix relation.

(25) a. ?Sam is one of those guys who if you talk to, you'll like him.
    b. ?Sam is one of those guys who the more you talk to, the more you like him.
    c. ??Sam is one of those guys who although I've talked to, I don't really know him.
    d. ??Sam is one of those guys who after I talked to, I respected him.

For similarities with comparatives, he notes that both the CC and comparative involve UDCs, as discussed in earlier sections. Example (26) illustrates the dependency claimed for the comparative as well as a comparative sentence that is ungrammatical by virtue of violating the Complex NP Constraint, supporting this conclusion (McCawley takes this for granted in his paper).

(26) a. Kim is more handsome$_1$ than Sandy is t$_1$.
    b. *Kim is more handsome$_1$ than people who know Sandy is t$_1$. 

30
Next, he gives an argument that the *more* in each clause of the CC is what he calls a 'real' comparative morpheme instead of just a marker. He considers *more* in the metalinguistic comparative a good example of a marker rather than a real comparative. His evidence is that, with a metalinguistic comparative (or other instance where *more* is a marker), only *more* and not *-er* or specialized forms like *better* or *worse* can be used, as shown in (27).

In real comparatives and in the CC, any of these forms can be used, as in (28).

(27) a. Roger is more happy than surprised. (McCawley's (8a))
   b. *Roger is happier than surprised.
   c. The dinner was more bad than good.
   d. *The dinner was worse than good.

(28) a. Roger is more happy/happier/less sad than his neighbor.
   b. The dinner was worse than the brunch.
   c. The more happy/happier/less sad Roger is, the happier his neighbor will be.
   d. The worse dinner is, the better brunch will be.

(B10) *Comparative Variety in the CC*: Any form of the comparative can occur in the CC.

This also points to a fact that is never stated in the literature but is assumed, which is that a comparative of some kind is obligatory in each clause of CCs, as (29) verifies.

(29) a. *The happy I am, the more I sing
   b. *The happier I am, the I sing.

(B11) *Obligatory Comparative*: Each CC clause requires a comparative form.

After establishing these similarities between CCs and conditionals on the one hand and comparatives on the other, he turns to dissimilarities, beginning with dissimilarities between CCs and conditionals and comparatives before turning to differences between the CC and the RCC.
One property that distinguishes CC from conditionals is that the omission of be is possible in the CC. This isn't true of the first clause in the RCC, and it's not true of conditionals or how much relatives, etc., as (30) shows. He does note, however, that in what he calls 'immateriality clauses', you see deletion of the verb just like with the CC (cf. (31)). Also, he does well to point out that not all copulas can be deleted from the CC, only those that are phrase final and following an NP that isn't specific, such as a proper name (my characterization of the pattern based on (32)).

(30) a. The warmer the day, the thirstier the flowers.
   b. *The flowers thirstier, the warmer the day.
   c. *If the flowers thirsty, then the day warm.
   d. *I wonder how thirsty the flowers.

(31) Remember that Fred is just a glorified clerk, however big his salary.

(32) a. *The more obnoxious Fred, the less attention you should pay him.
   b. *The happier the customer behaving, the more things you should try to sell them.

(B12) Copula Omission: a conjugated form of to be may be omitted at the end of a CC clause as long as the subject it would normally follow is not specific.

The next property that differentiates CCs and conditionals is that CCs cannot be negated clause internally, while conditionals and the RCC can be, but only from their matrix clause, as (33) shows. In (33), (a)-(d) show whether negation in each clause is felicitous for the CC and then the RCC. Examples in (e)-(h) show the same for conditionals and reversed conditionals. The important difference is the unacceptability of (b) on the one hand and (g) on the other, showing the difference between the CC and conditional.
(33) a. #The longer John doesn't wait, the angrier he gets. (=/= it is not the case that the longer John waits, the angrier he gets)
b. #The longer John waits, the angrier he doesn't get. (=/= it is not the case that the longer John waits, the angrier he gets)
c. John doesn't get angrier, the longer he waits. (he maintains a constant level)
d. #John gets angrier, the longer he doesn't wait. (=/= it is not the case that the longer John waits, the angrier he gets)
e. #If John doesn't wait, he gets angry. (=/= it is not the case that if John waits, he gets angry)
f. If John waits, he doesn't get angry.
g. John doesn't get angry if he waits.
h. #John gets angry if he doesn't wait. (=/= it is not the case that if John waits, he gets angry)

(B13) Negating a CC: The only way to negate the sentence-level meaning of the CC is by saying *It is not the case that X* where X is a CC.

(B14) Negating an RCC: An RCC's sentence-level meaning can be negated in the same way as the CC's or by negating the main verb in its first (matrix) clause.

McCawley has a syntactic story to explain this pattern. He claims that negation is a grammatical process that only applies to sentences that branch into an NP and VP (or V’). In the RCC, this configuration is met (see (35)), but in the CC, it is not (see (34)).

(B13) Negating a CC: The only way to negate the sentence-level meaning of the CC is by saying *It is not the case that X* where X is a CC.

(B14) Negating an RCC: An RCC's sentence-level meaning can be negated in the same way as the CC's or by negating the main verb in its first (matrix) clause.

McCawley has a syntactic story to explain this pattern. He claims that negation is a grammatical process that only applies to sentences that branch into an NP and VP (or V’). In the RCC, this configuration is met (see (35)), but in the CC, it is not (see (34)).
A similar pattern holds for *yes-no* questions, as in (36).

(36) a. Does Max get angrier, the longer he has to wait?
    b. *Does the longer Max has/have to wait, the angrier he get?
    c. *The longer Max has to wait, does the angrier he get?

(B15) *Yes-No Questions*: Only the first clause of the RCC can be used to form *yes-no* questions. (and neither clause in the CC can be used to form them)

Tag questions, however, may be formed from CCs or RCCs, which McCawley cannot account for.

(37) a. The longer Max has to wait, the angrier Lucy gets, doesn't she?
    b. Max gets angrier the longer Lucy has to wait, doesn't he?
    c. *The longer Max has to wait, the angrier Lucy gets, doesn't he?
    d. *Max gets angrier the longer Lucy has to wait, doesn't she?

(B16) *Tag Questions*: Tag questions can be formed on the basis of the second clause of a CC or first clause of an RCC but not their other clauses.

As McCawley points out, the pattern in (37) provides evidence for a subordinate-matrix structure for the CC and RCC because only the subject and auxiliary verb of matrix clauses can be copied in tag questions of conditionals and other subordinate-matrix clause pairs, and here we see that tags are formed from the second clause of the CC or the first clause of the RCC only, which would fit this same pattern.
McCawley next illustrates differences between the CC and the comparative. The only one that he mentions is that whereas comparative sentences have a than phrase, CCs do not. Also, according to McCawley (and Beck 1997, etc.), they cannot have a than phrase in either of their clauses (each of which must have a comparative, as discussed above).\textsuperscript{10}

(38) a. Nicole drives faster than Emily.
   b. *Nicole drives faster.
   c. *The faster we drive than Nicole, the sooner we'll get there.
   d. *The faster we drive, the sooner we'll get there than Nicole.
   e. *The faster than Nicole we drive, the sooner we'll get there.
   f. *The faster we drive, the sooner than Nicole we'll get there.

\textbf{(B17) Than-phrase Restriction:} CCs cannot occur with than phrases in either clause.

I have included (38e) and (38f) in order to demonstrate that the ungrammaticality of (38c) and (38d) is not merely a result of an illicit placement of the than phrase. Taken together, this set of examples show that it is the presence and not the placement of the than phrase that is the issue.

McCawley also discusses differences between the CC and RCC. As noted in (B6), the is lacking in the first clause of the RCC but is required in all other (R)CC clauses. He gives a functional story for why (B6) is no surprise. I agree with him that because the two clauses look the same in the CC (each beginning with the, each having the comparative fronted, etc.), reversing them would lead to an ambiguity about whether to understand the first clause as the protasis or apodosis, a problem that is solved by the

\textsuperscript{10} I am not committed to the grammaticality judgments represented in (38); here, I am just trying to represent the stated opinion of McCawley, Beck 1997 and others. (38b) is an example of a discourse comparative, which is merely infelicitous if its comparative phrase is not inferable from context, for example, and future chapters will explore the possibility that than phrases in CCs are merely infelicitous in certain contexts (which for the CC happen to be the majority of contexts) as well.
RCC’s not beginning with a *the*. McCawley is also the first to note that it is also the case that the first clause of the RCC, unlike all other (R)CC clauses, does not involve fronting of the comparative.

(39) a. The sun seems brighter, the fewer clouds there are in the sky.  
    b. *Brighter the sun seems, the fewer clouds there are in the sky.

(B18) *Lack of RCC 1st clause fronting:* The comparative is not fronted in the first main constituent of the RCC.

This co-occurrence between fronting behavior and presence of a *the* is what motivates the treatment of *the* as a movement trigger in my analysis in Chapter 5.

The CC and the RCC differ in other ways. Only the RCC can include what McCawley calls the 'notional' comparative *increase*, as in (40), McCawley's (17). Once again, this is restricted to the first clause of the RCC, as (40b) demonstrates in showing that it is not permitted in the second clause. To this generalization of McCawley's, I add the possibility of *more and more* in place of the first clause in the RCC in (41).

(40) a. His happiness increases, the more attention he gets.  
    b. *He gets happier, the attention he gets increases.  
    c. *The more attention he gets, his happiness increases.

(41) a. He is happier and happier, the more attention he gets.  
    b. #He gets happier, he gets more and more attention.  
    c. *The more attention he gets, he is happier and happier.

(B19) *RCC Flexibility:* The first clause of the RCC can take any syntactic form as long as it corresponds to the right kind of meaning (semantic term) to be combined with its second clause, whose syntactic form is fixed/does not show similar flexibility.

11 I have marked this as infelicitous because, though grammatical, it can only be understood as a conditional like *if he gets happier, he gets more and more attention*. As we will see in the next chapter, the meanings of sentences like this differ from the meaning of CCs.
I have created in the benchmark in (B19) to describe the acceptability of both (40a) and (41a). I do not need to include additional benchmarks to rule out the ungrammaticality of the other examples because (40c) and (41c) are already ruled out by (B4: Mandatory the) because neither has a the in the second main constituent, and (B11: Obligatory Comparative) rules out (40b) as well, since it does not have a comparative in the second clause. Interestingly, McCawley does not seem to notice the pattern in (41) or think that more and more is related to the (R)CC, as evidenced by the fact that, when discussing a particular construction in Chinese, he dismisses it as unrelated to the CC but related to the more and more construction.  

In the end, McCawley has two main issues with generativists (and particularly, lexicalists – those who attribute most or all idiosyncrasies to differences between lexical items). He first doubts that for every construction there will be some overt morphology to which one can attribute construction-specific properties. Second, taking the case of the CC, he thinks that the construction grammarians have the upper hand in showing how the properties of the CC arise from conditional and comparative constructions. Doing it Chomsky's way, he says, would mean building all of those similarities into the meaning of the in the CC, which would be duplicating properties that exist independently in multiple places. He says, "Chomsky's research program would force one...in effect to treat as peculiarities of the CC construction properties that it inherits from the more basic constructions" (1988:184-5). I will return to this point in Chapter 5 after giving my syntactic analysis to show how the commonalities between conditionals, comparatives,

---

12 Though McCawley has many interesting insights into the CCs of German and Mandarin Chinese, I will not consider arguments from languages other than English in my analysis of English CCs, as mentioned in Chapter 1, which is why that data is not discussed here. The same will be true for authors such as Beek 1997, den Dikken 2005 and Abeillé et al. 2006.
and CCs can be explained and not stipulated in a generative, lexicalist framework.

2.2 Beck 1997

Though Beck 1997 is primarily a semantic work, like all good work on the syntax-semantics interface, it includes a compositional syntax and semantics as well as information about how they are related. Here, then, is the syntactic part of her analysis. The only syntactic pattern that she adds to the data already discussed above is the observation that an overt measure phrase denoting a degree difference cannot appear in the (R)CC, which is not true of other comparative sentences. It is also not the case that a meaning roughly like what one might expect of a CC with measure phrases (were it grammatical) is impossible. If we take (42a) to try to mean something like 'there is a correlation between time and temperature such that for each hour later it got, the temperature rose by three degrees', that can be captured by a sentence like (43). Even though these are not exactly synonymous, Beck (and I) take this as sufficient evidence to think that the difference in (42) is syntactic.

(42) a. *The (by) one hour later it got, the (by) three degrees warmer it was.
   b. Today is three degrees warmer than yesterday.

(43) It got hotter every hour by three degrees.

(B20) No MP in the CC: Overt measure phrases are not possible in (R)CC clauses.

In terms of assigning a general sentence structure to the CC, Beck assumes that the CC is an instance of correlativization, following von Fintel (1994). Von Fintel mentions the CC briefly as one of five examples of the remains of the correlative construction in English and German. As we saw in the last chapter, a correlative clause ‘is picked up by
a correlate anaphor in the main clause” (1994:91). Von Fintel suggests a parallel structure for correlatives as in (44) (his (24)). He discussses the similarity between conditionals and correlatives, and the structure he assigns for conditionals is (45) (his (28)/(34)).

\[(44)\]

\[(45)\]

Von Fintel's idea is that *then* constrains a hidden restrictor argument of a quantifier \(Q\).

Thus, in (45), *if, then* and \(Q\) are all co-indexed. Relating (44) and (45), in the sentence *if \(p\), then \(q\)*, if *if \(p\) were taken to be a CP, then *if* would be the correlative operator and *then*, the correlate anaphor. The proposition-denoting strings \(p\) and \(q\) would take the places of the ellipses in the tree structure in (44).

Beck takes these structures into account in giving a more detailed and CC-specific analysis for the syntax of the sentence *the faster he runs, the faster he'll get tired*, as in (46) (the English version of her (11')).
In (46), the first clause of the CC is left dislocated and adjoined to a CP. The specifier of the matrix clause CP is itself a CP. Beck insightfully notes (but then disregards) a problematic prediction with respect to the syntax-semantics interface for this structure. That is, in every other instance of correlativization, the correlative clause is co-referential with a missing constituent in the matrix clause. But in the CC, nothing is being semantically identified, only correlated. This will be discussed further in the semantic chapters, but if it is the case that correlation and coreference are necessarily different, the CC would become the only exception to a general pattern of interpretation associated with a particular syntactic configuration.

For the RCC, she assumes that the subordinate clause has been extraposed. She gives the sentence *He gets tired faster, the faster he runs* the following structure:

In (46), the first clause of the CC is left dislocated and adjoined to a CP. The specifier of the matrix clause CP is itself a CP. Beck insightfully notes (but then disregards) a problematic prediction with respect to the syntax-semantics interface for this structure. That is, in every other instance of correlativization, the correlative clause is co-referential with a missing constituent in the matrix clause. But in the CC, nothing is being semantically identified, only correlated. This will be discussed further in the semantic chapters, but if it is the case that correlation and coreference are necessarily different, the CC would become the only exception to a general pattern of interpretation associated with a particular syntactic configuration.

For the RCC, she assumes that the subordinate clause has been extraposed. She gives the sentence *He gets tired faster, the faster he runs* the following structure:
Turning now from the macrostructure to the microstructure, Beck analyzes each *the + more* constituent as a DegP. The functional head of the DegP is the comparative morpheme, and *the* is a specifier. The fact that *the* occupies Spec DegP accounts for (B20: No MP in the CC) because many analyses of comparatives, such as Heim's, assign differential measure phrases to this same position, and so the occurrence of *the* would block the occurrence of *(by) three degrees*, etc.

This is all that is said about the syntax (besides what is said about the level of Logical Form), so there are many details that are left out that make it hard to evaluate Beck 1997 with respect to the benchmarks. For example, the question of whether this theory generates CCs with only a single *the* phrase, which would violate (B3: CC clauses don't stand alone) is unclear. One one hand, CPs must be able to stand alone as sentences since that is the category of the CC as a whole (and in accordance with (B2: CC as S)). However, if a CP can stand alone and each CC clause is a CP, then what would prevent each of those individual clauses from being a stand-alone sentence as well? If some kind of feature were adopted to mark this difference, then Beck would have additional explaining to do as to why a CP of one kind branches into two CPs of the other kind. In other words, without defined rules or an alternative solution, the analysis as is does not meet benchmark (B3) (or, if one would rather claim that no CP can stand alone, then it does not meet benchmark (B2)). This particular point is problematic for several authors, including Culicover and Jackendoff and Den Dikken. Similarly, there are many points that the author just doesn't cover but for which she may rely on the semantics to account for one of the patterns in this section. For instance, (B4) requires that *the* occur in a CC clause, but as we saw earlier, there are lexical items that can occupy the Spec DegP
position besides the, which could predict the possibility of sentences like *3 miles further he ran, 1 lb. more he lost. Alternatively, Beck could posit some semantic mismatch that would arise between the meanings of a measure phrase and the rest of a CC clause, but we don't know how it would be handled. Nonetheless, there are some benchmarks that she clearly doesn't meet, such as (B5: Mandatory Fronting), (B12: Copula Omission) and (B19: RCC Flexibility), and others that she clearly does, such as (B9: That in the CC), (B10: Comparative Variety in the CC) and (B17: Than-phrase Restriction). She can account for the presences of that in the CC by considering the gapped sentence a C', because it can be divided into a complementizer, type C, and a sentence (IP, VP, etc.). That can then fill the head C. She can also account for the variety of comparative morphemes used in the CC given that they are all of the same syntactic category and should thus be predicted to be interchangeable. We'll see in the next chapter how her semantic analysis fulfills (B17: Than-phrase Restriction). She cannot account for mandatory fronting with the given its lack of syntactic parallel with wh-constructions in this framework, and she cannot account for the possible omission of the copula only under certain circumstances (like when the NP is not a proper name). As mentioned before, further discussion of Beck is reserved for the next chapter, but what we have seen in this section is that few of the benchmarks developed so far are met solely on the basis of Beck's syntactic partial analysis.

2.3 Culicover and Jackendoff 1999, 2005

Culicover and Jackendoff 1999 (henceforth, C&J) was the first contemporary work to focus exclusively on the syntactic phenomena of the CC (though they then argue that
some of them should be handled by the semantics and pragmatics). They also renewed
the older constructionist/generative debate in current terms and sparked new interest in
the CC, as evidenced by the surge of literature published after their 1999 paper. Their
2005 book then further develops the same arguments and ties the CC to another
phenomenon that they call 'left-subordinating and' (LSand) in more detail. C&J criticize
mainstream syntactic theory for focusing too much on the regularities of language and
discounting its idiosyncrasies. They insist that grammarians must pay attention to
constructions 'on the periphery' in developing theories and hypothesizing about what
aspects of language are innate universals. C&J use the CC as the poster child for their
argument, stating that it requires an idiosyncratic analysis at 'all levels' (where this
presumably means that there are syntactic idiosyncrasies, semantic idiosyncrasies, etc.).

C&J, via their framework *Simpler Syntax*, also advance a view of the syntax-
semantics interface in which syntactic relationships between clauses and semantic
relationships between propositions do not have to match in terms of symmetry. In other
words, whereas most transformational grammarians (and others) assume that (syntactic)
subordinate clauses are consistently mapped to (semantic) material implication and that
(syntactic) coordinate clauses are consistently mapped to (semantic) conjunction, C&J
argue that some constructions in natural language are syntactically symmetric (coordinate
or quasi-coordinate) but semantically asymmetric (requiring material implication). They
argue that both CCs and LSand are examples of this kind of syntax-semantics mismatch
and say that their analyses argue against any level of syntax that mirrors semantics
The architecture of the grammar in the Simpler Syntax framework consists of phonological, syntactic, and semantic structures, their interfaces, and the lexicon. This is a departure from the syntactocentric view in which there might be many levels devoted to syntax alone. The semantic structures in the grammar are a departure from the mainstream as well. C&J define Conceptual Structure (CS) as “the mental structure which grammar encodes into communicable form” (2005: 20). By this they mean that they do not assume the existence of any logical structures, propositions, etc. to be a part of CS. I will assume, however, that CS has to be responsible for the parts of language that we consider 'semantics' for the purpose of this paper.

This section is broken into subsections as follows. I begin with arguments about the macrostructure of the CC according to C&J including the categories of the two main constituents and their relation to one another. Next, I discuss properties of the internal structure of the two main constituents. Then, I review C&J's arguments that are specific to the category of the separately. Finally, I return to general discussion of the theory and how it fares with respect to various benchmarks.

2.3.1 Macrostructure

C&J initially cannot decide whether the CC as a whole is of type IP (Inflectional Phrase) or CP (Complementizer Phrase). According to C&J, benchmark (B2: CC as S) favors the IP hypothesis, since that is the traditional type for sentences (they ignore the fact that this is not true in many syntactic frameworks), but they note that if both main CC constituents are CPs (which is what they argue), then a CP that branches into two CPs is less odd than an IP branching into two CPs, as it could be an instance of CP
recursion. Either way, as C&J note, this combination is a departure from the traditional patterns of grammar. As just mentioned, C&J deem both CC clauses CPs. This is primarily to account for (B9), which states that that can optionally appear in either clause, meaning that there must be a place for a complementizer. I will return to this issue in the discussion in 2.3.4. Note that the issue of whether (B2: CC as S) and (B3: CC clauses can't stand alone) are met is a problem for C&J as it was for Beck.

The next big question (and the focus of most of their paper) is what the relationship between the two CC clauses is (I return to calling the first clause \( C1 \) and the second clause \( C2 \)). C&J consider a number of alternatives, including the two most important: that \( C1 \) is a subordinate clause and \( C2 \) is a matrix clause, and that \( C1 \) and \( C2 \) are paratactic clauses, meaning that they are “jammed together...in an idiosyncratic configuration, anomalous from the point of view of X-bar theory” (1999:547). Other analyses they consider include that \( C1 \) and \( C2 \) are conjoined by an empty conjunction, that they are both arguments of an empty main verb, etc. They state that the empty conjunction hypothesis is indistinguishable from parataxis but then later state that \( C1 \) and \( C2 \) cannot be conjuncts because they are paratactic, not coordinated, which leads to confusion for future authors as to whether C&J think that CCs are syntactically coordinate structures (I return to discussing this in the last section under 2.3).

They present many pieces of evidence favoring the subordinate-matrix hypothesis, including almost all of the benchmarks in which the CC patterns with the conditional. These notably include backward pronominalization (B7) and the ability to form tag questions only in the would-be matrix clause, as in (B16). These also include semantic patterns with tense and modality as discussed by Fillmore, McCawley and others.
(described in Chapter 3). To this, they add a pattern that I will add as a benchmark.

Example (48) shows that in subjunctive contexts, only C2 and not C1 may appear with subjunctive morphology. It is similarly true of the RCC that the clause corresponding to what is semantically the antecedent can appear with subjunctive morphology, as in (49).

(48) 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).

(49) It is imperative that/ I demand that
   a. John pay(s) more, the more he eats.
   b. *John pay(s) more, the more he eat.

(B21) **Subjunctive:** Both CCs and RCCs are possible in subjunctive contexts, but the verb in the clause corresponding to what is semantically the consequent is the only one that receives subjunctive morphology reflecting this fact.

This is the same pattern we saw with tag questions in (37) that was encoded in (B16). In both of these cases, the pattern of the CC and RCC parallels the one found in conditionals such that only the main clause and not the subordinate clause may appear with subjunctive morphology or serve as the basis of a tag question. Thus, these favor the subordinate-matrix hypothesis.

There are two main pieces of evidence favoring the paratactic hypothesis, according to C&J. First, they say that unlike regular instances of subordinate clauses, the order of the matrix and subordinate clauses cannot be switched in the CC or RCC. Clearly, given the fact that the RCC is called the Reverse CC, McCawley and others thought that the CC could be reversed and that the result was the RCC. But C&J are correct that, if you treat the CC and RCC as separate constructions (and we have seen
above that they exhibit some different patterns), neither can be reversed and maintain the same meaning:

(50) a. The bigger the student, the older the parent =/= The older the parent, the bigger the student

b. Her dad is older, the taller she is =/= *The taller she is, her dad is older.

As we see in (50), only the counterpart of the RCC is actually ungrammatical, and no additional benchmarks are needed to account for that fact. If we take both the CC and RCC to describe an asymmetric semantic relationship and define the CC as the construction in which the antecedent precedes the consequent and the RCC as the construction in which the consequent precedes the antecedent, then the “reversed” RCC in (50b) should be a CC, but it violates (B3: CC clauses don't stand alone) and (B4: Mandatory the). If (50a) were an RCC, it would violate (B6: No the in the 1st clause of the RCC) in having a the in the first clause. As it is, (50a) does not violate any syntactic constraints, nor should it, given that both sentences in (50a) have the same syntactic structure (so switching CC clauses gives another CC). Even in the next chapter, I will hold off creating a semantic benchmark to encode an inequality like (50a) because in Chapter 6, I'll argue that these actually are equal in terms of their syntactico-semantic form (since correlation is a symmetric relation, etc.) but differ due to an asymmetry arising from the conservativity inherent in natural language operators. The point of (50) for C&J is that a common test for subordinate-matrix relationships is whether the clauses can be switched, and this example shows that they cannot in the CC and RCC, which makes them different from other adjunct subordinate-matrix structures as in (51).
(51) a. If I buy milk, I drink juice = I drink juice, if I buy milk.
b. Because she is older, she is taller = She is older because she is taller.

The second piece of evidence that casts doubt on the CC as a subordinate-matrix construction has to do with extraction. As we saw in section 2.1.1, CCs are UDCs, like topicalization, wh-questions, etc., so something in each clause (more, etc.) is extracted within each clause. C&J then look at whether any further extraction is possible when CCs are embedded within other UDCs. Here are some examples from C&J 2005:

(52) **Relative Clauses:**
a. This is the sort of problem which, the sooner you solve $t_i$, the more easily you can implement a solution to $t_i$. [Extraction from both C1 and C2]
b. This is the sort of problem which, the sooner you solve $t_i$, the more easily you’ll satisfy the folks up at corporate headquarters. [Extraction from C1]
c. The folks up at corporate headquarters are the sort of people who, the sooner you solve this problem, the more easily you’ll satisfy $t_i$. [Extraction from C2]
d. The sooner you solve this problem, the more easily you’ll satisfy the folks up at corporate headquarters.

(53) **Topicalization:**
a. This problem, the sooner you solve $t_i$, the more easily you’ll satisfy the folks up at corporate headquarters. [Extraction from C1]
b. ? The folks up at corporate headquarters, the sooner you solve this problem, the more easily you’ll satisfy $t_i$. [Extraction from C2]

(54) **It-cleft:**
a. It is this problem 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, (not the ones here at the regional office) who, the sooner you solve this problem, the more easily you’ll satisfy $t_i$. [Extraction from C2]

(B22) **C1 Extraction:** Extraction can occur from the first clause of the CC.

(B23) **C2 Extraction:** Extraction can occur from the second clause of the CC.

---

13 I use the terms *extraction* and *fronting* in talking about UDCs for historical purposes (to be consistent with past literature) and remain neutral about whether these UDC examples are derived by explicit movement operations.

14 I choose to break this pattern into three separate benchmarks because some theories account for some extraction patterns and not others.
(B24) *C1 & C2 Extraction:* Extraction can occur from both clauses of the CC at once.

As we see in these examples, extraction is possible from both clauses or from either clause by itself. In the cases of clefting and topicalization, we see a preference for extraction from the first clause. There are three main reasons that C&J find this puzzling (and on the basis of which they argue for its uniqueness). The first is that most UDCs are islands to extraction, meaning that none of these three possibilities should be available.15 Take the example of (55a), a sentence with an initial free relative clause that also has a UDC and allows extraction from neither conjunct (55b,c) nor both together (55e):

(55) a. Whatever he cooks in that crockpot, I'll eat with a spoon.
   b. * That's the crockpot that whatever he cooks in, I'll eat with a spoon.
   c. * That's the spoon that whatever he cooks in that crockpot, I'll eat with.
   d. Whatever he cooks in that crockpot, I'll eat from the crockpot.
   e. * That's the crockpot that whatever he cooks in, I'll eat from.

Second, and most importantly for the argument that C&J are making, they do not believe that this matches the extraction pattern of conditionals (which would indeed provide evidence against a subordinate-matrix form for the CC). The Constraint on Extraction Domains (CED) (Huang 1982) blocks extraction out of adjuncts such as the subordinate *if*-clause in a conditional. Thus, the extraction signature for sentences with an asymmetric relationship between clauses is predicted to be: extraction only from the matrix clause, not the subordinate clause, and therefore no extraction from both conjuncts. This is demonstrated for *because* subordinates and conditionals in (56) and (57).  

15 That UDCs are extraction islands is assumed by C&J as well as most contemporary syntactic textbooks. Nevertheless, as alluded to in 2.1.1, recent work such as Hofmeister and Sag 2010 disputes the validity and utility of this characterization.
(56) a. *What book will Robin understand linguistics better because he read(s)?
   b. What subject will Robin understand better because he read(s) Chomsky's book?
   c. *What book will Robin understand better because he read(s)?

(57) a. *What book will Robin understand linguistics better if he reads?
   b. What subject will Robin understand better if he reads Chomsky's book?
   c. *What book will Robin understand better if he reads?

Thus, if the CED were exceptionless (as it is for the two examples we've seen so far) and the CC clauses were in a subordinate relationship, we would expect extraction out of C2 exclusively in the CC. As we know from data such as (52)-(54), this does not match what we find with the CC. In fact, not only is extraction from C1 possible, it's actually better. This, then, is the primary evidence against the subordinate-matrix hypothesis and in favor of the paratactic hypothesis.

The third puzzle, however, appears to be a problem for the paratactic hypothesis insofar as parataxis is syntactically indistinguishable from coordination. Ross (1967) introduced the Coordinate Structure Constraint (CSC), stating that no conjunct of a coordinate structure (a sentence with clauses joined by a coordinating conjunction such as and, or, or but) nor any element contained in that conjunct can be moved/moved out of the conjunct, meaning that they cannot create UDC/filler-gap dependencies. For example, if we try to extract the object of buy from the coordinate structure John bought the book and read the magazine in order to form a question, the result is ungrammatical:

(58) *What book, did John buy t₁ and read the magazine?

---

16 This is often divided into two constraints: the Conjoint constraint (extraction of an entire conjunct), and the Element constraint (extraction of some part of a conjunct). I will only consider the Element constraint here since there is no evidence that the CC violates the Conjunction constraint.
Nevertheless, as Ross himself noted, there are certain exceptions to this rule, as in (59).

(59) What book, did John buy t, and read t? 

In this case, what book is the object of the verb in both of the coordinated conjuncts, whereas it was only the object of the verb in the first conjunct in (58). Thus, the extraction of the filler from the gap position can take place symmetrically, or Across the Board (ATB, Williams 1978), but not asymmetrically, from only one conjunct.

Extraction from either conjunct is predicted to be equally bad, c.f. the counterpart of (58): *What magazine did John buy a book and read t? Thus, if the CC were like a normal coordinate structure, extraction from C1 and C2 should be blocked unless it occurs in both in the same sentence, ATB. Once again, this does not match the pattern we saw in (52)-(54). Thus, we have evidence against a coordinate hypothesis as well as against a subordinate hypothesis.

C&J acknowledge that the number of patterns that provide evidence for the subordinate-matrix hypothesis vastly outnumber the patterns that provide evidence for a paratactic or coordinate hypothesis, but they argue that both kinds of patterns can be explained if we stop insisting that the (a)symmetry of the syntax and semantics must match. In other words, un-coupling the syntax and semantics gives us four possibilities rather than just two. In that case, the extraction pattern described above for subordinate clauses is only predicted to describe the behavior of clauses with an asymmetric syntax and semantics. Likewise, the pattern dictated by the CSC for coordinate clauses only applies to those that are both syntactically and semantically coordinate. C&J then propose that the CC belongs to the category of sentences that have a syntactic symmetry
with a semantic asymmetry (and because of this syntactic symmetry, they cannot be considered subordinate-matrix clauses). C&J further argue that the CSC is a constraint at the level of CS, not at the level of syntax. Arguing against a syntactic explanation of CSC violations puts C&J in league with Goldsmith 1985, Kuno 1987, and Kehler 2002, among others, who are arguing not for a treatment of the CC but for other exceptions to the CSC. This means that the CSC only applies to structures that are semantically symmetric, but since the CC was put into the 'semantically asymmetric, syntactically symmetric' category, the CSC is not predicted to restrict the CC in any way. The CED, on the other hand, is a constraint that C&J argue is syntactic. It blocks extraction from the subordinate clause of sentences with asymmetric syntactic relationships. Since the CC is not an instance of an asymmetric relationship in the syntax, the CED does not restrict its extraction. As a result, extraction should be licensed in the CC from either clause or both since neither the CSC nor the CED are operative on their kind of sentence. This finally matches the pattern we find in (52)-(54). They conclude, then, that CCs are paratactic and that they have symmetric syntax with an asymmetric semantics.

### 2.3.2 Structure of each CC clause

First I will describe the properties of the syntactic structure that C&J assign to both clauses and their relation to other structures. This will reveal the ways in which the CC is not necessarily unique. Then, I will describe two patterns that the CC does not share with these or other constructions.

From the last section, we already know that each clause is a CP. We also know that they are symmetric, so throughout this section, I give only a single CP structure, but
it extends to the analysis of both clauses. There are two main structural building blocks that C&J use to create their proposed structure for the CC: analyses of fronting (a parallel with other UDCs) and analyses of comparative sentences. I will not review specific analyses of other UDCs here, but the basic idea for any kind of fronting is that the fronted element occurs on the left periphery in the specifier (Spec) of XP and is marked with an index, $i$. Then, there is a trace, $t(i)$, somewhere in the X' that links the index to the place where it would have been in situ.

A standard analysis of comparatives, following Bresnan 1973, for the sentence in (60) appears in (61).  

(60) John is 4 inches taller than Bill is.

(61) In this way of doing things, measure phrases like 4 inches and degree constructions like more than Bill is are base generated in the specifier of AP. Here, because comparatives are considered quantificational, they are part of a Quantificational Phrase (QP). The

---

17 Given when Bresnan was writing, her structures included S's, NPs, etc. rather than CPs, IPs and DPs, but these differences aren't crucial to the point being made here.
comparative morpheme fills the $Q$ position and has a $\textit{than}$ phrase as its complement. This is because there is a relationship between a $Q$ morpheme and the complementizer that it licenses. For example, $\textit{more}$ licenses $\textit{than}$, but $\textit{as}$ licenses a second $\textit{as}$, etc. In the word order of many sentences, however, an adjective intervenes between $\textit{more}$ and its $\textit{than}$ phrase, and so for Bresnan and others, the CP is eventually extraposed to the right periphery and adjoined at a higher (IP) level. Jackendoff 1977, on the other hand, generates the XP constituent on the right to begin with. This is only notable in that Jackendoff 1977 is the model that C&J use, but this is one of the only ways in which it differs from the standard analysis.

In order to tweak this to fit the CC, C&J say that $\textit{the}$ is a specifier of the comparative in the same way that $\textit{a lot, 4 inches}$, etc. would be. Now, in assigning $\textit{the}$ to the Spec of QP, which is in turn the Spec of AP (so $\textit{the}$ is a specifier of the specifier of an AP), C&J are preserving Beck's intuition about why measure phrases and $\textit{the}$ do not co-occur since both occupy the same structural position. Then, in order to account for the fronting, the entire AP with the comparative morpheme and adjective will move to the Spec of the clausal CP, leaving a trace in the IP that it moves out of. This is the structure that we see in (62).
The next section takes this structure as its point of departure and notes that some predictions it makes are not borne out (nevertheless, this remains the final structure that they endorse).

Moving on to some of the idiosyncrasies that C&J note for the CC as compared to what one might expect on the basis of (62), there are three main patterns that they say make individual CC clauses weird. They are: possible omission of the copula, the subject requirement on imperatives and gerundives, and the properties to be discussed in section 2.3.3 below distinguishing the from other comparative specifiers.

Recall from benchmark (B12) that the copula can often be omitted in the CC. Given the parallel between the analysis of the comparative in (61) and the CC in (62), it might seem promising to analyze copula omission in the CC in terms of copula omission in the comparative. These can't be treated in parallel, however, because the copula can be omitted in comparatives with specific subjects, which is not true of the CC (as previously noted). In other words, the data in (63) show that CCs do not parallel comparatives in the way that we might expect based on (62).
(63) a. A dog is bigger than a cat (is). (comparative, non-specific)
   b. The bigger a dog (is), the bigger a cat (is). (CC, non-specific)
   c. Alex is taller than Colleen (is). (comparative, specific)
   d. *The taller Alex, the happier he will be. (CC, specific)

Another puzzle comes in the form of CCs and RCCs in subjunctive and imperative contexts. Imperatives can be formed on the matrix clause of the CC or RCC (despite the fact that C&J label these 'good' examples with question marks), but a subject is obligatory in the CC, but not the RCC, as (64) shows. (64a) and (64b) are CC and RCC examples of imperatives with subjects, and both are supposed to be fine. (64c) and (64d) demonstrate the ungrammaticality that supposedly results when the imperative is formed on the subordinate clause rather than the matrix for the CC and RCC. Finally, (64e) and (64f) demonstrate the difference in grammaticality between the CC and RCC for an imperative without a subject (like everyone) in the previous examples: the CC requires a subject.

(64) a. The more John eats, the tighter everyone keep your mouth shut about it, okay?
   b. Everyone keep your mouth shut tighter, the more John eats, okay?
   c. *The tighter everyone keep your mouth shut about John's eating, the more he eats, okay?
   d. *John eats more, the tighter everyone keep your mouth shut about it, okay?
   e. *The more John eats, the tighter keep your mouth shut about it, okay?
   f. Keep your mouth shut tighter, the more John eats, okay?

(B25) Imperatives: The clauses in the CC and RCC that correspond to semantic consequents can take on an imperative verb form when in imperative contexts, while their other clause cannot.

(B26) Subjects with Imperatives: In order for a CC clause to take on an imperative form, it must have a subject.
Then, we see the same pattern for CCs and RCCs taking the form of an accusative-gerundive complement (though C&J report limited acceptability of both C1 and C2 taking the accusative-gerundive form as opposed to C2 alone, as in (65b)). I include both him and his as possible subjects in the clauses that take the form of the complement because, though C&J use him in their examples, my dialect requires the possessive his.

(65) a. I can well imagine the more he eats, the fatter him/his getting.
    b. *I can well imagine the more him/his eating, the fatter him/his getting.
    c. I can well imagine him/his getting fatter, the more he eats.
    d. *I can well imagine the more him/his eating, the fatter he gets.
    e. *I can well imagine he gets fatter, the more him/his eating.
    f. *Bill can well imagine the more he eats, the fatter getting.
    g. Bill can well imagine getting fatter, the more he eats.

(B27) **Acc-Gerundives**: The clauses in the CC and RCC that correspond to semantic consequents can take on an accusative-gerundive form, while their other clause cannot.

(B28) **Subjects with Acc-Gerundives**: In order for a CC clause to take on an accusative-gerundive form, it must have a subject.

Here again, we see that what looks like a matrix clause licensing various inflections that the subordinate clause is not able to, and we also see a continued difference between the CC and RCC with respect to the subject requirement. C&J do not modify their analysis on the basis of these data and say nothing more about them.

### 2.3.3 Properties of The as a comparative specifier

Given a syntactic structure for a CC clause along the lines of (62), the is in the position of a specifier of the comparative – Spec of QP. There are three main patterns with respect to which the is different from other comparative specifiers such as so much, how much, a lot, etc. C&J actually mention four differences, but one is the possibility of
subject-auxiliary inversion mentioned by McCawley that is no longer productive in English and will thus be ignored here.

The first and second patterns are ones we have already seen above for the CC, but they will be restated here with a comparison to various comparative specifiers. They relate to the presence of a complementizer, (B9), and the obligatory rather than optional fronting, (B5). Taking the example of *so much and how much* as other comparative specifiers, *the* is the only one of the three that permits *that*, as in (66).

(66) a. The faster that you eat, ...
   b. *So much faster that did you eat...  
   c. *How much faster that you eat...

The CC and *how much* pattern together in contrast to *so much* with respect to being able to be in situ, as (67) shows:

(67) a. *You eat the more, you want the more.  
   b. You ate so much you made yourself sick.  
   c. *I wonder you ate how much.

The third pattern is the lack of preposition pied-piping in the CC. Here once again we see *so much* and *how much* pattern in opposition to the CC. They permit the preposition to be fronted or in situ. CCs, on the other hand, only permit the in situ position, which is why (68a) but not (68b) is grammatical.

(68) a. The more people Bill talks to,...  
   b. *To the more people Bill talks,...  
   c. So many people does Bill talk to that...  
   d. To so many people does Bill talk that...  
   e. I wonder how many people Bill talks to.  
   f. I wonder to how many people Bill talks.
Preposition Pied-Piping: The CC does not permit prepositions to be extracted along with the more when it precedes them.

Note that this pattern is not due a general constraint against fronting PPs in the CC given that examples like the more out of breath you are,... are grammatical.

These are all of the ways that C&J mention that the is unlike other specifiers of QP. On this basis, however, they do not modify their structure to change the category of the or put it in another location. Thus, these are all benchmarks that C&J have nothing to say about.

2.3.4 Discussion

Looking first at the big picture, I find that C&J share a common approach with Fillmore and McCawley: come up with as plausible a structure as possible for the CC, then focus on the myriad ways in which the data suggest complications for the best possible straw man. Conclusion: CCs are idiosyncratic and cannot be captured sufficiently in the kinds of grammars that we have developed to this point. Now, C&J admit that Minimalists could use functional heads, strong features and attraction to account for this data, but their point is that those things would be specific to this construction, and so an analysis along those lines isn't independently motivated over an analysis that just states which aspects are specific to this construction. I think this is a reasonable point, and both Borsley and den Dikken support it, as we will see in the next two sections. In my analysis, too, some aspects of the grammatical form of the are specific to the CC, but I will justify in Chapter 5 the reasons for which I think that this is the most reasonable way to do things and not ad hoc. In fact, den Dikken set out to show
that it is possible to handle the CC lexically without any additional grammatical rules, and though he does not succeed, this is exactly what my analysis accomplishes.

C&J remark that the CC construction is a case where ‘the mapping of syntactic structure into conceptual structure is more or less arbitrary’ (568). But they also note that many languages have a similar structure that maps to the correlational meaning of the English CC. So what do they mean? The literature takes them to say that the syntax-semantics link in CCs is totally arbitrary, and then Borsley and den Dikken disagree, but to varying degrees. Den Dikken says that the syntactic structure and subsequent LF mapping into interpretation is the same in every language, which is too strong a claim in the other direction, and Borsley represents the middle ground, as we will see in the following sections. The main claim that Borsley levels against C&J is that they do not develop a formal analysis. I find this both true and false. It is false in the sense that they posit a syntactic structure that makes certain predictions, and they also make an argument that within Simpler Syntax, separating the syntactic and semantic argument relations and letting the semantic argument relations explain syntactic patterns will account for what would otherwise be mispredictions of the structure in combination with the constraints on it that are believed to hold. In both of these senses, they are developing a formal analysis to account for the data about the CC. Borsley is correct, however, in that there are many patterns besides extraction that the C&J analysis mispredicts (including most of the benchmarks developed in this section), and so there is not a working formal analysis that accounts for the full range of data described in their article. This is the most obvious flaw in the theory that C&J propose, but even if it doesn't cover the full empirical pattern of the CC, it could still be the case that, for the issues they do account for, their approach
is onto something. In particular, it is the burden of future authors to match the story that C&J tell about extraction, which is something that Borsley himself neglects. Given the importance of the extraction facts, they are worth talking about in more detail.

I disagree with C&J's conclusion that the syntax of the CC is paratactic. I believe that the CC is indeed a subordinate-matrix relationship. I believe that the RCC is similarly a subordinate-matrix relationship but that it differs from the CC in other particular aspects of its syntax, which leads to a difference in patterns between the CC and RCC. I agree that the CSC is not a syntactic constraint. I also agree that it does not apply to the CC because, when the conservativity of the quantificational operator is taken into account, it is not a symmetric structure in terms of semantico-pragmatic relations. I disagree that the CED is a syntactic constraint or that it applies to conditional antecedents in the first place (or even that it is a real constraint on extraction domains). Thus, even though the conditional, CC, and RCC are asymmetric at both levels, the CED does not apply at either level, thereby predicting that extraction from any clause (or combination) should be possible.

In order to argue that the CC is a subordinate-matrix construction, we must re-evaluate the evidence against subordination. The first issue was that the CC and RCC could not have their clauses switched. While it is true that some subordinate clauses can appear on either side of a matrix clause without changing the meaning of the sentence, this is not true of all of them. Though it was shown above for *if x, y* statements, it is not true of *if x, then y* statements, as (69) shows.

(69) If I buy milk, then I drink juice =/= *Then I drink juice, if I buy milk.*
Despite (69), conditionals of both the if and if, then variety are considered subordinate-matrix relationships, so it seems plausible that the CC and RCC pattern with if, then. The second argument concerned the extraction facts, especially with respect to the CED, where it was suggested that the first clause of the CC does not share the pattern of other subordinate adjuncts in being islands for extraction. First, I will show that C&J’s account in terms of separating syntax and semantics and having the constraints apply at different levels cannot correctly predict the pattern of the CC by claiming parataxis. Then, I will show how to overcome the apparent problem posed by the CED for the analysis of the CC as a subordinate-matrix relationship.

First we have to clarify what parataxis really is, whether it is actually different from coordination, and how we would expect it to pattern. C&J seem to (perhaps unintentionally) go back and forth between whether coordination and parataxis are the same or different. The only property distinguishing parataxis from coordination is the presence of a coordinating conjunction in the latter. C&J thus call paratactic structures “quasi-coordinate” (C&J 1999, 551). This means that coordination and parataxis are not distinguished on the basis of symmetry. Since coordination and parataxis are strictly syntactic relationships, this means that either could ostensibly be paired with either a symmetric or asymmetric semantics (according to C&J), but both must be syntactically symmetric. In the tradition of Transformational Grammar, extraction from both clauses of a symmetric structure is considered ATB extraction, whereas extraction from multiple complements of the same verb, each of which allow extraction, such as *Who did you tell t that you would pay a call on t?* is considered parasitic gapping (as long as one gap is licensed by another). It is surprising, then, that C&J state that the extraction from both
clauses that is possible in the CC is an instance of parasitic gapping rather than ATB on
the basis that the CC is not a coordinate structure. They also conclude that C1 is neither
an adjunct nor a conjunct, but given what they say earlier about the paratactic and
coordinate hypotheses not making testably different predictions, I don't see how they can
argue against conjunction.

If it is true that the CC is an instance of parataxis in which each clause contains a
UDC, then it should pattern similarly to other paratactic constructions that contain
fronting (modulo the affect of whatever constraints apply to their semantic form).
Example (55), repeated here, is just such an example. If, like the CC, free relatives had
an asymmetric semantics, then extraction should be possible in all three cases in
(55b,c,e). If it had a symmetric semantics, the CSC would apply and block (55b,c) but
license (55e). Neither of these alternatives fit the pattern in (55). In fact, even under the
hypothesis that fronting blocks further extraction, that only explains why (55b) and (55e)
are ungrammatical; in (55c), extraction is blocked from a non-UDC clause. Thus, we
might attribute the ungrammaticality of (55c) (at least) to parataxis. And if the CC
contains UDCs of its own and is paratactic, then why is extraction possible from any of
the clauses? None of this is explained by the theory that C&J put forth.

(55) a. Whatever he cooks in that crockpot, I'll eat with a spoon.
   b. * That's the crockpot that whatever he cooks in, I'll eat with a spoon.
   c. * That's the spoon that whatever he cooks in that crockpot, I'll eat with.
   d. Whatever he cooks in that crockpot, I'll eat from the crockpot.
   e. * That's the crockpot that whatever he cooks in, I'll eat from.

Thus, I conclude that the hypothesis that the CC is paratactic is not adequately supported.
The second issue, whether extraction from C1 shows that the CC cannot be a subordinate clause, is also a case in which the actual data do not match what we would predict them to be on the basis of theoretical constraints. Despite what the CED predicts, extraction from a subordinate *if*-clause is acceptable in most cases, as in (70). Similarly, (71) and (72) show that extraction is possible from the second clause and both together.

(70) That's the house that if Terry had left me a key to, I could have watered his plants.
(71) Those are the plants that if Terry had left me a key, I could have watered.
(72) Those are the plants that if Terry had bought, I would have had to care for.

Here, we find a pattern that does actually match the pattern of the CC exactly, and it is a subordinate-matrix relationship between clauses. In fact, in the literature on semantic, pragmatic, and cognitive explanations for extraction, such as Kehler 2002, extraction from the subordinate clause is predicted to be better than extraction from the matrix clause, which is exactly what we saw in the CC cases of Topicalization, etc. in (53) and (54). For Kehler, this is related to the cause-effect relationship that represents the discourse function of each clause and the fact that the cause is more likely to be the topic of conversation, which is in turn related to Kuno 1987. Even though this is diametrically opposed to what the CED predicts, it seems to be validated by empirical evidence. The table below contains both the predicted patterns (as indicated by author's names) and empirical patterns discussed in this section. It is clear from this chart that the CC and conditional share the same pattern, which does not match any of the other patterns.
I conclude that Huang's predicted pattern of subordination does not apply to all subordinate-matrix clause relationships and that one exception is the conditional. Thus, the antecedent of a conditional is a subordinate clause, and since its pattern matches the pattern of the CC, I further conclude that $C_1$ is a subordinate clause and $C_2$, a matrix clause. In summary, then, I find no reason to separate conditionals and CCs on the basis of either of the tests that C&J claim are evidence against the CC as a subordinate-matrix clause. I think that the evidence is sufficient to reject the paratactic hypothesis and assume subordination.

2.4 Borsley 2004a,b

Borsley is the only author to date whose work is sufficiently concrete in terms of both structures and rules that it is possible to see what predictions it makes in every area. Borsley gives an adequate account of the syntactic patterns in the CC. He does so by adding non-standard head-filler clauses (which is typical in HPSG), non-standard clause types (such as correl-clause), making a special lexical entry for *more* when it appears in CCs, creating new constraints for these new clause types, etc. Though many of these additions are standard within HPSG or a particular approach to HPSG that is
construction-oriented, the question for Borsley is less whether his theory is adequate than whether it is explanatory. A theory that just lists the patterns of the CC is also a descriptively adequate theory but clearly not explanatory. Insofar as Borsley's analysis ties certain patterns of the CC into more general natural language patterns, it is at least partially explanatory, but I will argue that some aspects of the theory are ad hoc. In going through Borsley's analysis, I work in the opposite direction than I do with C&J, starting with how the *the more* is combined and working toward how the clauses are combined. I once again end with discussion of the theory.

2.4.1 Microstructure of each CC clause

In Head-Driven Phrase Structure Grammar (HPSG), identifying the heads of various phrases is crucial. Because, as we saw in (B1), the CC is a UDC, it is a filler-gap construction. In such cases, we expect that conditions at the gap site will be reflected by the properties of the filler. Further, within a phrase-structure grammar with a notion of headedness like HPSG, we expect those conditions to be projected by heads. This aids us in determining what should be the head of various phrases. The following are three kinds of connectivity that help to identify head status (Bob Levine, p.c.). Because these are all tied to UDCs, they are relevant for the CC in its capacity a UDC and as such, I include them as sub-benchmarks of (B1).

(73) a. Robin put the book *the edge/*careful/near the edge.
   b. *Which edge/*How *careful/*Near which edge did Robin put the book?
   c. The *more (the) edge/*more careless/nearer to the edge we put the books, the more likely they'll topple.

(B1.1) *Major Category Connectivity*: The syntactic category of the filler and gap must match.
(74) a. You said that certain people *was/were coming to the party.
    b. Which people did you say *was/were coming to the party?
    c. The fewer people Robin believes *is/are coming to the party, the happier he'll be.

(B1.2) **Number Connectivity**: The number of the filler and gap must match.

(75) a. We *constructed/*produced/made headway up the river.
    b. I wonder how much headway we *constructed/*produced/made up the river.
    c. The more headway we *construct/*produce/make up the river, the more of a workout we'll get.

(B1.3) **Idiom chunk Connectivity**: idiom restrictions from the gap constrain the identity of the filler.

In each of the examples in (73)-(75), (a) shows what is grammatical in a simple sentence (non-UDC), (b) shows that the same pattern as in (a) holds for a classic UDC such as a question or relative clause, and (c) shows that the same pattern holds for the CC. The data in (73) exemplify major category connectivity because what is at issue is the syntactic category of the argument. Here, *put* requires a prepositional phrase and not a noun phrase or adjective, regardless of whether it is local as in (a) or long-distance as in (b) and (c). (74) is evidence of number connectivity because in each case, only the plural and not the singular verbal morphology is grammatical given the plural noun phrase *people*. Finally, in (75) we see something called idiom chunk connectivity illustrated with the idiom 'to make headway'. Since there is not a similar idiom 'to construct headway', etc., the others are ungrammatical. I will return to these in chapter 5 where I discuss how my analysis predicts these connectivity facts, but their relevance here is in understanding the background assumptions in Borsley's work that form the basis of his decisions about headedness in the CC.
For a CC clause such as *the more books I read*, there are two daughters: *the more books*, and *I read*. Borsley takes *I read* to be the head daughter, and *read* to be its head. In the non-head daughter, *books* is the head of *the more books*, and *more* is the head of *the more*. He also identifies *read* as the head of *I read*, which would once again be the standard choice within the HPSG framework. As for *I read* as the head daughter, Borsley claims that *the more books* is a definite NP argument of *read*, in which case *read* would indeed normally be the head. *I* is also an NP argument of *read*.

Because neither *the* nor *more* is the head of its phrase (*the more books*), neither is guaranteed to appear in the CC. Given benchmarks (B4: Mandatory *the*) and (B11: Obligatory Comparative), however, we know that both *are* actually required in any CC phrase. To account for this, Borsley first posits a NON-LOCAL feature FILLERFORM (FFORM for short) that has values *the* and *none* (for when there isn't a *the*). This is not ad hoc insofar as the same is done for *wh*-fronting, which has its own NONLOCAL feature that is either *wh* or *none* in its values. Then he requires that there be a *the* value for FFORM in any CC sentence in order to guarantee that you don't get CCs without a *the* present (I'll return to how this is done in the section on the macrostructure). The FFORM feature is subject to a special Linear Precedence (LP) condition when its value is *the*. The condition states that items with FFORM *the* precede items with FFORM *none*. This is needed to rule out examples where something like a preposition (or anything else) appears before *the* in the CC, as (B29: Preposition Pied-Piping) requires. This ensures that in pied-piped CCs like *the more hostages' stories I hear, the more...*, the *the*-phrase is always on a left branch of the filler.
Next, to guarantee that you also get a comparative word appearing with *the* in a CC clause, Borsley makes *the* the specifier selected by a special comparative morpheme. So there is one comparative morpheme that selects a *than* phrase complement, which the *more* in the CC does not do, but the *more* in the CC requires a *the* specifier, which the ordinary *more* does not do. The structure of this phrase is as in (76) (Borsley 2004a's (49)).

(76)

So far, in terms of additions to the grammar and lexicon, we have the motivated FFORM feature and lexical entry for *the* on the one hand and an ad hoc LP condition and new lexical item for *more* on the other.

Moving on to the form of a *the* clause, Borsley says that *the* clauses are both head-filler phrases and clauses, which is typical in HPSG in order for them to inherit both clausal and head-filler kinds of properties. The *the* clause is given an official title, *the-cl*. Borsley also splits the category of head-filler phrases in two to accommodate the CC. The two types are standard head-filler phrases and non-standard head-filler phrases, of which *the-cl* is the only example. Borsley considers this move necessary in order to account for benchmark (B9), which allows *that* in CC clauses. Because this is not true of
other head-filler phrases, they cannot result from the same constraints as the CC, and constraints are specified for particular kinds of phrases and clauses. In (77) we see the clauses with and without a complementizer with their head daughters in bold.

(77) a. The more books I read, ...
b. The more books that I read, ...

Both of these heads contain a verb, but one also has a complementizer. Borsley states, following Ginzburg and Sag 2000, that heads can be 'pure verbs' (like read in (77a)), gerunds, or complementizers (like that in (77b)) that are either saturated (have taken all of their arguments) or have a non-canonical subject. There are different HEAD feature values for heads with a verb or gerund as compared to those with complementizers, however. The constraints on the the-cl don't make reference to these values since they are acceptable with or without a complementizer. But now that Borsley has let any head-filler phrase be headed by a verb or complementizer, he must disallow head-filler phrases with complementizers other than the CC. Thus, he then has to create an extra constraint for the category of standard head-filler phrases saying that they are sensitive to the HEAD value of their head daughter, requiring it to be a verb or gerund type and not a complementizer type.

The fact that the more comes before she eats in the more she eats,... and similar sentences can be attributed to a general fact about fillers in English that they always precede their heads. Thus, the LP constraint already present in HPSG saying that the non-head daughter must precede the head daughter applies to the the-cl as well. This satisfies benchmark (B5: Mandatory Fronting). Finally, there is a constraint (alluded to
above) requiring that sentences with an FFORM item are instances of head-filler-structures and have a non-head daughter with an FFORM feature valued `the`. In other words, *the-cl* head-filler-phrase/clause hybrids are required to have some daughter with a `the` value for the FFORM feature.

The Generalized Head Feature Principle (Ginzburg and Sag 2000) states that a phrase inherits many of the syntactic and semantic properties of its head daughter by default. One of the exceptions to this default rule concerns all head-filler structures. In these cases, the filler discharges a member of the SLASH value set of its head sister so that the head sister does not inherit that particular SLASH value. The Attribute Value Matrix (AVM) in (78) illustrates some of the requirements discussed above.

(78)
Here we see that [4] is the LOCAL value of the non-head daughter that is part of the SLASH value for the head daughter but is absent from the SLASH value of the *the-cl* as a whole. This is true of all filler-gap constructions within HPSG, all of which violate the GHFP. Because the GHFP is stated as a default condition and filler-gap constraints supersede the default, this is not a problem with Borsley's analysis. The v that appears in the HEAD feature of the *the-cl* stands for *verb* and indicates that the head of the head daughter is a pure verb. This would be replaced with a *c* in the case of a complementizer in *the-cl*. The only thing appearing in (78) that has not yet been discussed is the valence feature CORREL in the *the-cl* AVM, whose description I now turn to.

### 2.4.2 Macrostructure of the CC

Borsley introduces a new feature CORREL in order to distinguish clauses beginning with *the* from other kinds of clauses. It is typical in HPSG for the feature that distinguishes clauses to be different from the NONLOCAL feature that governs the placement of a particular part of the clause (as in QUE v. INTERROG for questions). CORREL is proposed for use outside of the CC for conditional sentences and *as...so...* sentences as well. Thus, CORREL has the following possible values: *the, as, so, if, then* and *none*.

Not surprisingly, CORREL makes an appearance in various constraints. The constraint that requires a daughter with FFORM *the* also requires that the clause itself be CORREL *the* and have a finite verb. Up to this point, we have seen that Borsley comes up with an analysis for CCs and RCCs that makes a number of correct predictions with respect to some of the benchmarks and captures the parallel between CCs and other kinds
of clauses and UDCs. We have also seen a variety of technical machinery, much of which is standard in HPSG. There are some additions to the standard, including the LP rule preventing anything from occurring before *the* and the *the-cl* constraint. We have seen the addition of two new lexical items, one for *the* that is justified on the basis of the historical evidence presented in chapter 1 and one for *more*. Finally, there were a few constraints that have been added that do not specifically treat CCs but were created in opposition to the CC. For example, the existence of the category of standard head-filler phrases and its constraints, though not particular to the CC, would not exist were it not for the CC.

Turning now to the treatment of the whole CC and not just a single clause, Borsley introduces a new kind of clause (parallel to the *head-filler clause*) called the *correlative-clause*. Just like the *the-cl* was part clause and part head-filler phrase, the *correl-cl* is similarly part clause and part head-adjunct phrase. It has three subtypes in order to account for the three kinds of sentences Borsley is interested in: the (R)CC, the conditional, and the *as...so* construction. The three new phrase/clause types that go along with these are the *c-c-cl*, the *if-then-cl*, and the *as-so-cl*. Because *correl-cl* is an umbrella for the three others, its constraint just specifies that the CORREL value of the clause is *none* and that it has a verbal head (despite the fact that its head daughter may have a HEAD value that is a complementizer) and one non-modifier daughter (MOD *none*). The fact that the HEAD value of the whole *correlative* clause must be verbal is just a stipulation to unify it with other kinds of sentences. Borsley then separately posits constraints for each of the three sub-clause types saying that they must have two daughters with CORREL features of certain values (*if* and *then*, for example). This way,
he has four constraints for *correlative* clauses in English – one for the overarching
category and three for the different variants.

Then he defines a constraint on the *correl-cl* such that the head has a FFORM *the*
feature just like its daughter. Also, this constraint puts no restriction on the finiteness of
the verb form in the head, but it does require a finite verb in the non-head daughter. This
accounts for the pattern with imperatives, etc. in (B25).

Now, in order to account for the relationship of two *the-cl* in the CC, Borsley
distinguishes between standard head-adjunct phrases and non-standard ones and says that
the clause type *c-c-clause* is a non-standard head-adjunct phrase, while the other *correl-
cl* are standard ones. The reason that the *c-c-clause* will not match the constraint already
posited for standard head-adjunct phrases is that a *the* clause cannot stand on its own, but
when it combines with another, it can, cf. (B2: CC as S) and (B3: CC clauses don't stand
alone). (Though technically *then* and *so* clauses can't stand on their own either, if those
lexical items aren't present, they can, unlike the CC). That means that a crucial property
of that kind of phrase changes when it is modified, unlike what was required in the
earlier constraint (and what is observed in all other English adverb + sentence
arrangements). The way that he chooses to account for this is have the CORREL feature
value change from *the* to *none* when it is modified, so the full CC and RCC are unified in
being CORREL *none* at the sentence level.

As we have seen, the conditional, the *as...so* construction and the CC are all head-
adjunct phrases. Normally, heads are modified by adjuncts on the basis of the value of
the MOD (Modify) feature, and for Borsley, one CC clause will be the adjunct of the
other. Borsley says that the phrase starting with *the* in the RCC is the modifier. He says
that the value of the MOD feature for the the clause in the RCC is $S[imp\text{-}comp]$, which I take to mean that it would form a sentence with an implicit comparison in it. In the CC, the first the clause has the same value for the MOD feature, but the second clause has the MOD value none, since it is supposed to be the matrix clause that gets modified by the other the clause. The constraint on the clauses reflects this (the MOD feature is specified to be either $S[imp\text{-}comp]$ or none). This is only true of the constraint on the clause adverbials; if and as clauses will not have a specified restriction for the MOD feature to a certain kind of content. Borsley assumes that if and as are complementizers, which means that they don't require any special phrase types, but then the class of if, as, and the are not unified, since the isn't a complementizer (it's a specifier, as discussed above). Additionally, there is a head-adjunct phrase constraint (but this would be needed anyway, as in Pollard and Sag 1994) that says that the phrase that results when the sentence is modified by an adjunct is the same kind of phrase that the sentence was prior to modification. All of this leads to a lot of new kinds of phrases and clauses, constraints, and even features. An AVM that meets all of the relevant constraints appears in (79).
An AVM for an RCC appears in (80).
In this way of doing things (dividing the CC and the RCC and saying that only the RCC is a standard head-adjunct phrase), it seems mysterious that the clauses in an RCC cannot be reversed, as we saw in section 2.3. In order to meet this benchmark, Borsley introduces another LP constraint saying that whenever you have two clauses where one is marked CORREL and the other marked CORREL none, the CORREL none clause must come first. Thus, we have yet another linear precedence constraint required uniquely for the CC/RCC.

The remainder of this section is devoted to an explanation of how Borsley meets other benchmarks. It is worth noting that he gives his own characterization of the environments in which copula deletion can occur. He says that it is possible to omit the copula if: (a) its complement is fronted, (b) it is the main verb of the construction, and (c)
that is not present. All three conditions are met in (81), but (82a) violates the first, (82b)
and (82c) violate the second, and (82c) and (82d) violate the third.

(81) The more intelligent the students, the better the marks.

(82) a. *The more intelligent the students, the more marks given.
    b. *The more intelligent the students, the better the marks will.
    c. *The more intelligent the students, the better it seems that the marks.
    d. *The more intelligent that the students, the better that the marks.

Borsley has the following analysis of copula deletion in the CC. He assumes that it is a
special property of the verb *be* that it can be phonologically null. In other words, there is
a NULL feature for which *be* is the only verb to have a '⁺' value (NULL is a HEAD
feature). Then he says that head-filler phrases (and them only) can have a null head. The
lexical entry for this special form of the copula is NULL + and requires that the
complement is fronted and not in-situ (based on its requirements on the COMPS and
SLASH features). In addition to being ad hoc, this analysis only works because *the* and
*more* are not heads and so it will always be the verb (or complementizer!) that rules out
the right examples.

One of the most elegant features of Borsley's analysis is the way in which he
handles the possibility of tag questions and subjunctive morphology in the second but not
first clause of the CC, as per (B16) and (B21). Because *she reads* is the head of the
matrix clause and it is what will be a question or have subjunctive morphology, Borsley
doesn't have to say anything special to handle these because his rules are written so that a
phrase or clause has the same semantic and syntactic properties as its head daughter
unless a difference is required by some constraint, etc., which is similarly a feature of
other phrase structure grammars like GPSG. His analysis also predicts subject-auxiliary inversion in any clause of the CC or other correlative clause. Though McCawley and C&J both mention the historical possibility of this inversion in the matrix clause, it is certainly not possible in the subordinate clause, so it needs to be blocked from occurring in one or both clauses. The same is true for conditionals and other correlatives. For this purpose, he posits a feature INV, as is standard in GPSG and HPSG. If INV has a positive value, inversion should occur. Borsley predicts that clauses with if and as won't invert because if and as are complementizers and therefore not specified as having an INV feature (or are specified as -INV).

We have seen (in this section and the last) that Borlsey, via a set of phrases, restrictions on those phrases, features that those constraints make reference to, and linear precedence rules, accounts for nearly every benchmark set out in the previous sections. The only benchmarks that he neglects are (B8) – subordinate gapping, (B13,14) – negating a CC/RCC, (B15) – forming Y/N questions, (B19) – the flexibility of what appears in the first clause of the RCC, (B20) – the lack of measure phrases in CCs, (B22-24) – the extraction pattern of the CC, and (B25-28) – the verb form of CCs and RCCs in imperative and accusative-gerundive complements and the subject restriction on the CCs in those cases. Though this may seem like a lot of benchmarks remaining, this is by far the fewest remaining unaccounted for by any author.

2.4.3 Discussion

It's clear by now that Borlsey achieves broad coverage of the benchmarks. Of course, it is desirable to meet all of the benchmarks, but there are some, such as the
extraction pattern in the CC, that Borsley might reasonably believe to not be syntactic (perhaps related to general processing) that his theory, being a syntactic theory, should not try to address. Borsley's aim does not include a compositional analysis of the semantics of the CC, but if an analysis were put forth that had as broad a coverage of the syntactic benchmarks we've seen in this chapter as well as broad coverage of the semantic benchmarks and worked as a compositional system deriving the correct meaning for the CC, that theory would be preferable. A few other concerns are mentioned below.

The first issue has to do with the status of the FFORM feature. Borsley categorizes it as a NONLOCAL feature in keeping with other UDCs, but the non-locality of this feature isn't well justified (as Borsley notes) because the cannot appear in-situ. Next, recall from above that only *c-c-clauses* were restricted to an 'implicit comparison' value for the MOD feature; *if* and *as* clauses will not have a specified restriction for the MOD feature to a certain kind of content. Especially since he's trying to account for sentences like *I understand more as I read more*, which are nearly synonymous with CCs and RCCs like *I understand more, the more I read*, I would think that this would be a restriction that *as* and *the* would share. Perhaps Borsley would have a separate restriction for *as* that is explicit comparison, or perhaps since *as* can occur in sentences that are not synonymous with CCs, its wider distribution requires not limiting itself to an implicit comparison MOD value.

Another concern is that by introducing the CORREL feature for all root clauses, Borsley needs to find a way for all those root clauses other than CCs, conditionals and *as...so* sentences to be specified as CORREL none. Otherwise, overgeneration will
occur. In his articles, Borsley does not specify how this will happen, though it is clearly necessary. It remains to be seen how and how easily this could be done in HPSG.

The primary question for Borsley's analysis, however, is whether it is significantly better than a simple description of the patterns in the CC, which is what C&J predicted of any analysis of the CC in a contemporary syntactic framework. As mentioned earlier in this section, one could argue that Borsley's analysis is better than a description in that the machinery used to account for many of the CC's patterns is the same machinery used to account for other UDCs. This parallel is very clear within Borsley's system. I think that C&J still have a point, though, because I have highlighted throughout this section each piece of machinery that exists either only to handle the CC or to handle non-CC data but only because it is now in opposition to a feature value of the CC, etc. Each of these additions is justified by their necessity in meeting one or more of the benchmarks for the CC, but in some ways, that is is the point that C&J are making: the behavior of the CC is different enough from other English sentences that there will always need to be construction-specific additions to account for it. This foreshadows the unity between Borsley's analysis and a Construction Grammar-type analysis descended from some of the earlier works discussed here: 'constructions' are essentially a more complex kind of phrase structure rule.

2.5 Den Dikken 2005

Den Dikken argues that the similarity in CCs cross-linguistically, which he spends the bulk of the article articulating and defending, suggest that they are a universal structure with limited variability. His goal, like Borsley's, is to prove C&J wrong by
providing an analysis within a particular framework, in this case, Minimalism. As we will see, he falls short primarily insofar as he does not give a set of principles and parameters that his proposed structures are beholden to, and without them, it is difficult to determine whether his approach does capture the empirical patterns encoded in the benchmarks laid out in this chapter. His basic analysis appears in (83) (his (11)).

Some traits that this structure shares with the one in C&J are as follows: both are symmetric structures, both permit the overall category to be CP, both assign the CP category to both CC clauses, both consider the gapped sentence in each CC clause to be a C', and both locate the *the* and *more* somewhere in the specifier of a CP. This is largely the end of the similarities.

For den Dikken, the English *the* in the CC is a degree head, and it forms a constituent with the fronted comparative, which it takes as a complement. Many other positions in (83) remain unoccupied or occupied by a silent element in Modern English, but den Dikken proposes (83) as a universal structure that accounts for all languages across all time periods, and the example that den Dikken returns to again and again to
demonstrate a case where every position is filled is (84) (his (15)), from Lylyan.

(84) By how much the lesse he looked for this discourse, by so much the more he lyked it.

For Modern English, den Dikken says that neither P position is filled, both Q positions are filled by a silent/null element, both QP specifiers (Op and Dem, for 'operator' and 'demonstrative') are filled by null elements, both Deg positions are filled by the, and both APs contain a comparative form.

Beginning to look at (83) in more detail, we see that Borsley attributes a syntactic correlative structure to the English CC. He argues for this on the basis of historical evidence, giving examples where another relativizer, such as how much in (84), is present in the first clause (the would-be relative clause in the correlative sentence) but never the second. As we saw above, this is no longer true of English CCs because both can have a that relative, but den Dikken doesn't recognize that this weakens his point. For him, once a correlative, always a correlative, despite the apparent change in their syntactic properties. He does say that he can't explain why that appears in the second clause of English CCs, but he does not try to account for this, meaning that his theory cannot capture benchmark (B9: That in the CC). Also, since English “highest-subject” relatives without a wh word cannot omit that, as in (85), it's strange that for speakers like me, that is unnecessary in both clauses of the CC in which what is being compared is the subject.

(85) The people *(that) arrived late brought food.
Den Dikken provides an additional argument that CCs are like correlative clauses in that neither allows stacking; both are limited to two clauses (unless there is a conjunction of clauses in the matrix position). He further states that constructions that do not allow stacking cannot be paratactic. He does not comment on what they are instead.

Whereas Borsley focused on the microstructural benchmarks advanced by C&J and others but ignored the extraction facts they presented, den Dikken does the opposite. He discusses the extraction facts but not the data related to tag questions, etc. About extraction, he says simply that Dutch and German don't allow extraction in any of the cases where extraction is possible in English, and if we take them to be the well-behaved cases, then the fact that extraction is possible in English even though it also has this universal relative+matrix structure is due to the fact that English is generally permissive of extraction of nominals from islands (in other words, he is saying that all constraints on extraction are easily violable). He also makes an argument that what look like cases of ATB extraction in CCs are actually instances of parasitic gapping, which C&J already proposed themselves. Since these two kinds of gapping are only distinguished theory-internally, I will not consider this argument here.

As for the individual CC clauses, Den Dikken says that there are a variety of kinds of SUBCLAUSE (subordinate clauses) but that the fact that the CC is a correlative construction tells us that the SUBCL in the CC is a relative clause, which is why you see obligatory fronting. This is a nice aspect of den Dikken's theory given that you don't always see fronting in the matrix, e.g. in the RCC and so this explanation accounts for (B18: Lack of RCC 1st clause fronting).
Most of den Dikken's arguments relate to aspects of the microsyntax given that that is where most of the innovation in his theory as compared to C&J is located. Den Dikken thinks that the QP (formerly called the comparative specifier) such as so much not present in contemporary English is actually covert, which “explains” why no other DegP occurs and gives an element to contain the operator meaning in the construction. So any compositional meaning built on den Dikken's analysis would have null morphology contributing the wh operator-like meaning in both the CC and RCC. But then, the fact that much the better was a licit form historically is one of the reasons that den Dikken criticizes C&Js approach of positing that the is simply the specifier of a comparative like more. In other words, he is disputing the benchmark in (B20) that measure phrases and other kinds of comparative specifiers are not possible in CC clauses. Once again, however, den Dikken is making an argument on the basis of English forms that are not contemporary.

In looking more closely not just at the claim that QP contains a null morpheme but that the QP operator and demonstrative are similarly null, a reviewer rightly asked den Dikken how you could have a null demonstrative and still effectively demonstrate. He replied that this null demonstrative isn't the deictic demonstrative, it's more like another kind. I would guess that the reviewer would respond that these two senses seem to always co-occur in the same morpheme or word, asking why he would assume that they could be different, as he is proposing they must be (if it would be impossible to have a null deictic but not the one he is proposing). One place he does not posit a null element is in the preposition slot. This is because den Dikken has to assume that there is a null QP and null operator or demonstrative in order to account for why no other measure
phrase is possible, as per (B20: No MP in the CC), but there is not a similar requirement for the PP or preposition, so he assumes that that does not exist in a null form when it is not present. But in that case, what is den Dikken really saying about how the meaning of the CC is built? If the preposition is sometimes there and sometimes not, then no crucial part of the CC meaning should require it. But then how will he explain what contribution it makes when it does appear? This might be fine if the preposition had a non-existent contribution to meaning when it did occur, but as the examples show, the cases he is considering involve the difference between how much and by how much, which are significantly different in that how much requires a degree, while by how much requires a degree difference. It is unclear to me how den Dikken would resolve this issue.

There are two other concerns about the null operators. Whatever is in Deg' (the more) comes along with the null operator in Q (responsible for the fronting), but that makes it the only null operator in English that pied-pipes (according to him), which smacks of idiosyncrasy. Den Dikken does not think this is ad hoc because it's not in the kind of environment that the other null operators are in, and so it doesn't have the same restrictions. On one hand he's right, but on the other, this is still something that distinguishes the CC in contemporary English from any other sentence of English – that it has a null operator in an environment unlike any other. And though creating several new null lexical items is technically a lexical solution, it is a little misleading of him to conclude that there is nothing grammatically unique about the CC at the end of his paper, given all of these empty operators, etc.

As for the question of why you can't have a pronounced word or phrase where the null operator is in contemporary English CCs (since you don't find any), den Dikken
simply says that he can't answer that at this time. He thinks that Borsley is wrong in making a restriction limiting *the* to the left edge of the CC, however, because of examples like *so much the more* or *all the more*.

The degree head that hosts *the* in the CC also receives some attention. Den Dikken devotes a section of his paper to trying to justify that *the* isn't a standard of comparison like *than 3 ft.* but is instead a regular measure phrase like *3 ft.*. He claims that it cannot be a standard of comparison because those are not possible in CCs, but this is circular reasoning. Surely, if they (the *the*')s were standards of comparison, that would explain why other standards of comparison can't occur in the CC.

Despite some of my reservations about den Dikken's arguments (above), he is saying that the comparative should be the complement of the Deg0, in this case, *the*. That actually matches my analysis to a significant extent, as we'll see in Chapter 5. In general, then, there are two main types of problems facing this account. The first is identified by Abeillé and Borsley 2008, who criticize den Dikken for saying that he'll show how the CC can be handled straight-forwardly in a Principles and Parameters kind of syntax without ever stating which principles and parameters he takes to be involved in the CC. I agree that this is a major deficit, especially in combination with the fact that den Dikken does not specifically address how his analysis captures the patterns represented by most of the benchmarks presented in this chapter, instead concentrating on his own set of benchmarks developed on the basis of data from CCs in other languages. It is a dangerous methodological choice to base an entire analysis on the assumption that CCs have the same structure and meaning in every language. As an example of the kind of problem that a lack of parameter specification poses, let's look at
what den Dikken says (or doesn't) about adjunction. When he says that the two CC clauses are adjoined, either he's assuming that this process of adjunction is the same as other for other adjuncts, in which case he's making false predictions about the differences between C1 and other subordinate clauses (as noted in previous sections), or he's assuming that there is something special about this kind of adjunction, in which case he needs to specify what it is in order to see whether it does make the right predictions about the differences between C1 and other subordinate clauses. So either he doesn't have an analysis (as Borsley says), or his analysis is wrong. The second primary kind of shortcoming is that den Dikken often argues against a straw man theory rather than the actual claims of C&J. Though C&J are not clear about the way in which differences between parataxis and coordination are expected to play out, den Dikken doesn't mention this but spends most of his time in the paper trying to show that the CC isn't like a regular coordinate structure, but that's not something C&J would disagree with.

In fact, den Dikken 2005 inherits some of the problems for C&J on the basis of not being significantly different from C&J in terms of how the macrostructure of the CC is analyzed. Though den Dikken is explicitly saying that the CC is a remnant of correlative in English and C&J are not, C&J are saying that the CC is paratactic, and correlative structures such as where she goes, I will go there are some of the best examples of parataxis in contemporary English. As we saw in 2.3, however, the extraction pattern of these paratactic sentences do not match the extraction pattern of the CC and is actually its opposite. Den Dikken's analysis would not be able to explain this difference, but the extraction pattern of the CC was one of its most notable idiosyncrasies. To some extent, then, it seems that den Dikken is arguing that if you
ignore the data that make them look idiosyncratic, he can give an analysis of CCs that shows that they are not idiosyncratic. Another problem with his assumption that all CCs are actual instances of correlativization relates to the meaning of the CC. Den Dikken notes that correlativization is subject to a “strict matching constraint” in the semantics (which it is), but he fails to recognize that that constraint is one of co-reference, a strictness that the CC does not display, as we will see in the next chapter.

In his conclusion, den Dikken claims to have shown that none of the syntactic properties of the CC are idiosyncratic, only the lexical properties. However, even if that were technically true, there are still certain lexical items needed for the CC that are different from other lexical items in the language. Now, granted, that doesn't make the CC any more odd than the cleft construction or another sentence type where there are particular lexical items associated with a particular syntax, but den Dikken takes it a step further by saying that the lexicon is necessarily a compilation of sui generis entries because that's what makes them arbitrary. But this is not quite right – the point is that unless he's claiming that *the* is the only word in the CC that requires an extra lexical entry, then there are some extra entries for *more*, etc. like there are for Borsley, and even though the lexicon is arbitrary, it still has to be argued that the lexical items in the CC aren't specialized to that construction in a way that undermines their regularity of behavior in the CC as compared to elsewhere in language. Also, in the kind of Minimalism where Merge, Move and Agree are the only basic grammatical rules and other grammatical properties are contributed by lexical heads (thus, lexically-specified information), a language's grammatical properties depend upon what functional heads it has. As we saw above, den Dikken requires null functional heads for each clause of the
CC as well as to combine the two clauses (in the case where the CC is symmetric), meaning that there are phonologically null lexical items accounting for all of the idiosyncratic properties of the CC language by language. This contradicts den Dikken's claim that his approach doesn't require anything specific to the CC construction.

Abeillé and Borsley 2008 (A&B) have some critiques of den Dikken 2005 along this line as well. Borsley’s approach to the CC, as we saw above, proposed a lot of non-lexical idiosyncratic grammatical categories/constraints to account for differences between constructions of English, and the same is true for A&B’s analysis of CCs cross-linguistically if the analyses of each language were compared to one another. Den Dikken represents the near-opposite position that all inter-language variation is lexical. A&B point out that if it is similarity to a free relative that gives the first clause of a CC its properties (such as copula deletion), then it remains a mystery how we can explain the fact that the second CC clause has those same properties despite not being a free relative. Den Dikken notes for Dutch (where both clauses can have the complementizer after the filler, as in English) that the second clause seems to 'mimic the internal structure' of the other clause, but he doesn't spell out how this would be done. A&B speculate that any way to do this via parameters would lead to the creation of parameters that are specific to this construction, which, naturally, den Dikken is trying to avoid.

2.6 Conclusion

In this chapter, we have seen a variety of descriptions of the syntactic pattern of the CC as well as a few theories, concrete to different degrees. We saw arguments from construction grammarians and their sympathizers as well as theories proposed by non-
construction grammarians working in a variety of syntactic frameworks. In the end, none of the theories met every benchmark, but some fared better than others. Borsley, in particular, fared best in terms of empirical coverage. His is also the theory in which the most machinery was added in order to account for the CC, but this could just be because some of the others were not as detailed when it came to explaining how patterns would be accounted for. It is debatable whether this supports the basic thrust of the construction grammar argument that any analysis of the CC will necessarily be so ad hoc as to not be any better than a simple description of the phenomena. The benchmarks, re-listed together below, provide that description, and the analysis in Chapter 4 is my own attempt to give an analysis that is superior to just listing them. I now turn to describing the semantics of the CC, providing a second set of benchmarks against which one can measure an analysis of the CC.

2.7 Syntactic Benchmarks

(B1) **CC as UDC**: Each CC clause contains an unbounded dependency and is therefore subject to island constraints such as the Complex Noun Phrase Constraint, the Empty Category Principle, and the Strong and Weak \textit{wh}-island constraints.  
(B1.1) **Major Category Connectivity**: The syntactic category of the filler and gap must match.  
(B1.2) **Number Connectivity**: The number of the filler and gap must match.  
(B1.3) **Idiom chunk Connectivity**: idiom restrictions from the gap constrain the identity of the filler. 
(B2) **CC as S**: CCs are of the same syntactic category as other sentences of English. 
(B3) **CC clauses don't stand alone**: whenever one the more constituent occurs, a second the more constituent must follow to form a sentence. 
(B4) **Mandatory** The: The is required phrase-initially in both main CC constituents. 
(B5) **Mandatory Fronting**: the comparative (and whatever part of the filler it brings with it) cannot appear in situ in either CC main constituent. 
(B6) **No the in the 1st clause of the RCC**: In the RCC, no the can appear initially in the first main constituent.  
(B7) **Backward Pronominalization**: Pronominalization is possible in either direction in the CC (between clauses).
(B8) **Subordinate Gapping**: The CC, like other subordinate-matrix structures, does not allow gapping between its clauses.

(B9) That in the CC: That can optionally appear before the gapped clause in a CC clause.

(B10) **Comparative Variety in the CC**: Any form of the comparative can occur in the CC.

(B11) **Obligatory Comparative**: Each CC clause requires a comparative form.

(B12) **Copula Omission**: A conjugated form of to be may be omitted at the end of a CC clause as long as the subject it would normally follow is not specific.

(B13) **Negating a CC**: The only way to negate the sentence-level meaning of the CC is by saying It is not the case that X where X is a CC.

(B14) **Negating an RCC**: An RCC's sentence-level meaning can be negated in the same way as the CC's or by negating the main verb in its first (matrix) clause.

(B15) **Yes-No Questions**: Only the first clause of the RCC can be used to form yes-no questions.

(B16) **Tag Questions**: Tag questions can be formed on the basis of the second clause of a CC or first clause of an RCC but not their other clauses.

(B17) **Than-phrase Restriction**: CCs cannot occur with than phrases in either clause.

(B18) **Lack of RCC 1st clause fronting**: The comparative is not fronted in the first main constituent of the RCC.

(B19) **RCC Flexibility**: The first clause of the RCC can take any syntactic form as long as it corresponds to the right kind of meaning (semantic term) to be combined with its second clause, whose syntactic form is fixed/does not show similar flexibility.

(B20) **No MP in the CC**: Overt measure phrases are not possible in (R)CC clauses.

(B21) **Subjunctive**: Both CCs and RCCs are possible in subjunctive contexts, but the verb in the clause corresponding to what is semantically the consequent is the only one that receives subjunctive morphology reflecting this fact.

(B22) **C1 Extraction**: Extraction can occur from the first clause of the CC.

(B23) **C2 Extraction**: Extraction can occur from the second clause of the CC.

(B24) **C1 & C2 Extraction**: Extraction can occur from both clauses of the CC at once.

(B25) **Imperatives**: The clauses in the CC and RCC that correspond to semantic antecedents can take on an imperative verb form when in imperative contexts, while their other clause cannot.

(B26) **Subjects with Imperatives**: In order for a CC clause to take on an imperative form, it must have a subject.

(B27) **Acc-Gerundives**: The clauses in the CC and RCC that correspond to semantic antecedents can take on an accusative-gerundive form while their other clause cannot.

(B28) **Subjects with Acc-Gerundives**: In order for a CC clause to take on an accusative-gerundive form, it must have a subject.

(B29) **Preposition Pied-Piping**: The CC does not permit prepositions to be extracted along with the more.
This chapter is devoted to semantic benchmarks and analyses of Comparative Correlatives. As in the last chapter, I move through the literature chronologically and anytime a set of data is presented that shows a pattern that any theory would need to capture, I introduce this pattern as a numbered benchmark (B1), (B2), etc. so that they stand out from regular example numbers. Not all of the benchmarks presented in a given section are observations of the author under review. I will try to be clear about which are theirs and which are mine. The chapter concludes with a list of these benchmarks repeated together in one location.

3.1 Earliest Analyses of the Semantics of CCs

3.1.1 Thiersch 1982

Though the sentence in (1) is ruled out by benchmark (B4: Mandatory the) because its clauses do not begin with the, I include this here just to remind us that the is not simply a determiner filling a slot that another determiner might be substituted into, nor should it have the meaning typically attributed to the definite article. ¹⁸ Thiersch made the observation that (1) is not acceptable in his discussion of the history of the for the purposes of justifying his syntactic proposal, but it is also relevant as we begin to describe the meaning of CCs.

¹⁸ This does not exclude the possibility that both the’s encode semantic definiteness since demonstratives, pronouns, etc. are all semantically definite (Roberts 2003).
Thiersch is the first to describe the meaning of the CC despite its not being his central concern. Just as in the syntax, he is interested in drawing a parallel between CCs and comparatives. He says that the meaning of *John is taller than Mary* involves an operator EXCEEDS with two arguments – the extent to which John is tall, and the extent to which Mary is tall (this returns 'true' in case the first extent does in fact exceed the second extent). Similarly, he argues that the CC has an operator called IS-PROPORTIONAL-TO that takes two extents, so for the example *The harder John tries, the more often he succeeds*, the two extents are the extent to which John tries hard and the extent to which he succeeds. The exact meaning of the operator in the CC remains a mystery (there is no meaning postulate or other indication as to when IS-PROPORTIONAL-TO returns 'true'), so there's no way to know whether it would accurately capture the truth conditions of the CC.

One of the things that is interesting about this characterization, however, is that it reveals a common source of confusion. One one hand, you have two comparative morphemes in the CC, one in each clause. That would suggest that there are two comparisons happening in the CC. But neither of these relates to the connection between the two clauses (or if one did, what would the presence of the other indicate?). Nevertheless, there is a sense in which degrees are compared across the two clauses, but then what part of the sentence is responsible for that comparison? Thiersch ignores the contribution of the comparatives in his characterization of the CC's meaning, focusing on the operator that relates them, but this is, in a sense, a third operation comparing degrees

(1)*A better the player, a better his game.
that is at work in the CC. We will see below that other theorists have accounted for these in very different ways, often not mentioning the possible confusion.

3.1.2 Fillmore 1987, Fillmore, Kay, and O'Conner 1988

Fillmore 1987 is the first article to describe semantic patterns of the CC. He assumes that the CC asserts a dependency relationship between the values of two scales, where each clause sets up one of the scales. In other words, he takes the meaning of C1 to denote an independent variable and the meaning of C2 to denote a dependent variable. He does not give a formal representation of this meaning, and this is all he says about it.

Fillmore then claims that tense and modality combinations within the clauses lead to different meanings in the following way: past tense in both clauses yields a generalization about the past; simple present in both clauses yields a general principle (that should be timeless); and simple present in the left-hand clause combined with future will in the right-hand clause yields a contingent prediction. His examples (9)-(11) appear in (2).

(2) a. The more he drank, the sillier he got. (Past Generalization)
b. The more he drinks, the sillier he gets. (General Principle)
c. The faster we drive, the sooner we'll get there. (Contingent Prediction)

Fillmore shows that the same pattern is true of the conditionals as in (3).

(3) a. If he robbed someone, then he went to jail. (Past Generalization)
b. If you rob someone, you go to jail. (General Principle)
c. If you rob someone, you will go to jail. (Contingent Prediction)
A further parallel that Fillmore mentions and subsequent literature has picked up on is that in both conditionals and CCs, the future tense interpretation of *will* is not usually possible in the antecedent clause.\textsuperscript{19} To that, I add the RCC:

(4) If you (*will) cook, then Mary will eat what you make.

(5) The more food you (*will) cook, the more Mary will eat

(6) Mary will eat more, the more food you (*will) cook.

(B30) *Future tense: CCs and RCCs do not occur with future will in their antecedents.*

Fillmore notes one further semantic parallel between conditionals and CCs in the form of permitting negative polarity items (NPIs) in their antecedent clause, as (7) and (8) demonstrate (I have changed the negative polarity items, etc. in these examples in order to make a stronger point). The RCC data are added in (9). NPIs include *any, ever, hope in hell* and *the least bit of*.

(7) a. If she has the least bit of fun, she'll smile.
    b. #If she smiles, then she's had the least bit of fun.
    c. If he ever thinks about peace, he might act on it someday.
    d. #If he experiences war, then he might ever think about peace.

(8) a. The more she has the least bit of fun, the more she'll smile.
    b. #The more she smiles, the more I can tell she's having the least bit of fun.
    c. The more people that ever think about peace, the more peaceful the world will be.
    d. #The more wars we see, the more we will ever think about peace.

(9) a. She'll smile more, the more she has the least bit of fun.
    b. #I can tell she's having the least bit of fun, the more she smiles.
    c. The world will be more peaceful, the more people that ever think about peace.
    d. #We will ever think about peace more, the more wars we see.

\textsuperscript{19} Though see Kaufmann 2005 for exceptions to this generalization for the conditional, such as (i), which equally apply to the CC, as in (ii) (and thus are not relevant here).

(i) If the worker will be running errands, ask to see driver's license and proof of insurance. [NYT 08/08/94]
(ii) The more the worker will be running errands, the more important it is that s/he has a driver's license.
(B31) *NPIs in the CC*: Negative Polarity Items can occur in the antecedent but not consequent clause of the CC and RCC.

For Fillmore, these parallels join the syntactic evidence in proving that CCs are a kind of conditional. Fillmore says that “it seems clear from [these facts] that any successful description of these phenomena require us to assemble *for this single construction* an organized body of facts which cannot simply be taken for granted as following from other facts independently knowable from the grammar of English” (1987:166). This conclusion about the CC is based on the fact that he thinks that conditionals are similarly a sui generis construction, and so the parallels argue for the same treatment for the CC.

It is certainly true that conditionals exhibit certain properties that distinguish them from other kinds of sentences, such as the presence of *if*, but contemporary theories agree that conditionals should be able to be accounted for by the same rules and lexical items (with the exception of lexical items that only occur in the conditional) as any other sentence of the language (cf. von Fintel 1994 and subsequent). Thus, the parallel that Fillmore brings to light between conditionals and CCs, his primary evidence for both being constructions, would now be considered evidence that CCs, like conditionals, should be able to be given a generative analysis. And indeed, giving the CC a conditional analysis is exactly what Beck 1997 does. Interestingly, then, C&J go back to arguing that CCs are somehow beyond the bounds of what the grammar can account for while maintaining that conditionals fit within those same bounds.

Further support for the idea that both conditionals and CCs can be accounted for by a general mechanism at work in natural language that extends to other phenomena
comes from other kinds of sentences that Fillmore does not mention. These include
generic sentences and sentences with free relative clauses. The three patterns listed by
Fillmore for CCs are shown in (10)-(12) and (13)-(15) for these constructions. I have
illustrated the generic with plural kind NPs, but the same is true of generic indefinites.

(10) a. Dodos that were born in the summer had colorful beaks.  (Past Generalization)
    b. Pukekos that are from New Zealand have blue feathers.  (General Principle)
    c. Kiwi that habituate to the light will increasingly be in danger. (Contingent
       Prediction)

(11) Cats that (*will) mate in the wild (will) have higher fertility rates.

(12) a. Men who have a hope in hell of making it are educated.
    b. #Men who are entrepreneurs have a hope in hell of getting an education.
    c. Men with any sense avoid installment schemes.
    d. #Men with sense avoid any installment schemes.

(13) a. Whatever she did, I did, too.  (Past Generalization)
    b. Whatever she does, I do, too.  (General Principle)
    c. Whatever she does, I will do, too. (Contingent Prediction)

(14) However often he (*will) greet(s) the train, that's how often I will greet it.

(15) a. Whenever she has a hope in hell of winning, she loses her poker face.
    b. #Whenever she is winning, she has a hope in hell of keeping her poker face.
    c. Whenever someone has any sense, they move to Ohio.
    d. #Whenever someone is in their right mind, they move to any Midwestern city.

Focusing now on the NPI data, I will show that the CC is slightly different from any of
these constructions in a couple of respects. Ladusaw 1979 is the seminal reference on
NPIs and their licensing. In contemporary terms, his proposal shows that operators have
an ability to license negative or positive polarity items (PPIs) in any of their arguments
on the basis of entailment properties between the sub- and superset relations of
extensions of the operator's scope in that argument. Specifically, if there is a downward
entailing/monotone decreasing (DE/MD) relationship (closure under subsets), then an NPI is licensed, and if there is an upward entailing/monotone increasing (UE/MI) relationship (closure under supersets), then a PPI is licensed (actually, a PPI may be licensed anytime there is not a MD/DE relation). In (16), the entailment goes from the subset to the superset in the right argument, making it MI/UE on the right, and on the left, the entailment goes from the superset to the subset, making it MD/DE on the left. In (17), the entailment goes from the superset to the subset in both arguments, making it MD/DE in both. The data in (18) and (19) show whether NPIs and PPIs (such as somewhat) occur in each argument position.

(16) a. All basketball players eat broccoli.
    b. All basketball players eat green vegetables.
    c. All people eat green vegetables.

(17) a. No basketball players eat broccoli.
    b. No basketball players eat green vegetables.
    c. No people eat green vegetables.

(18) a. All basketball players with any sense eat broccoli.
    b. All basketball players who are somewhat smart eat green vegetables.
    c. *All basketball players have any sense.
    d. All basketball players are somewhat smart.

(19) a. No basketball players with any sense eat broccoli.
    b. No basketball players who are somewhat smart eat green vegetables.
    c. No basketball players have any sense.
    d. *No basketball players are somewhat smart.20

Here, (18c,d) and (19c,d) pattern as we would expect on the basis of entailment patterns. Because all is UE in the right argument, a PPI is licensed, but because it is not DE, the NPI is not licensed. The reverse is true for no because it has the reverse monotonicity.

20 This sentence could be felicitous as a denial. If somewhat is intonationally prominent in a certain way, this sentence could be taken to mean that the speaker rejects another's claim that basketball players are somewhat smart and wishes to imply that they are extremely smart, etc.
pattern of all in the right argument. The examples in (18a,b) and (19a,b), however, do not fit the pattern predicted by Ladusaw, since both all and no are DE on their left argument. This means that the left argument should license NPIs but not PPIs. As we see, however, both are licensed. A similar puzzle for Ladusaw is noted in Heim 1984. She shows that antecedents of conditionals, generic sentences and most are all cases of non-DE environments that license NPIs anyway (she develops a story in which strengthening inferences from context may license NPIs in these cases; for more recent accounts, see e.g. Gajewski 2010). Many theories of conditionals such as Kratzer 1991 take the conditional to have a default generic operator that takes the antecedent and consequent as arguments, which would predict the parallel between conditionals and generics. Similarly, the generic and most are often lumped together in a single category insofar as both have meanings that require a majority but allow exceptions (i.e. are not universals) and are difficult to pin down exactly (is most 60%? 95%? does it depend on context? etc.). In the next few pages, I will argue that most can be distinguished from conditionals and generics in terms of its entailment pattern and, more importantly, that the CC patterns with most rather than the conditional in this respect. I will make the same argument on the basis of another phenomenon in a later section.

Most is non-monotonic (NM), meaning that it is neither DE/MD nor UE/MI, in its left argument (but licenses NPIs, as per the puzzle mentioned above) and upward monotonic in its second argument, licensing PPIs, as (20) and (21) show. To see that its left argument is NM, consider the following models. In the first model, no basketball players eat green vegetables, but they are a minority of the population as a whole, and the rest of the population does eat green vegetables. In that case, (20c) would be true, while
(20b) would not, showing that the superset sentence's meaning does not entail the subset sentence's meaning. In the second model, the basketball coach makes all of his players eat green vegetables to stay healthy, but a minority of the general population eats them. In this case, it is (20b) that is true, while (20c) is not, showing that the entailment doesn't hold from the subset to the superset, either. As for the right argument, (20a) entails (20b) but not vice versa.

(20) a. Most basketball players eat broccoli.
    b. Most basketball players eat green vegetables.
    c. Most people eat green vegetables.

(21) a. Most basketball players with any sense eat broccoli.
    b. Most basketball players who are somewhat smart eat green vegetables.
    c. *Most basketball players have any sense.
    d. Most basketball players are somewhat intelligent.

In (21), Ladusaw would predict that only (b) and (d) would be grammatical, but despite the lack of monotonicity in the left argument of *most, both NPIs and PPIs are licensed there. Turning to the conditional, we see the same NPI/PPI pattern as with *most (compare (21) to (23)).

(22) a. If Evan is a men's basketball player, then he eats broccoli.
    b. If Evan is a men's basketball player, then he eats green vegetables.
    c. If Evan is a man, then he eats green vegetables.

(23) a. If basketball players have any sense, then they eat broccoli.
    b. If basketball players are somewhat smart, they eat green vegetables.
    c. *If basketball players eat green vegetables, then they have any sense.
    d. If basketball players eat green vegetables, then they are somewhat intelligent.

At first glance, the pattern of entailment in the conditional looks like it is MD in the left argument, which would make it not surprising that NPIs are licensed there. Whenever
(22c) is true, (22b) will necessarily be true. Heim shows, however, that some example pairs with conditionals, such as (24) are not MD, which means that the antecedent of a conditional is not necessarily MD across all examples, which makes it NM.

(23) a. If a tiger has stripes, it is orange.
   b. If an albino tiger has stripes, it is orange.

In summary, both most and the conditional are NM in their left argument. What I would like to draw attention to is the fact that the determination of non-monotonicity was made differently in the two cases. In the case of most, we could construct models for any example sentence in which the entailment would not hold in either direction. I will call any argument position determined to be NM in this fashion NM-1. In the case of the conditional, we had to construct a particular kind of example that differed from others in being non-MD. I will call any argument position determined to be NM in this fashion NM-2. Returning now to the CC, we saw in (8) that it patterns like both most and the conditional in licensing NPIs in its left but not right argument position. Given this new distinction between most and the conditional in terms of how NM is determined, we can now see whether the CC is NM-1 or NM-2 (or something else altogether).

(24) a. The more women you know, the more popular you'll be.
   b. The more people you know, the more popular you'll be.

Beginning with the left argument since that is the one that draws out the difference between most and the conditional, we want to determine whether there is an entailment between (24a) and (24b) in either direction. Consider the models below.
Table 3.1: Model for (24)

<table>
<thead>
<tr>
<th>Person</th>
<th># of women</th>
<th># of men</th>
<th># of people</th>
<th>popularity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Person 1</td>
<td>14</td>
<td>20</td>
<td>34</td>
<td>unpopular</td>
</tr>
<tr>
<td>Person 2</td>
<td>16</td>
<td>49</td>
<td>65</td>
<td>not very popular</td>
</tr>
<tr>
<td>Person 3</td>
<td>23</td>
<td>17</td>
<td>40</td>
<td>somewhat popular</td>
</tr>
<tr>
<td>Person 4</td>
<td>40</td>
<td>13</td>
<td>53</td>
<td>popular</td>
</tr>
<tr>
<td>Person 5</td>
<td>47</td>
<td>31</td>
<td>78</td>
<td>most popular</td>
</tr>
</tbody>
</table>

Table 3.2: Alternative Model for (24)

<table>
<thead>
<tr>
<th>Person</th>
<th># of women</th>
<th># of men</th>
<th># of people</th>
<th>popularity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Person 1</td>
<td>14</td>
<td>20</td>
<td>34</td>
<td>unpopular</td>
</tr>
<tr>
<td>Person 2</td>
<td>23</td>
<td>17</td>
<td>40</td>
<td>not very popular</td>
</tr>
<tr>
<td>Person 3</td>
<td>40</td>
<td>13</td>
<td>53</td>
<td>somewhat popular</td>
</tr>
<tr>
<td>Person 4</td>
<td>16</td>
<td>49</td>
<td>65</td>
<td>popular</td>
</tr>
<tr>
<td>Person 5</td>
<td>47</td>
<td>31</td>
<td>78</td>
<td>most popular</td>
</tr>
</tbody>
</table>

The model in Table 1 makes (24a) true. We see that in all cases where one person knows more women than another, the person who knows more women is more popular. We can also see that (24b) is not true in this model, since Person 2 knows more people than Person 3 or Person 4 but is less popular than either of them. The model in Table 2, on the other hand, makes (24b) but not (24a) true. So, the left argument of the CC is NM, and given that we could construct models such as those in Tables 1 and 2 for any CC, it is NM-1 like most. Despite its non-monotonicity, it licenses not only NPIs as in (8) but also PPIs in that argument position, as (25) demonstrates.

(25) The more colors John somewhat likes, the more difficult his decision.

Moving on to test the right argument position of the CC, we find that, unlike any of the earlier quantifiers and constructions that it patterned with in other ways, the CC is NM-1.
In (26), it is once again possible to construct models that make either (a) or (b) but not both true. In fact, if we reinterpret the column headings in Tables 1 and 2 to indicate the number of people one will meet rather than the number of people that one knows, then the model in Table 1 verifies (26a) but not (26b), and the model in Table 2 verifies (26b) but not (26a). Thus, the CC is NM-1 on both arguments. This conclusion is especially interesting given the discussion of symmetry v. asymmetry in syntactic and semantic form that we've seen in the discussion of the literature so far. With respect to entailment patterns (but not actual NPI licensing), then, the CC is symmetric, unlike the conditional.

Going back to the actual NPI/PPI occurrence in the CC, its entailment behavior predicts a lack of any NPI in either argument of the CC if Ladusaw is right about the connection between monotonicity and polarity licensing. Fillmore noted that NPIs were licensed in the left but not right argument, as in (8). It seems to me, however, that there are sentences in which an NPI occurs in the right argument of a CC that sound grammatical. In (27), I give two examples of CCs with NPIs in the consequent, but (27a) seems more natural than (27b).

(27) a. The happier people are, the less often they are the least bit tired.
    b. *The richer people are, the more often they are the least bit sad.

One noticeable difference between (27a) and (27b) is that the one where the NPI is okay contains a less rather than a more. Thus, one question we should ask is whether less may be a negative (DE) operator itself that licenses the NPI in (27a). Hoeksema 1983 does in
fact claim that comparatives license polarity items, but only in the than clause of clausal comparatives (as opposed to phrasal comparatives), and no difference between more and less is noted. I include the data in (28) for the reader to judge, but I do not see a difference in acceptability between more and less on a large scale, ruling that out as a potential explanation for the difference in (27). This is all I will say about the pattern in (27) here, as I leave it for future research.

(28) a. Jane, though ugly, has more of a chance of getting the job than women who are the least bit pretty (because she is more talented than they are).
   b. Jane, though pretty, has less of a chance of getting the job than women who are the least bit intelligent (because she's so dumb).

Moving onto PPIs in the CC, they seem to be licensed in both arguments, as in (29).

(29) a. The more kids that somewhat like green veggies, the healthier their generation.
   b. The more kinds of ice cream he tries, the more he somewhat likes ice cream.

This is another respect in which the CC's pattern is a bit odd. In the other argument positions that are non-monotonic (of the constructions that we've considered here), both NPIs and PPIs are licensed. Given that the CC is NM in both its arguments, that would suggest that the CC should be able to occur with NPIs or PPIs in either clause. Nevertheless, in at least some cases such as (8), NPIs in the consequent are not good.

One final note about polarity item patterns concerns other constructions. In the interest of space, I will not present examples for the full range of constructions here, but I want to point out that generic sentences pattern like conditionals and that the adverb of quantification usually patterns like most with respect to all of these polarity patterns. A chart summarizing the characteristics of these constructions appears in Table 3.
The fact that CCs pattern on the one hand like so many other constructions and on the other unlike anything other than *most* and *usually* argues against both of Fillmore's main claims: that the CC is a kind of conditional construction, and that both of these constructions are too unique to be given a generative grammar account. Of course, just because the CC and conditional share a pattern with a handful of other constructions, it does not mean that he is necessarily wrong about the second argument, but at a certain point, a pattern that is repeated throughout a language is likely one that is not specific to one construction and inherited by others.

Fillmore, Kay, and O’Connor 1988, following Fillmore 1987, state that a CC “expresses a correlation between an independent variable and a dependent variable,” and they give the sentence *the more carefully you do your work, the easier it will get* a paraphrase of 'the degree to which you do your work carefully will determine the degree to which your work gets easy' (1988:506). Beck 1997 shows why this cannot be right, to which I will add an additional piece of evidence in the next section. Fillmore et al. also state that the CC is “free of pragmatic commitments” and assume that each of its two clauses is propositional. Given that propositions correspond to sentences in the syntax, it
would be surprising for each clause to be propositional given that neither can stand alone. As for what is meant by 'pragmatic commitments', it is hard to tell, but the subsequent sections will show the ways in which contextual factors strongly influence particular parts of the meaning of CCs.

3.2 Beck 1997

Beck 1997 is the seminal work on the semantics (and syntax-semantics interface) of CCs and RCCs. She argues that CCs are a kind of conditional and that the comparative morphemes in the CC are the same as those we see in other comparative sentences, and she gives the CC a compositional analysis based on these two qualities. Though she illustrates most of her points with German CC and RCC sentences, she explicitly says that German and English CCs should receive identical analyses. Thus, throughout this section, I will consider her theory as it applies to English. This section is organized as follows. I begin by discussing some of the data that Beck presents (or their English translations). In subsection 3.2.1, I go through her analysis in detail, and in subsection 3.2.2 I present some of my thoughts and observations about this analysis.

Beck first notes the ways in which CCs are not identical to conditionals in terms of their meanings. Conditionals as a whole do not necessarily express something comparative in each of their clauses (so they have a wider variety of meanings), and when they do contain comparatives, than phrases are present, unlike the CC (as we saw in (B17: Than-phrase Restriction)). On the basis of these generalizations, she claims that the CC is a restricted kind of conditional meaning. In the conclusion, she adds another difference, saying that the main difference between CCs and conditionals is that CCs
require quantification over at least two variables and conditionals, only one (this will be explained in more detail later). Once again, however, as long as conditionals are not argued to preclude quantification over multiple variables, one could continue to argue that CC meanings are a subset of conditional meanings.

To this set of differences, I would like to add a further case of the conditional having a less restricted meaning than the CC. Many authors give conditional sentences like (30) as paraphrases for (31) without recognizing that they are not equivalent.

(30) If you run faster, then you're more likely to win.

(31) The faster you run, the more likely you are to win.

Only (30) can be interpreted as a statement about one particular faster speed rather than as a generalization about any possible faster speed, however. We can easily construct situations in which (31) would be false but (30) true, showing a lack of synonymy between the two sentences. Suppose we are discussing marathons. In running a marathon, you actually don't want to run too fast or you'll run out of steam too early and not finish the race, so (31) is false because after some cutoff speed, you jeopardize finishing, let alone winning. But it could be that you are talking to a friend before the race and you know her normal marathon pace and also that she is capable of running a little faster without dropping out of the race. In this case, you can utter (30) and have it mean that if she runs one half of a mile faster per hour, she is more likely to win the marathon, without committing yourself to any statements about any other faster speeds or the truth of this statement on any other day or for any other race. In other words, if you uttered (31), you would be suggesting that running the fastest speed possible is the best
way to win, in this race or any other, whereas uttering (30) merely suggests that there is some faster speed that would lead to better results but not that it is the fastest she is capable of running. This is, in my opinion, the primary meaning difference between the CC and the conditional.

We are now ready to dive into the sentence-level meaning of a CC. For the example CC in (32), the meaning in prose according to Beck is given in (33) and a more formal representation of hers in (34) (this isn't an actual example of hers, but it is consistent with the way she describes all the other examples).²¹

(32) The faster we drive, the sooner we’ll get there

(33) For all hypothetical situations (possible worlds), if we drive faster in situation 1 than in situation 2, then we will get there sooner in situation 1 than in situation 2.

(34) \[\forall w_1, w_2 [\text{We drive faster in } (w_1) \text{ than } (w_2) \rightarrow \text{We get there sooner in } (w_1) \text{ than } (w_2)]\]

In (34), we see universal quantification over two possible world arguments. Beck notes that quantification is possible over three different kinds of variables: individuals, times, or worlds, as in (35)-(37) (Beck's (15)-(17)).

(35) The slimier an attorney looks, the more successful he is. (individuals)

(36) The hotter it was, the more tired Uli was. (times)

(37) The better Otto is prepared, the better his talk will be. (worlds)

(B32) Quantificational Variability: CCs and RCCs have meanings that can correlate degrees with respect to a number of different natural language types, including worlds, times and individuals.

²¹ As we will see throughout this section, Beck also consistently mixes object language and metalanguage in the same terms, which she considers to be easier for the reader.
As we will see when we look at the specifics of her analysis, she will claim that *the* is ambiguous depending upon which of these kinds of variables is quantified over. Her ambiguity-based approach yields an unwelcome result in that it predicts that we shouldn't be able to find examples quantifying over something besides individuals, times/time intervals or worlds. This prohibition includes combinations of these basic variables (such as 'stages' of an individual, which involve both times and individuals). If we quantify over a single type of variable, the others presumably remain fixed in value. But we can construct many examples where the quantification applies to pairs of individuals and times, individuals and worlds, etc. One of these cases is mentioned by Beck as a case where indefinites receive wide scope, as in example (38) and her meaning for it as given in (39).

(38) The slimier an attorney gets, the more successful he becomes.

(39) \[ \text{Gen } x \ [\text{attorney}(x)] [\forall t_1, t_2 \ [x \text{ is slimier at } t_1 \text{ than at } t_2] \Rightarrow [x \text{ is more successful at } t_1 \text{ than at } t_2]] \]

Here, we see that she takes the indefinite *an attorney* to be bound by a generic operator, which has wide scope with respect to the universal quantification over times. We will return to this issue more later, but for now, this represents what Beck considers the fourth kind of meaning possible in a CC. It also shows how she intends to cover cases in which there is a combination of types of variables within her system (so that these are not counterexamples to what it is capable of handling).
The difference between (39) and an analysis of (35) along the lines of (34) (but with individuals), for example, raises another issue. That issue is that many CC examples can be understood in multiple ways, so there are two or more readings of a single CC. An example appears in (42). In the context in (40), the natural interpretation quantifies over students (individuals) at a single time (their age/height on the first day of school), whereas in the context in (41), the most natural interpretation is one that compares a child's height across times. Per Beck's (39), (41) does seem to also assert that this would be true for other children besides the little boy in question. If we substituted the boy's name for a child in (42), that reading would be ruled out.

(40) Two teachers talking to each other after the first day of class in a new school year:
   A: It's a funny coincidence that the girls are taller than the boys in my class.
   B: Oh really? With my students...

(41) A mom and her mother-in-law talking about their son/grandson:
   A: Can you believe how tall your little boy is now?
   B: Well, the doctor keeps telling me that at this age, ...

(42) The older a child is, the taller he is.

So far, we've established that CCs can range over individuals, times, worlds, and combinations thereof, and a single sentence can have multiple readings with respect to what is quantified over (which are typically disambiguated by context). This parallels what is known about the conditional. The basic range of quantification similar to that in (35)-(37) is shown for conditionals in (43)-(45) below.

(43) If a player can shoot, he can't dribble. (individuals)
(44) If Turner shoots, he scores. (times)

(45) If he were to go pro, he wouldn't graduate. (worlds)

Beck continues by introducing other similarities between conditionals and CCs. One relates to anaphora. Let's assume, as most of the literature does, that there are only four main types of anaphora in English: deixis, referential anaphora, bound variable anaphora, and e-type anaphora. Consider some of the properties of these different types. Because they are all 'anaphoric', what they have in common is a relation between two lexical items or phrases across one or more sentences where one item is dependent on the other for its interpretation. In order to see how they differ, refer to the following three examples, each of which illustrates one of the types named above.

(46) Mary is lucky, and she is happy.

(47) Every man won with the horse he bet on.

(48) If a farmer owns a donkey, he beats it.

Ignoring deictic uses of pronouns (when *she* would be used while pointing to a person not mentioned, for example), all of the pronouns in sentences (46)-(48) are anaphoric in one of the remaining three ways. *She* in (46) is referentially anaphoric, *he* in (47) is a bound variable anaphor, and *it* in (48) is an e-type anaphor. In the case of (46), *Mary* refers to *mary' and *she* refers back to *Mary* in order to retrieve *mary' as its meaning as well. This simple co-reference is not possible in (47), where there is no particular man mentioned. (47) is a statement about all men (within a certain domain), and *he* needs to pick up as its referent each of those men. So neither *every man* nor *he* refers to a
constant of type $e$; instead, the meaning of every man contains a variable ranging over men and he retrieves as its discourse referent that variable. Thus, the meaning of he in (47) ends up being a variable that is bound by every man. Finally, (48) is an example of donkey anaphora, and, beginning with Evans 1980, many have claimed that neither co-reference nor bound variables will work to describe the right interpretation for the anaphoric relationship between a donkey and it. There are a number of different characterizations of the problem(s) donkey sentences present, but one way of describing the problem is as follows. If someone utters (48) and we know a farmer who has multiple donkeys, we interpret this sentence as claiming that the farmer beats all of the donkeys that he owns. So, the desired resulting interpretation is similar to what we see with bound variable anaphora, where for all $x$ such that $x$ is a donkey, $x$ gets beaten. However, the antecedent here is quantificational and the pronoun is not in its scope, which is what makes it a separate category (e-type). There are those who maintain that donkey anaphora can be accounted for via bound variable anaphora, saying that a donkey only introduces a variable $x$ with a restriction donkey'($x$) but not an existential quantifier. Then, at the sentence or discourse-level, a quantifier (in this case, universal) binds all free variables (Kamp 1981, Heim 1982). This kind of approach to the analysis of donkey anaphora is known as Unselective Binding.

As in (48), conditionals are one of the primary environments for donkey anaphora. Beck notes that it is possible in the CC as well, as (49) shows.

(49) The more often a farmer milks a cow, the more he appreciates it.

(B33) Donkey anaphora: The CC and RCC license donkey anaphora between their clauses.
Given the discussion immediately above, Beck's analysis will need to include some way of handling this anaphora, and she uses the unselective binding approach described there.

Other similarities between conditionals and CCs that Beck lists include the following. In both CCs and conditionals, the subordinate clause restricts the domain of cases under consideration and the matrix clause asserts something about those cases. In both, when you quantify over worlds, you ignore some where something miraculous or unexpected happens (so when you say *if he went to the store, he forgot his money*, you're excluding worlds in which he gets run over by a car on his way to the market). She says that counterfactual statements are possible in both, as in (50) and (51), though (51) certainly sounds odd to me.

(50) If he had run faster, he would be more tired now.

(51) The faster he had run, the more tired he would be now.

Last but certainly not least (as this is one of her prime arguments), Beck discusses adverbs of quantification (AQs). There is a long history of work on conditionals (see von Fintel 1994 and Bhatt & Pancheva 2006 for overviews), and one of its most influential observations is the extent to which conditional antecedents serve as domain restrictors of AQs such as *always, sometimes, usually, seldom, normally*, etc. Like conditionals, AQs similarly show a wide range of quantification, leading Lewis 1975 to hypothesize that what is being quantified over with AQs (and conditionals and therefore, potentially, CCs) was a *case*. A case is any admissible value assignment for all the variables that occur free in an open sentence modified by an AQ (according to Heim). This importantly
includes all participant, time, and world variables. So basically, the AQ is a kind of operator in a tripartite structure of the form \([\text{Operator}] [\text{Restrictor}] [\text{Nuclear Scope}]\) (Heim 1982, Kamp 1981). Other operators include modals, quantificational determiners, tense/reference times, and focus-sensitive \textit{only} or negation (Roberts 1995) (among others, such as the comparative). In the case of AQs (and also modals), the restrictor is typically given by an \textit{if}-clause, but they can be implicit, meaning that their value is determined almost entirely by context, as in examples like \textit{Usually, I go to the store} where we're already talking about what I do after school, etc. Similarly, it is possible to have an \textit{if}-clause restrictor without an explicit AQ or modal, as in (52a) and all of the examples we have seen up to this point. Not surprisingly, Beck argues that CCs once again parallel conditionals in showing the same behavior, as in (53). We'll see the extent to which this observation by Beck is crucial to her analysis and others' analyses in subsequent sections.

(52) a. If I buy fruit, I eat it before it goes bad.
   b. Usually, if I buy fruit, I eat it before it goes bad.
   c. If I can buy fruit, I eat it before it goes bad.

(53) a. The more fruit I eat, the less often I get sick.
   b. Usually, the more fruit I eat, the less often I get sick.
   c. The more fruit I can eat, the less chocolate I crave.

(B34) \textit{Overt AQs}: The CC can appear with an overt Adverb of Quantification.

In both cases shown here, Beck argues that the (a) variant without an explicit AQ is understood as having universal (or possibly generic) scope. Beginning with Kratzer (1986), an implicit epistemic modal or universal/generic AQ has been posited as existing
in sentences with a bare conditional antecedent. Similarly, then, it could be extended to
the analysis of the CC, in which case the universal quantification in the CC would not be
a part of the lexical meaning of the. This is the approach Beck takes.

Moving now from conditionals to comparatives, Beck is committed to giving an
analysis of CCs in which the comparative morpheme is a real comparison, as elsewhere
in English. Previous work on the semantics of comparatives is abundant and varied (see
Heim 2006 and Kennedy 1999 for an overview). First, more should have a consistent
lexical entry across sentences such as (54)-(57) if at all possible, and that same lexical
syntax and semantics should not be able to be used to derive sentences such as (58) in
order to account for its ungrammaticality.

(54) Jo owes more than $3. (Simple Comparative)
(55) Jo owes more than Bo owes. (Comparative Subdeletion)
(56) Jo owes more than Bo. (Phrasal Comparative)
(57) Jo's ladder is shorter than her house is tall. (Cross Polar 'Nomaly')
(58) *Jo is shorter than Bo is tall. (Cross Polar Anomaly)

Much of Kennedy's early work centers around these issues, but he also addresses issues
of scope in Kennedy 1995, which is similarly of concern to Schwarzschild and Wilkinson
2002. The interest in scope comes in part from Russell 1905; Russell's yacht example as
in (59) has been argued to show that comparatives are a type of definite description, and
as a consequence, may have wide or narrow scope with respect to an intensional
predicate or operator. On the narrow scope reading, Jones believes something

---

22 This term, coined by Büring (2007) contrasts these sentences with Kennedy (1999)'s 'Cross Polar
Anomalies', as in (19). The Cross Polar Nomalies are a subset of comparative subdeletion sentences.
contradictory, but on the wide scope reading, Jones' impression of the yacht's size is greater than its actual size in the real world.

(59) Jones thinks Smith's yacht is larger than it is.

This is why comparatives are taken to be scopal operators, as alluded to in the discussion of AQs. Beck reviews the main benchmarks that might suggest that the comparative in the CC was particular to that construction: (B17) and (B20) – that than phrases and measure phrases do not appear in CC clauses. To partially combat the idea that these benchmarks are specific to the CC, she gives a range of comparative constructions as in (60). Several of these have comparatives without than phrases.

(60) a. Alex got more and more tired.  
   b. It got hotter every hour by three degrees.  
   c. With every step, Alex grew more tired (than he had been before).  
   d. As Alex advanced up the cliff, he grew more tired.  
   e. Each apple was more succulent (than the last).

She does not give an analysis of any of these kinds of examples, however.

3.2.1 Beck's analysis

In order to maintain that the -er appearing in CCs is nevertheless a regular comparative morpheme, Beck uses a variant of the analysis of comparatives in von Stechow 1984 and Heim 1985, 1990, which is in turn influenced by Bresnan 1973 (which we saw in the last chapter). She gives two entries for more/-er; one for when it occurs without a measure phrase, as in (61), and one for when it occurs with a measure phrase, as in (62), and in both cases, she takes degrees of type $d$ with variable names $d$, $d_1$, $d_2$. 

117
etc. to be part of the natural language ontology.

\[(61) \text{[-er]}(D_1)(D_2) = 1 \text{ iff the max } d_2 \mid D_2(d_2) > \text{ the max } d_1 \mid D_1(d_1)\]

\[(62) \text{[-er]}(D_1)(d)(D_2) = 1 \text{ iff } [\text{the max } d_2 \mid D_2(d_2)] = [d + \text{ the max } d_1 \mid D_1(d_1)]\]

In both (61) and (62), each of the \(D\)'s is of type \(\langle d, t \rangle\) and represents the type of an expression such as \textit{we drive fast} or \textit{you are tall}. In its application to a simple sentence such as \textit{Roy is taller than Ray}, -\textit{er} (from (61)) would take \textit{than Ray (is tall)} as its first argument (where \textit{than} is assumed to be vacuous and \textit{is tall} is assumed to be present but silent), and then it would take a second degree predicate \textit{Roy is tall}. It would return true of the maximum degree of Roy's height is greater than the maximum degree of Ray's height. If we were instead looking at the sentence \textit{Roy is 2 inches taller than Ray}, we would use (62), which would return true if Roy's height was equal to two inches plus Ray's height.

As Beck notes, she uses a definite in the comparative ('the max') where others (von Stechow 1984, Heim 1990) use an existential quantifier over degrees. She recognizes that this could lead to mispredictions about negation, negative polarity, etc. in terms of their interaction with the maximality operator of the definite, but she uses this analysis of comparison anyway because it is easier for analyzing the CC. She gives the following LF for a comparative sentence with a specified differential measure phrase.
Here we see a discontinuous analysis of the comparative in which it never forms a meaningful expression with an adjective or adverb apart from the rest of the clausal material. Aside from the analysis of the CC, the main contribution of Beck's paper is a strong argument that this is necessary, so we will return to this again later.

I now turn to discussing aspects of the sentence-level meaning that Beck gives CCs. Then, using the meaning of the comparative and the meaning of the CC as a whole, we will see how Beck chooses the meaning of *the* in order to give a compositional analysis. (65) shows the meaning that she gives for the sentence in (64).

(64) The better prepared Otto is, the better his talk will be.

(65) \[
\forall w_1 w_2 \quad [\exists d [d > 0 \& \text{the max } d_1[[\text{well}(d_1, \\
\lambda x[\text{prepared}_w(x))(\text{Otto}) = d + \text{the} \\text{max } d_2[\text{well}(d_2, \lambda x[\text{prepared}_w(x))(\text{Otto})]]] \\
\Rightarrow [\exists d [d > 0 \& \text{the max } d_1[\text{good}_w(d_1, \text{Otto's talk})] = d + \text{the max } d_2[\text{good}_w(d_2, \text{Otto's talk})]]]
\]
Basically, (65) says that for all pairs of situations in which there exists a positive non-zero degree difference between Otto's degree of preparedness at one versus the other, then there also exists a positive non-zero degree difference between the degree to which his talk is good in the first situation as compared to the second. As Beck points out, this meaning says nothing about the respective sizes of the differences in preparedness and talk-goodnessss, nor does it say anything about the general relationship between preparedness and goodness of talks. Thus, this meaning is weaker than what Thiersch, Fillmore, etc. would have it be. In particular, the idea of a dependent and independent variable is missing. Beck rightly shows that a stronger meaning would make incorrect predictions, however, which I discuss next.

One possibility that she does not mention is one that is even stricter than the previous literature proposes. The degrees or degree differences in one clause could need to be identical, in which case there would be co-reference between the clauses. This is clearly not required given that in most of the examples we have seen so far, the degrees in each clause are not even of the same kind (height in inches as compared to age in months or years, for example), let alone the same degree and kind/scale.

The next most restrictive possibility after co-reference is linear dependence or what Beck calls 'proportionality'. This would require one kind of degree expressed in a CC (in one of its clauses) to be a function of the other, and a linear function at that. This, too, is too restrictive given the existence of felicitous CCs that express an exponential or other non-linear function, such as *the larger a number, the larger its square*. If we relaxed this condition to include any functional dependence in which one value could be determined by rule or algorithm from the other, this would include exponential,
logarithmic, etc. cases. It would also be consistent with the intuition of past authors that one clause gives an independent variable and the other, a dependent variable. This is also too strong, however, because there does not have to be an algorithm or rule relating the values denoted by each clause. For example, evaluate the truth of (66) within the model given below it. Because the number of goals scored increases as the temperature increases, this sentence is judged to be true. But note that while an increase of two degrees (from 65 to 67) led to a one-goal increase, a one degree increase (from 67 to 68) yielded a two-goal increase, showing that only the fact of increasing and not the size of the increase is correlated between the two things being measured.\footnote{Despite an assumption in the literature that CCs are cross-linguistically identical in both their syntax and semantics, it seems that some languages, including Croatian and Hindi, do require some kind of semantic proportionality in CCs, making examples such as (66) infelicitous in the given model. We will see this in Brasoveanu's 2009 discussion of Romanian.}

\begin{equation}
\text{(66) The warmer it was, the more goals they scored.}
\end{equation}

\begin{align*}
\text{# of goals scored:} & \quad 1 \quad 3 \quad 4 \quad 6 \quad 7 \quad 10 \\
\text{temperature in deg:} & \quad 60 \quad 65 \quad 67 \quad 68 \quad 72 \quad 76
\end{align*}

Because this kind of 'proportional' analysis does not adequately reflect the truth conditions of the CC, contemporary theories in which this sentence-level meaning is ascribed to CCs will not be considered here. This includes Hsiao 2003.

Even if the relationship between the degrees in the two clauses is neither linear nor algorithmic, it may still be a monotonically-increasing function. However, even this is too strong because the relationship is not functional. In order to be a function, each element of the domain could only have a single associated value in the codomain. However, as Beck shows, (67) is judged to be true in a model like the one below it in
which there are multiple codomain values \((3, 4)\) associated with a single domain value \((25)\).

(67) The warmer it was, the more goals I scored

<table>
<thead>
<tr>
<th>Game</th>
<th>Temperature</th>
<th>Goals</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>15</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>20</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>25</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>25</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>25</td>
<td>4</td>
</tr>
<tr>
<td>6</td>
<td>30</td>
<td>5</td>
</tr>
<tr>
<td>7</td>
<td>35</td>
<td>7</td>
</tr>
</tbody>
</table>

Despite this, Beck sometimes says that the CC asserts that a monotonic function exists between the degrees in the two clauses. I assume that what she means is that there is a relation (free from the restriction on functions) between two functions that are monotonically correlated as per the definition in (68).

(68) Monotonic Correlation: Two functions \(f\) and \(g\) with the same domain and pre-ordered codomains are monotonically correlated if and only if
\[
\forall x \forall y \ [g(x) < g(y) \rightarrow f(x) < f(y)]
\]

Here, we see that the idea of monotonicity \((x>y \rightarrow f(x)>f(y))\) is preserved but the relationship between \(f\) and \(g\) does not have to be a function (I leave this proof to the reader). In consulting the sentence-level meanings that Beck gives for a variety of sentences (some of which are given above), it is clear that she has concluded that the CC denotes the existence of a monotonic correlation between degrees with respect to two different properties (one per clause) for pairs of individuals, times or worlds. So, for the example the faster we drive, the sooner we'll get there, \(x\) and \(y\) are variables over
different possible worlds in which we drive a certain speed and get somewhere at a particular time, \( g \) is the function that gives you the driving speed for a particular world, and \( f \) is the function that gives you the arrival time for a particular world.

There is one further dimension of the CC's meaning that is sometimes confused with whether the CC denotes a proportional dependence. It has to do with whether CCs necessarily express causation or just correlation. CCs are often used to express natural relationships, such as the one between speed and time in (32), or other cause-effect relationships. As Beck points out, however, it is possible to use a CC to merely note a correlation rather than a causation, as in example (69).

\[(69)\] The more prolific a semanticist, the taller her husband.

Though I doubt that anyone would argue that a generalization like (69) could have a causal root, I believe that clearer evidence comes from the fact that it is possible to explicitly deny causation without leading to unacceptability, as in (70).

\[(70)\] The more ice cream people buy, the more murders occur, but ice cream doesn't cause murder (it's just that both rise in warm weather).

Despite this, the interpretation of CCs as expressing a causal relationship is tempting, I think, for the following reason. Cause-effect relationships are asymmetric. If \( x \) causes \( y \), it is not the case that \( y \) causes \( x \). Correlational relationships, on the other hand, are symmetric. If \( x \) is correlated with \( y \), then \( y \) is necessarily correlated with \( x \). This relates to some of the ideas in the previous chapter with respect to symmetry or asymmetry in the syntax and semantics. I propose that the reason the CC seems to express cause-effect
relationships is because they seem asymmetric. In particular, when a CC is used to express a known cause-effect relationship, the causal factor must be denoted by the subordinate clause and the effect factor, the matrix clause. Consider (71) in which the clauses cannot be switched with a preservation of meaning.

(71) a. The taller the child, the older the parent.
    b. The older the parent, the taller the child.

<table>
<thead>
<tr>
<th>Age of parent:</th>
<th>40</th>
<th>42</th>
<th>45</th>
<th>47</th>
</tr>
</thead>
<tbody>
<tr>
<td>Height of child:</td>
<td>4'7''</td>
<td>4'8''</td>
<td>4'8''</td>
<td>5'</td>
</tr>
</tbody>
</table>

Normally, any change that results in a truth-conditional difference is thought to be encoded in the semantic form. However, there are actually two distinct differences between (71a) and (71b). The first, which is responsible for the difference in truth conditions, is that the first clause's meaning restricts the domain, so different subordinate clauses lead to different domain restrictions. In particular, (71a) doesn't consider any pairs of children that are the same height, while (71b) doesn't consider any pairs of parents who are the same age. Thus, consider the model under (71). It verifies (71a) but falsifies (71b). All differences in children's heights lead to positive non-zero age differences, but not all age differences lead to non-zero height differences, given that both the 42 and 45-year-old parents have children of the same height. This is one asymmetry, and the second is related to causation. Compare (71), which could be a correlation or cause-effect relationship to (72), which we know to be causal.

(72) a. The faster we drive, the sooner we get there.
    b. #The sooner we get there, the faster we drive.
    c. The sooner we get there, the faster we will have driven.
(73) a. If we drive faster, we will get there sooner.
b. #If we get there sooner, we will drive faster.
c. If we get there sooner, we will have driven faster.

Because arrival time does not affect speed but rather is affected by it, (72b) and (73b) cannot mean what (72a) and (73a) do, so even if these just mean something very odd rather than being infelicitous, there is a clear difference between them and their reversed counterparts. They can only mean that as result of an earlier arrival time, they will drive faster on another occasion in the future. I included (73) to show that conditionals are also asymmetric structures in which cause-effect relationships must be encoded with the cause in the subordinate clause and the effect in the matrix. Given this parallel, then, there is a potential argument for causality in the CC. However, it is not always the case that switching clauses leads to infelicity. Taking the earlier example that seems to have no possible causal link and is just an accidental generalization, they are felicitous in either order.

(74) a. The more prolific a semanticist, the taller her husband.
b. The taller a semanticist's husband, the more prolific she is.

In (74), the clauses can be switched without affecting the meaning of the CC, which is facilitated by the one-to-one relationship between husbands and wives but is also due to the lack of causality. The combination of the fact that some CCs can have their clauses switched and the fact that causality can be denied without leading to a contradiction proves that causality is not part of the truth-conditional meaning of the CC. In order to account for the fact that, when causal relationships are expressed, the cause is denoted by the subordinate clause, I propose that there is an asymmetry inherent to any conservative
operator due to its conservativity that is responsible for this effect (this will be defined in 3.2.2.1). In other words, the correlation itself is symmetric, but because the correlational operator is conservative, there is an asymmetry such that the thing likely to be the cause should be the first argument of the operator. This would explain why it is explicitly defeasible, and, in cases where real world knowledge prevents a causal interpretation, limited by context.

Now that we have defended a semantic form for the CC as a whole and seen what Beck takes to be the semantics of the comparative, here is her analysis of each clause in a CC. The Logical Form for the better prepared Otto is appears in (75) (where German je is taken to be identical to English the), and her lexical entry for the, in (76). I've also restated the lexical entry for -er in (77).

(75) \[ \text{je}'(w_1, w_2)(-\text{-er}')((\lambda w \lambda d[\text{well}(d, \lambda x[\text{prepared}_w(x)])](\text{Otto}))) \]

(76) \[[\text{the}]](w_1, w_2) ([[\text{-er}]])(D_{<d,\text{t}>>}) \text{ iff } [[\text{-er}]](D(w_1))(D(w_2))

(77) \[[\text{-er}]](D_1)(d)(D_2) = 1 \text{ iff } \left[ \text{the max } d_2 \mid D_2(d_2) = [d + \text{the max } d_1 \mid D_1(d_1)] \right]

Beginning with (75), notice that Otto is x well-prepared is the meaning of the C’. It would normally be something of type <\text{d}, \text{t}> that the comparative could take as argument. Because it will be an argument of the in this case, and the requires something of type
<s,<d,t>>, Beck assumes that type-driven $\lambda$ abstraction takes place that makes the C' the $\lambda w \lambda d...$ term in (75) rather than a $\lambda d...$ term. In the DegP, the takes two implicit world variables as its first argument. Its second argument is the comparative morpheme, and then the C'. Once the takes these arguments, it is a proposition. It is true if there is a degree difference between the value of the degree predicate applied to the first world and the value of the same degree predicate applied to the second world. This is related to the discussion of Beck in chapter 2 in that she is relying on something syntactic (but it is unclear what that would be) to disallow the possibility of a the clause standing alone as a sentence (prohibited by (B3)), since each clause, as a proposition, might otherwise stand alone.

Looking closely at (76) and (77), the comparative in (77) requires three arguments: two degree predicates and a degree. However, in (76), the degree argument is missing. Beck assumes that when no overt degree is present, the comparative existentially binds its degree argument, which is where the $\exists d$ binding comes from in (78), below, which is Beck's derivation of the meaning from (75).

(78)

\[
\begin{align*}
[[j e' (w_1, w_2) (-er')] \ (\lambda w \lambda d[\text{well}(d, \\
\lambda x[\text{prepared}_{w_1}(x)])(\text{Otto})]) \\
\text{iff } \exists d[d > 0 \& \ [[-er']][[[\lambda d[\text{well}(d, \lambda x[\text{prepared}_{w_1}(x)])(\text{Otto})]]]]] \\
\text{iff } \exists d[d > 0 \& \text{ the max } d]\text{well}(d, \lambda x[\text{prepared}_{w_2}(x)])(\text{Otto})] = \\
\quad d + \text{ the max } d[\text{well}(d, \lambda x[\text{prepared}_{w_1}(x)])(\text{Otto})]
\end{align*}
\]

The way that the and -er work together to create the meaning in (78) is as follows: part of the's meaning is to take one function from worlds to degree predicates and turn it into two degree predicates, each denoting the degree of something in a different situation.
Another part of the's meaning is to feed both of these degree predicates to the comparative morpheme, which would otherwise be looking for one in the form of a than phrase. As a result of this meaning for the, Beck's analysis will satisfy benchmarks (B11), (B17) and (B20): requiring a comparative morpheme and prohibiting than phrases and measure phrases. The comparative is required because it is an argument of the, and the than phrase is prohibited because the makes sure that both degree predicate arguments of -er are already filled, leaving no room for a than phrase. Similarly, because the specifies what arguments -er takes and they do not include a value for d, the guarantees the existential binding of a degree variable rather than the possibility of a measure phrase.

Looking now at the combination of the clauses, Beck takes the quantification in the CC to be implicit unless it is determined by an overt AQ, per Kratzer's account of conditionals. Beck's sentence-level LF for the CC is thus as in (79) for the sentence the better prepared Otto is, the better his talk will be, where ∀ on the left branch is the implicit AQ that takes each clause's meaning as argument and binds each's world variables. At the top CP node, we see that ∀ takes two sets of pairs of worlds, but we saw above that each clause returned a truth value (modulo contextual retrieval of values for its free variables). Beck reconciles this type mismatch by once again assuming type-driven λ binding for the free variables in each of the CC's clauses. The same is true for the RCC in (80) (Otto's talk will be better, the better prepared he is), where the subordinate clause is on the right, and the implicit universal quantification with it.
In (80), note that the DegP in the left (matrix) clause contains a \textit{the/je/umso}. This means that for the analysis of English, Beck assumes a null \textit{the} at work in the RCC. Based on
the difference in the structures in (79) and (80), Beck says that one would expect scope differences in the CC v. RCC. Then she shows that negation in the first clause of the RCC can have wide scope, but not negation in a CC clause (as we've seen before). She also ties the lack of negation in a CC to the inability to have negative quantifiers in comparatives (like *John is taller than no child*). The same is true for the CC (*the taller someone is, the smaller his feet aren't*). The point is that the narrow scope reading doesn't make sense in either of these cases, whereas it does for some other cases, which are the cases where negation is licit but still narrow scope. For the CC, this includes (81). Based on the benchmarks in the last chapter, we know that it is not possible for this negation to take wide scope, but Beck's point is that that is true of comparative sentences as well, which is one more indication that her analysis of CCs (in which the same meaning for the comparative as in these other comparative sentences is used) is on the right track.

(81) a. The more you don't brush your teeth, the more cavities you'll have.
    b. The more you brush your teeth, the more likely you won't have to get dentures.

(B35) Narrow scope negation in CCs: CC clauses can contain narrow scope negation where it is felicitous with respect to the comparative.

As Beck points out, this is a quirk of negation. Other quantifiers such as *every* can receive wide or narrow scope in both the comparative and the CC.

So far we have seen Beck's analysis for CCs and RCCs with quantification over worlds. For quantification over times and over individuals, she posits new lexical entries for *the* that are only different from the above in the type of the covert variables they take as their first argument. The implicit AQ that Beck assumes presumably binds any
number, kind, etc. of free variables indiscriminately, so it will not need separate lexical entries for various types of variables. And since the degree predicate at C' gets coerced to take another argument anyway, the kind of argument created via \( \lambda \)-abstraction will always be in accord with the variable that the requires, so there is no lexical ambiguity within the C' either.

For quantification over individuals, the individual variables are introduced via definites or indefinites that contain not just an individual-type variable, but also a restriction on that variable. Beck assumes that these denote 'open sentences' – a property based on the denotation of the noun with a free variable as its argument. She also assigns the content of these noun phrases to be inside the degree description in C', as we see in (82), which is the sentence-level form for the sentence the slimier an attorney, the more successful he is. She compares this to what we might expect as a final form as in (83) but points out that her system is unable to derive (83).

\begin{align*}
\text{(82)}
[[(\forall (\lambda xy [ie'(x, y) (-er') (\lambda x \lambda d [\text{attorney}(x) \& d\text{-slimy}(x)])])
(\lambda xy [ie'(x, y) (-er') (\lambda x \lambda d [d\text{-successful}(x)])])]]
\iff \forall xy [\exists d [d > 0 \& \text{the max } d[\text{attorney}(y) \& d\text{-slimy}(y)] =
 d + \text{the max } d[\text{attorney}(x) \& d\text{-slimy}(x)])]
\Rightarrow \exists d' [d' > 0 \& \text{the max } d[d\text{-successful}(y)] = d' +
 \text{the max } d[d\text{-successful}(x)])]
\end{align*}

\begin{align*}
\text{(83)}
\forall xy [\text{attorney}(x) \& \text{attorney}(y) \& \exists d [d > 0 \& \text{the max } d[d\text{-slimy}(y)] =
 d + \text{the max } d[d\text{-slimy}(x)])]
\Rightarrow \exists d' [d' > 0 \& \text{the max } d[d\text{-successful}(y)] = d' + \text{the max } d[d\text{-successful}(x)])]
\end{align*}

Beck then suggests that this is a potential advantage of her system. She says that (82) and (83) only differ truth-conditionally in the case where there are no attorneys. The
latter would be true if there aren't any, whereas hers would return no truth value since its presupposition wouldn't be met. She says that her intuition is that the CC in question is not felicitous if there are no attorneys in the worlds in question, favoring her (82).

Nevertheless, she does point out a few reservations about her analysis of quantification over individuals (Beck 1996). The first is that any kind of requirement as in Heim 1982 that the first clause have an indefinite or the second clause have a pronoun has to get stated at the level of combination with *the*. This is different from the conditional, where it is stipulated at the sentence level. The second concern is that any time there is more than one indefinite in a CC clause, there would have to be a new form of *the* that takes an extra individual argument. Also, in the case of multiple individual variables in the degree description (C'), the order in which the variables are bound matters. So, which pronouns in the second clause refer back to which indefinites in the main clause just depends on which you \(\lambda\)-bind first. She points out that she has not way of accounting for cases where these relationships are judged by speakers to be fixed (i.e. unambiguous examples), since you could always bind them in either direction from a technical perspective.

In summary, we have seen in this section Beck's 1997 analysis of CCs and RCCs and how it accounts for a number of benchmarks. She is able to maintain a meaning for the comparative morpheme in the CC in which it plays the same role that it usually does, with *the* and some covert elements (AQs, world/time/individual variables) doing all the heavy lifting. The *the* in each clause has the same meaning, and there is a null *the* in the first clause of the RCC. She manages to technically observe resource sensitivity (using each argument only once), but then the meaning of *the* makes it such that the degree
description with its adjective, restrictions on individuals, etc. gets used twice. This is what is crucial to build a CC meaning that represents a monotonic correlation. The monotonic correlation itself is created from the meanings of various CC elements rather than being stipulated in its entirety by a single lexical item or constructional meaning. Her analysis shows that a compositional treatment of the CC is possible.

Another implication of her theory relates to the interpretation of comparatives. Beck specifically sates that it is not possible to have her analysis of CCs and have a continuous comparative (one in which an adjective and -er combine together as a constituent independently of other words or morphemes). It's easy to see why this has to be so, because for her, the comparative morpheme is used once in each clause, but the adjective's meaning is used twice (and there are other aspects of the meaning that need to get used twice as well). If they were an indivisible unit, there would be no way to duplicate one along with other parts of the sentence without duplicating the other. Beck therefore argues that \([Adjective + -er]\) cannot be an atom of interpretation for CCs and that morphological boundaries are visible to LF.

3.2.2 Discussion

Overall, Beck has the best working hypothesis on the meaning of the CC and RCC. Nevertheless, the analysis has several kinds of problems. Some are those that Beck mentions for her own analysis, especially as related to the treatment of examples in which quantification takes place over individuals. Some are issues that I brought up in the process of describing the analysis in order to illustrate what predictions it makes, as in the problem of not accounting for (B3: CC clauses don't stand alone). Some are issues
related to analyses she sketches but does not detail. These include questions about her analysis of examples in which there is a non-specific indefinite with quantification over times.

For Beck, the preferred reading for a sentence like *the slimier an attorney is, the more successful he is* is the one quantifying over individuals, not times, which is the opposite of what Lin 2007 predicts for Mandarin, as we'll see in the next section.

Because he believes that his analysis extends to English, he questions her judgment. He also asks, if her assumed interpretation in (84) is right, how do you get both a generic operator and the universal operator in a single conditional?

(84) \[ \text{Gen } x \left[ \text{attorney}(x) \right] \left[ \forall t_1, t_2 \right] \left[ x \text{ is slimier at } t_1 \text{ than at } t_2 \right] \Rightarrow \left[ x \text{ is more successful at } t_1 \text{ than at } t_2 \right] \]

I don’t see Lin's concern as a problem for Beck since any unbound variable needs a binder and we often have sentences in which multiple operators are at work, as in all cases of multiple quantifier scopes (including those in conditionals). The question is then what operators are at play, and Lin is right that, on the face of it, the conditional (or its implicit adverb of quantification) is the obvious candidate but wouldn't be contributing two different quantificational forces, scopes, etc. I think Beck must be assuming, however, that generic indefinites, quite apart from conditionals, have a default generic operator of their own. But Beck would need to explain why the indefinites in the CC are not always analyzed as generic indefinites, given that the CC always expresses a generalization since currently they are not generic if they are the basis for quantification over individuals, only when quantification occurs on the basis of another type of
variable. Lin also points out that it's not clear why the generic indefinite should necessarily take wide scope.

The remaining set of problems for Beck's analysis are the most damning. As I will argue in the next several subsections, CCs are not universal, they cannot appear with the same range of AQs as the conditional, and they sometimes appear with than phrases. Beck's analysis predicts default universal quantification, the same range of AQs as with the conditional (since for her, the CC is a conditional), and the impossibility of a than phrase in any CC or RCC.

3.2.2.1 Quantificational Force

As we will see in later sections, every analysis of the semantics of CCs to date has assumed universal quantification in the CC and RCC. Given the universality in the definition of monotonic correlation and the assumption that CCs express monotonic correlations, it is easy to understand why that is tempting. The first piece of evidence against having universal quantification (even by default, as with Beck's analysis) is both the most obvious and the easiest to dismiss: the fact that native speakers accept CCs as true in models where there are a few exceptions to the monotonic correlation. This is true of most sentences that are thought to be universal for a couple of reasons.

Quantifiers in natural language have restricted domains (except in special cases where the domain is really the entire set of worlds, individuals, etc.). When we say that everyone was at the party, we probably mean that everyone you might expect to be there was at the party, such as our mutual group of friends, not every member of the group of Homo sapiens. Sentences with quantificational determiners (generalized quantifiers) are
taken to have tripartite structures of the form [Operator] [Restrictor] [Nuclear Scope] 
(Heim 1982, Kamp 1981). The quantifier is the operator, its common noun argument is 
the restrictor, and the verb phrase is the nuclear scope. But in the example of everyone 
being at the party, the restrictor *one in everyone* is an explicit restriction to humans but it 
does not fully restrict the domain of quantification. There is an implicit restriction that is 
pragmatic that tells us who we are actually considering. Given the possibility of implicit 
restriction, then, it is possible that any exceptions to the monotonic correlation in the CC 
are just part of a set that is not being considered for pragmatic reasons.

There is also a vagueness inherent in a number of places in natural language. The 
fact that the sentence *France is hexagonal* is judged to be true despite the fact that the 
country's shore lines and borders couldn't possibly meet the definition of a hexagon is 
one example. Given this, one could say that CCs allow some exceptions on the basis of 
vagueness, even among those cases in the pragmatically-restricted domain. Generic NPs 
are another place where there is some vagueness in that we may say that dogs 
(generically) have four legs, but just because a dog has lost a leg, that doesn't falsify the 
rule (in fact, we then consider the anomaly an 'exception to the rule' rather than revising 
the rule). Conditionals can also be generic statements (*if it is a dog, it has four legs*), and 
similarly, they permit some violations while remaining true. So far, then, Beck could 
maintain her parallel between the conditional and the CC and say that both have generic 
rather than universal force. And in fact, most authors do not even distinguish between 
universal and generic force as a default for conditionals and wouldn't think twice about 
extending the “universal/generic” moniker for implicit quantificational force to CCs (cf. 
von Fintel 1997). 

136
Because generic force does not translate into a more concrete formal representation, it is difficult to distinguish from other quantifiers as well. In particular, a quantifier like *most* also requires that the majority of restricted cases meet the requirements of the matrix predicate or clause, so even though there may be an intuitive difference between the generic operator and *most*, it remains just that: intuitive. In what remains of this section, however, I propose a way of distinguishing *most* and other similar operators (majority, etc.) from the generic operator. We will then see, just as we saw above with the case of NPIs and entailment patterns, that the CC patterns like *most* rather than like generics or conditionals. The similarities between antecedents of conditionals and CCs will then be explained in terms of the fact that both happen to be Generalized Quantifiers. A bit of background is in order next.

Generalized Quantifiers (GQs) are operators of type \(<xt, <xt,t>>\) where \(x\) is a variable over individuals, worlds, etc. Barwise and Cooper 1981 propose that all natural language GQs are *conservative*, defined as in (85).

\[(85) \text{Conservativity} \text{ is the property of being a predicate (OP) on two properties such that } \text{OP}(A,B) \text{ is equivalent to } \text{OP}(A, A&B).\]

In prose, (85) means that the things you're talking about in the restrictor are the same things you're talking about in the nuclear scope. Chierchia (1995) argues that conservativity is at the heart of donkey anaphora; the fact that the farmers and donkeys discussed in the first clause are interpreted as being coreferential with those in the second clause has to do with the fact that the same situations are under consideration in both clauses. Typical GQs include *every, most, none, some, few, half, at least 2*, etc. De Swart
1991 also argues that Adverbs of Quantification (AQs) are GQs, including *usually*, *occasionally*, *always*, *never*, etc.

Kratzer 1986, 1991 and von Fintel 1994, inter alia, apply the GQ theory to conditionals. I have taken their analysis as the default throughout this chapter since Beck does and it was important to understand how her theory works. They argue that the *if* clause restricts an operator, and that the matrix clause is the nuclear scope argument. What operator is this? As we have seen, it could be an overt AQ, as in (86).

(86) a. If I buy coffee, I always buy a donut.
    b. Sometimes, if it rains, there is a rainbow later.
    c. If I have a dream, it's never about zombies.
    d. Occasionally, if it stops snowing, it starts raining.

Because of the possibility of an overt AQ determining the quantificational force, Kratzer (and others, following her) doesn't want to assign a particular quantificational force to any of the lexical items in the conditional (i.e. it's not always existential, but it's also not always universal, etc.). So she proposes that in the absence of an overt AQ, there is an implicit AQ operator that has 'universal/generic' force (so that the default interpretation of donkey sentences is a generic one).

In Beck's analysis, the implicit operator (and the other overt AQs) suffers from some of the same problems as unselective binding accounts of conditionals (like Kamp 1981 and Heim 1982). One of these is the proportionality problem. Proportional quantifiers are ones that require reference to the number of elements in a set, like *most*, *half*, *usually*, *30%*, etc. When one of these is added to a conditional, as in (87), it makes the wrong prediction in a certain set of models, such as the one in (88). As Heim 1990
points out, native speakers judge the sentence to be false in this model, but the theory predicts the sentence to be true in the same model.

(87) If a farmer owns a donkey, he is usually rich.
(88) Model: There are 100 farmers with donkeys, 99 have one and are all poor and the last one has 200 donkeys and is rich.

The trouble is that with proportional quantifiers, one is interested in the number/ratio of cases that verify the matrix predicate, so what you are counting matters. Thus, in the case of (87), it matters whether you are counting farmers or farmer-donkey pairs. If you instead look at a sentence like (87) but with usually removed, there is a more stringent burden of proof and counting by farmers and farmer-donkey pairs will both yield a false result in keeping with native speaker intuitions. This will become relevant presently.

Note that there are many analyses of conditionals including Kratzer 1986 and Heim 1990 that do not suffer from this particular problem. But for our purposes, it doesn't matter what the actual analysis of conditionals should be because it turns out that CCs aren't like conditionals. However, it turns out that when the 'proportionality problem' arises is key to understanding what the analysis of CCs should be. For Beck, on the other hand, given her unselective binding account of the CC, the following examples will already be problematic for the same reason that (87)/(88) is problematic for unselective binding theories of conditionals.

The proportionality problem that we saw for conditional sentences with an overt proportional AQ arises in CCs without any kind of additional proportional operator. For example, consider sentences (89), (90) and (91) as well as their respective models.
(89) The more a man loves a woman, the more flowers he brings her.

Model for (89): we know 5 men. One has nine girlfriends, and his love for each depends upon her behavior. The nicer she is, the more he loves her, and the more flowers he brings her. The other four men each have one girlfriend, and even though their love for their girlfriends grows over time, they only bought them flowers during the first couple months of their relationship and now they don't buy flowers no matter what.

(90) The more closely related a man is to a woman, the less likely he is to marry her.

Model for (90): we're in Austen Pride & Prejudice land. One guy (Darcy) has, say, five possible women to marry including Ms. Bennet (no relation at all), his best friend's sister (a distant relative) and his rich first cousin. He's most likely to marry Ms. Bennet, somewhat likely to marry his friend's sister, and there's a next-to-null chance he's going to marry his cousin. The other two women in the running fit the same trend. Then there are three other guys who all went to primary school with Darcy and are absolutely going to marry their first cousins because they follow the rules and that's what was arranged by their parents and they're not considering any other women as possibilities.

(91) The better a dog knows a command, the more treats he will receive.

Model for (91): There are 10 dogs that know any commands. Two of them are star pupils: Simon and Chloe know 9 commands to varying degrees. The other 8 dogs each know only one command. Here is a list of those 8 dogs, how often the respond correctly to the command, and how many treats they receive per day:

<table>
<thead>
<tr>
<th>Name</th>
<th>%correct response</th>
<th>#of treats</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mia</td>
<td>10</td>
<td>6</td>
</tr>
<tr>
<td>Mitka</td>
<td>14</td>
<td>7</td>
</tr>
<tr>
<td>Sammy</td>
<td>15</td>
<td>8</td>
</tr>
<tr>
<td>Taffy</td>
<td>20</td>
<td>10</td>
</tr>
<tr>
<td>Murphy</td>
<td>22</td>
<td>12</td>
</tr>
<tr>
<td>Alf</td>
<td>25</td>
<td>13</td>
</tr>
<tr>
<td>Basil</td>
<td>30</td>
<td>14</td>
</tr>
<tr>
<td>Mr. Muffles</td>
<td>33</td>
<td>17</td>
</tr>
</tbody>
</table>

Chloe and Simon's owners reward every command equally regardless of how often their dog gets it right. They give one treat per command per day, so Simon and Chloe both get 9 treats a day, even though they respond to the 9 commands with accuracy that varies between 35 and 95% correct.

Note that none of these sentences has an explicit proportionality operator (that we recognize, such as most or usually). If the model of CCs as conditionals was correct and
the default operator at work in these sentences had a universal interpretation, all three of 
these examples should be judged false by native speakers because each has exceptions to 
the correlation expressed by the CC. If, instead, there were a proportional quantifier that 
was a part of the meaning of the CC, we would predict the following pattern, which 
matches native speaker intuitions.

(92) a. (89) False because 4/5 guys don't fit the pattern. 
   b. (90) False because 3/4 guys don't fit the statement. 
   c. (91) True because 8/10 dogs fit the pattern.

In order to verify that my sample population of native speakers was not anomalous, I also 
asked them about conditional sentences without a proportional quantifier, such as (93)-
(95). If conditionals and CCs truly had the same default quantification, the pattern with 
the conditionals should match the pattern with CCs. That would be surprising from the 
perspective of the conditional literature in which the 'proportion problem' is reserved for 
example with explicit proportional operators, but it was worth testing. I asked about 
conditionals in several different kinds of models. Obviously, this included the kinds of 
models needed to test for the proportionality effect, but it also included models that 
would clearly verify or falsify the sentence in question, just so that I could see whether 
people's judgements were stable across those examples (I thought that this was important 
given how nuanced these kinds of judgements are). I reserve for future work any attempt 
to test these judgements experimentally. After each conditional in question, I include its 
model and the predicted responses based on different analyses of the conditional. As 
with the CCs, these are just a sample of the full range of examples I asked about, and I 
varied the models so that there were some where the proportion was just barely a
majority and others where it was a significant difference as well as some where there was a binary difference (rich v. not rich) and others where it was multi-valued, but the results I report below were true regardless of these differences.

(93) If a person earns a degree, s/he is rich.

Model for (93): There are 10 people with degrees. 9 have one degree and are poor. 1 has 10 degrees and is rich.

This sentence will be judged to be:
T if pairs (of people and degrees) are counted and the S has a proportional interpretation
F if pairs are counted and the sentence has a universal interpretation
F if individuals are counted and the sentence has a proportional interpretation
F if individuals are counted and the sentence has a universal interpretation

(94) If a grocer sells an apple, she is relieved.

Model for (94): There are 32 grocers who sell apples. They are all relieved when they sell apples.

This sentence will be judged to be:
T if pairs (of grocers and apples) are counted and the S has a proportional interpretation
T if pairs are counted and the sentence has a universal interpretation
T if individuals are counted and the sentence has a proportional interpretation
T if individuals are counted and the sentence has a universal interpretation

(95) If a dog knows a command, it is intelligent.

Model for (95): There are 200 dogs that know commands. 167 know one command and are intelligent. 33 know two commands and are dumb.

This sentence will be judged to be:
T if pairs (of dogs and commands) are counted and the S has a proportional interpretation
F if pairs are counted and the sentence has a universal interpretation
T if individuals are counted and the sentence has a proportional interpretation
F if individuals are counted and the sentence has a universal interpretation

The cases of most interest were the cases like (95) in which a T or F definitively differentiated proportional and universal interpretations regardless of whether one was
counting pairs or individuals. The results revealed that people found (93) false, (94) true and (95) either false or odd. Though some people found (95) 'weird' or 'not strictly true but I'd let it slide', this was still markedly different from the response with the CCs where, for parallel contexts, people said that the sentences were true without any comments about oddity, etc. It is also not surprising that people would be unsure about how to judge sentences like (95) given that, as discussed above, even universal/generic statements do permit some exceptions, so it may have been a tricky case.

To summarize, as you can see in terms of the overlap and the overall general patterns of response, there is a difference between the conditional and the CC with respect to examples without an overt proportional item judged in a model like those used to demonstrate the proportionality problem for unselective binding theories. What was a solid true (false) judgment in the CC becomes 'not sure' or false (true) with the conditional. What we're really seeing here is the difference between the generic and the 'most' (noticeable majority) reading. Based on this evidence, I conclude that the quantificational force at play in the CC is proportional, not universal, and I establish the following benchmark:

(B36) Quantificational Force in the CC: The (unmodified) CC's quantificational force is neither universal nor generic but rather patterns with most in terms of native speaker judgements of truth in proportional contexts.

3.2.2.2 AQs in CCs

If CCs are like conditionals in the Kratzerian treatment, which is the one used in Beck 1997 and Lin 2007, then the semantic structure of a CC like the faster we drive, the sooner we get there is something similar to (96) where the the's are not responsible for
the quantificational force in the CC, and instead a covert/implicit operator is.

(96) (implicit 'for all') (the faster we drive – Restrictor) (the sooner we'll get there – Nuclear Scope)

This predicts that, as with the conditional, any AQ should be possible modifying the CC and replacing the implicit operator. Beck illustrates the similarity by using the example of mostly/frequently, which does appear with both (as we saw above), but occasionally and other AQs seem to be able to restrict CCs less readily than conditionals, as we see in (97) and (98).

(97) a. Always, the faster we drive, the sooner we'll get there.24
b. Sometimes, the more educated John becomes, the better his career prospects.
c. Occasionally, the more the dogs pull, the better.
d. Seldom the more we give, the more we get in return.
e. Rarely the worse the fight, the less violence occurs as a result.
f. Never, the faster we drive, the later we get there.

(98) a. Always, if you need anything, call me, no matter what time day or night.
b. Sometimes if John becomes more educated, his career prospects are better; those are the times when he gets government certifications. All other degrees and continuing education don't help.
c. Occasionally, if the dogs pull hard, it helps me up a hill.
d. Seldom will we arrive late if we leave thirty minutes early.
e. Rarely if the fibroid is unusually large, surgery may be recommended.25
f. Never if people have similar salaries will the court award alimony.

We will return to further discussion of these examples below. The immediate point is that the existence of questionable and infelicitous sentences in (97) suggests that the proportional operator in the CC is not a default that can be over-ridden but is rather an operator that is part of the meaning of a lexical item (or something conventional within

24 I did find the following example in a play from 1700 (The Way of the World by William Congreve):
(i) The greater the coxcomb, always the more the scandal; for a woman who is not a fool can have but one


144
the sentence). The contrast with the examples in (98) shows that the CC is unlike the conditional in this respect. Further evidence comes from the following chart of Google counts for conditionals and CCs with various preceding AQs:

<table>
<thead>
<tr>
<th>AQ Type</th>
<th>CC Total Hits</th>
<th>Conditional Total Hits</th>
</tr>
</thead>
<tbody>
<tr>
<td>All AQs + No AQ</td>
<td>111,000,000</td>
<td>783,000,000</td>
</tr>
<tr>
<td>Always</td>
<td>38,100</td>
<td>2,090,000</td>
</tr>
<tr>
<td>Usually</td>
<td>5,290,000</td>
<td>3,020,000</td>
</tr>
<tr>
<td>Often</td>
<td>7,070,000</td>
<td>2,770,000</td>
</tr>
<tr>
<td>Sometimes</td>
<td>6,700,000</td>
<td>5,090,000</td>
</tr>
<tr>
<td>Occasionally</td>
<td>9</td>
<td>730,000</td>
</tr>
<tr>
<td>Seldom</td>
<td>0</td>
<td>2,230,000</td>
</tr>
<tr>
<td>Rarely</td>
<td>4</td>
<td>6,390,000</td>
</tr>
<tr>
<td>Never</td>
<td>5,950</td>
<td>540,000</td>
</tr>
</tbody>
</table>

Table 3.4: CC search term “the more the”; Conditional search term “if the” (eliminating “not always”, etc.)

On the basis of the values in this table, there is no way to maintain that the conditional and CC are parallel with respect to AQ modification. Even factoring in the difference in frequency between CCs and conditionals, it would still be surprising that CCs occur so much more often with usually or often than conditionals and that CCs occurring with occasionally, seldom and rarely are literally in the single digits. Of course, there is more work to be done in describing these patterns apart from presenting frequency counts. For example, almost all of the hits for CCs with always were cases in which the CC was being quoted, in the context of someone complaining “it's always 'the more...’” where they mean that someone is constantly telling them this. Also, none of the hits for rarely or occasionally were actual matches; most had the AQ at the end of a sentence prior to an unmodified CC.
Then, too, the counts for the conditional do not tell the whole story, either. The hits for *seldom* often involved subject-verb inversion in the matrix clause (e.g. *seldom if the child is born deaf do the parents not opt for surgery*) which is not possible in the CC, as we know (and though it would be interesting to look for instances of *seldom* modifying RCCs partially for this reason, they are much more difficult to search for reliably). The data for *rarely*, however (to my surprise) were largely exactly the kind of instances that were the target. *Never if the* often yielded hits such as *It's now or never if they think...* or others that were not instances of sentence-initial modification of a conditional. Regardless, even if you could come up with stories for why those might not work with the CC for other (grammatical) reasons, the difference between *sometimes* and *occasionally* would remain a mystery. Looking closely at those two shows that, even if the relative infrequency of *occasionally* were accounted for in addition to the overall frequency difference between conditionals and CCs, *sometimes* occurs more often than expected in CCs and *occasionally* occurs less often than expected in CCs. Detailed observations about these patterns are beyond the scope of this dissertation, but the point is that the fact that not all AQs are possible makes the idea that the proportional quantificational force in the CC is encoded in *the* rather than an implicit operator plausible.

The remaining question, then, is how to explain which AQs can occur with the CC. If we are to accept the idea that *the* encodes proportional force, we must explain why a CC can occur with certain other AQs. This is related to exactly what the nature of the proportional force in the CC actually is. I would like to put forth the following functional story: the CC in English is a primary way in which humans can convey
information about correlations or co-occurring changes. Though we may, in contemporary society, think about correlation in terms of correlational coefficients and statistical significance, there must be a much more primitive sense that humans have for when different variables are related, which may sometimes even be at odds with what statistics reveal. This is presumably slightly different in different people such that some are willing to assume a correlation in light of less evidence than others. I propose that it is this intrinsic notion of correlation that is expressed through the utterance of a CC in natural language. Thus, though the quantificational force inherent to a CC patterns with most, it is not, strictly speaking, identical to most. For some, a simple majority of cases in a relationship of monotonic correlation may be enough to establish the truth of the CC, while for others, 95% of cases might be required. For the remainder of this thesis, I will call this quantificational force *correlational force*, and I will encode it in the operator CORREL, meant to be parallel to MOST as an operator that takes two arguments and returns a truth value if there is a certain relation between those arguments (in this case, the relation of correlation). In my analysis in the next chapter, we'll see that both of CORREL's arguments are sets of pairs, where each of those pairs contains two items that differ along some dimension. Then CORREL requires that a majority (or more, etc. depending on what each individual speaker takes as necessary for establishing a correlation) of the pairs in the denotation of the antecedent to also be pairs in the denotation of the consequent (representing a change along a second dimension).

To summarize what I've said above, the meaning of the CC describes a correlation, a situation in which two or more attributes/measurements over the same group co-vary to a certain degree in a certain direction. If there is a difference in one
attribute, there is a difference in the other a certain percent of the time (co-variation).

Given this meaning of the CC, it is easy to see that it is black and white: two things are correlated, or they are not; there is no in-between. This is true of contemporary statistic significance as well: a correlation is either significant or it is not; it cannot be somewhat significant in technical terms. Saying that a correlation is stronger than expected, e.g. that the relationship it expresses always holds is fine, but saying that it occasionally holds does not make sense, unless you are relativizing the results to different domains and saying that in some data sets it does and in others it doesn't. Now let's consider the meanings of various AQs. If sometimes, occasionally, etc. have as their meaning that there exists at least one time such that the denotation of the sentence holds, then it is not inconsistent with their entailment that the denotation hold at all times. This would depend upon whether speakers think that the following is a contradiction: sometimes, if I buy shoes, I wear them the next day; in fact, I always do. If this is judged to be a contradiction, it would mean that the 'not often' association with these AQs is entailed, in which case AQs like occasionally, sometimes, and seldom contradict what is being added to the Common Ground by the meaning of the CC (by majority) and should be infelicitous. If it is judged not to be a contradiction, then the 'not often' force of these AQs is an implicature, in which case their appearance with a CC would still be infelicitous, but because their Quantity implicature conflicts with what is entailed by the CC. Either way, these AQs would not be able to modify the CC, just as they would not be able to modify conditionals that already contain a majority-requiring quantifier such as most, as demonstrated in (99).
(99) a. #Some years, if I file my taxes most years, I'll get a refund.
    b. #Occasionally, if I often go to the park, I'll see dogs.
    c. #Seldom do I always write checks, if I can use a debit card instead.

Always, often, usually, etc., on the other hand, should be able to be accommodated by the Common Ground as they merely strengthen the majority sense of the CORREL operator. This would nicely account for the difference in the AQ pattern between CCs and conditionals if it weren't for the fact that sometimes, unlike the others in the non-majority category, does appear with the CC (and not just marginally, it appears in CCs more than conditionals!). I hypothesize that the non-majority AQs that occur with CCs quantify over a different domain than the one quantified over by the proportional quantifier inherent to the CC, and I further propose that their presence strongly implicates that the CC being modified is a correlation that holds contrary to expectation. Let's consider some particular examples.

(100) Sometimes the more force is used, the less effective it is.\(^{26}\)

(101) Sometimes, the more the dogs pull, the better.\(^{27}\)

In (100), the meaning of the CC quantifies over pairs of possible worlds or hypothetical situations in which the military uses different amounts of force for a given operation, saying that if more force is used, the operation will be less effective. The AQ sometimes is quantifying over operations, however, saying that there are some kinds of operations for which this correlation holds, namely those in which there is a greater chance of collateral damage (from the discourse context on the website). It's also the case that this...

\(^{26}\) http://www.press.uchicago.edu/Misc/Chicago/841519.html

\(^{27}\) http://www.pedigreedatabase.com/gsd/bulletins_read/318452.html?pagen=1
CC is the opposite of what some might expect. Anyone who watches T.V., for example, will see constant (fake) evidence that when someone needs to get information from someone else, more force (in the form of torture) leads to a more effective outcome (getting the information). Another example I found said that, when backpacking in Eastern Europe, it's typically better to go by yourself or with one other person in order to stay with and get to know local people, but that sometimes, the more the merrier, meaning that if you don't care that much about mixing with the local culture, then more people will always add up to more fun. So *sometimes* was restricting the domain in the same way that *if you're not interested in local culture* would. The difference is that *sometimes* is less explicit in its semantic domain restriction, relying more on the pragmatic retrieval of past discourse information to determine which situations were relevant. Just as prior discourse can determine the restriction of *sometimes*, so too can real world knowledge. This is related to the 'contrary to expectation' association with *sometimes* when used with a CC comes in, I believe. It's that you can determine from world knowledge when this relationship holds as opposed to a relationship in the opposite direction (the cline in the first clause correlating with increases as opposed to decreases or vice versa in the second clause). This is certainly the case in (101), written by a woman who lives in a hilly area of Duluth, Minnesota. Real world knowledge tells us that people walking dogs normally find it annoying when a dog pulls forcefully on its leash but that the dog's force would mean less effort required by the dog-walker on inclines. Thus, the *sometimes* restriction indicates that the correlation is true for those cases in which you are walking uphill with dogs. It also indicates, via its implicature or entailment that this is not true of very many cases, that the instances in which you are
walking uphill do not constitute a majority of time spent walking dogs and that this correlation does not hold at other times.

This explanation explains why we simultaneously find many felicitous examples of CCs with *sometimes*, but it also explains why some examples such as (102), restated from (97b), are not typically felicitous.

(102) #Sometimes, the more educated John becomes, the better his career prospects.

This example was created to demonstrate what happens when *sometimes* and the CORREL operator quantify over the same domain, both in terms of semantic type and size/makeup. Here, we are restricted to a single individual and events in the real world so that the most natural interpretation is in terms of pairs of times at which John has different levels of education. *Sometimes*, too, is best understood as quantifying over times. As you can see, this leads to infelicity as predicted above. In a context where you understand *sometimes* to mean something different, such as *some decades*, then it becomes felicitous again. Note that this is different from what we find with conditionals, where AQs can quantify over the same or different domains as compared to their conditional antecedents. Compare (97b) to (98b) or (103) to (104).

(103) #Some years, the greater the total John made that year, the happier he was.

(104) Some years, if John made lots of money that year, he was happy.

Thus, whereas conditional antecedents can provide the complete restriction of an AQ, CCs can only provide the complete restriction of AQs that are consistent with their
quantificational force, and otherwise, they must only provide a partial restriction.

This explanation covers all of the data so far except for the difference between *sometimes* and other less-than-most operators in the CC. If we assume that *sometimes* and *occasionally* mean approximately the same thing and that they share a similar syntactic distribution pattern (as opposed to say *seldom*), then the lack of *occasionally* with the CC is the most puzzling. However, there is a difference in implication between them that explains the difference. *Occasionally* implies that the times when the event it is modifying occurs are random or not predictable. When set of times is predictable, *sometimes* is used. So, if someone were talking about a chocolate shop they liked and you happen to meet friends for chocolate there every other Thursday, you might respond that you sometimes go there with friends, but it is less likely that you would say that you occasionally go there with friends. You might say that you occasionally go there with friends if every random time you happen to stop there, you have friends with you. Because the CC necessarily represents a pattern, therefore, *sometimes* will appear more often than *occasionally*. All this leads to the following benchmark:

(B37) *Indefeasible Force*: The correlational force of the CC cannot be altered by an AQ.

We have now seen a number of ways in which the CC differs from the conditional. But many of the benchmarks adopted on the basis of Beck's data show parallels between CCs and conditionals, so my claim that CCs are not conditionals requires an explanation of these similarities.

The primary similarities between the CC and the conditional were the benchmarks (B7: Backward Pronominalization) and (B30)-(B34). Those included
patterns related to NPIs and AQs, both of which have been shown to have differences when considered closely. The other patterns related to donkey anaphora, a lack of future tense, backward pronominalization, and the ability of both to quantify over a variety of different semantic types. These similarities can easily be explained on the basis of the fact that both are generalizations and both are instances of the overarching category of Generalized Quantifiers (if one accepts a GQ theory of conditionality). We can see that all of these properties exist outside of the CC and conditional in other sentence types that are GQs and sentences that denote generalizations.

(104) Donkey anaphora: Every farmer who owns a donkey beats it.

(105) No future tense: #Every time Sue will board a train, John will wait to greet it.

(106) Backward pronominalization: a. However long he, has to wait, John, is irritated.
    b. #He, is irritated, however long John, has to wait.

(107) Variability of Quantification: a. Everyone fishes. (individuals)
    b. Every time I fish, I catch something. (situations)

In fact, treating the first clause of CCs as a conservative GQ actually explains the presence of donkey anaphora if Chierchia (1995) is right that Conservativity is at the root of all donkey-sentence-hood. Now, it should be noted that because it is based on the work of Kratzer, and von Fintel's analyses of conditionals, Beck's analysis of the CC is also a Generalized Quantifier account of the CC. However, for her, the CC is only a GQ because the conditional is, and the GQ's quantificational force comes from an adverb of quantification that may be implicit or explicit, as in the conditional. I have argued in these past sections that the quantification force is fixed, more similar to most (another GQ) than the conditional, and part of the conventional meaning of a lexical item in the
CC. Thus, these are two very different approaches to the CC as involving a GQ.

3.2.2.3 Than Phrases in CCs

In this section, I evaluate Beck's claim that than phrases cannot occur in the CC. I show that for many speakers, they can, but that this is not a problem for Beck because these than phrases parallel another phenomenon in which there are multiple elements that would fill the comparative's than phrase argument with only a single comparative morpheme.

Beck deems all sentences such as (108) ungrammatical, suggesting that regular comparatives, but not the comparatives in CCs, can occur with overt than-phrases. Her analysis requires that the feed the comparative morpheme a single argument twice, filling the position where a than phrase would occur, effectively explaining why an overt than phrase is missing in the CC.

(108) *The slimier one lawyer than another, the more efficient he is.

There are times, however, when a comparison is used to restrict the domain of consideration. In a sentence like (109), there is no comparative morpheme, but we understand that the phrase as compared to is restricting the domain for us to judge whether someone is slimy. For some speakers, CCs with than phrases are alternative paraphrases for sentences like the one in (109). These appear in (110)-(112) along with their paraphrases. Unlike the than phrase in (108), which could have been uttered more simply as the slimier the lawyer, the more efficient he is, the than phrases in these CCs potentially make a difference in the truth conditions of the CC.
(109) As compared to one lawyer, the slimier another is, the more efficient he is.

(110) a. The faster the cat is than the dog, the more likely it is to get away.
    b. As compared to the dog chasing it, the faster the cat, the more likely it is to get away.

(111) a. The faster we drive than 35 mph, the more expensive our speeding ticket will be.
    b. Starting at 35 mph, the faster we drive, the more expensive our speeding ticket.

(112) a. The faster you wash dishes than I dry them, the more water we’ll end up with on the counter.
    b. As compared to my drying speed, the faster you wash dishes, the more water we’ll end up with on the counter.

(113) The faster you wash dishes, the more water we’ll end up with on the counter.

In order to understand what a sentence like (112a) means and how it makes a truth-conditional difference, compare it to (113). First picture a scenario in which a dishwasher becomes more efficient at washing dishes with each plate he washes, placing each one on the sink ledge once he has finished, where the dish-dryer is supposed to pick it up immediately so that water does not keep dripping onto the counter and subsequently, the floor. However, if the dish dryer becomes less efficient with each plate he dries (possibly because his dish towel becomes more saturated with each dish and thus less able to dry it quickly), then as the washer’s rate increases and the dryer’s rate decreases, the plates will sit on the ledge longer and longer between being put down by the washer and picked up by the dryer, thus leading to more and more water spilling onto the floor. Now picture a model in which, toward the end, the washer realizes that the dryer has slowed down and water is getting on the floor, so he slows down a bit too (but by that time, it's too late; more water is going everywhere). That model is given in (114).
(114) Your dishwashing speed  
    My drying speed  
    Difference (wash – dry)  

<table>
<thead>
<tr>
<th></th>
<th>3</th>
<th>4</th>
<th>6</th>
<th>7</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Water on counter</td>
<td>0</td>
<td>3</td>
<td>6</td>
<td>8</td>
<td>12</td>
</tr>
</tbody>
</table>

In this model, (112a) is true, but (113) is false (if we assume that with so few data points, all of them have to fit the pattern for a sentence to be true). That's because there is a monotonic correlation between the differences and the amount of water on the counter, but not between the dishwashing speed and the amount of water on the counter. So the contribution of the *than* phrase makes a difference in the truth conditions of the sentence. There is a difference between (112a) and some of the other CCs with *than* phrases such as (111a). The value that the *than* phrase contributes is fixed in (111), but variable in (112) (as we saw in (114)). Interestingly, of the many, many native speakers that I have asked for judgements about CCs with *than* phrases, there are four basic groups of speakers. Some find all of the *than* phrase examples in (110)-(112) felicitous and some find none felicitous, but the remaining speakers typically feel very strongly that either examples like (111) or examples like (112) but not the other are felicitous. In other words, when I ask about a long list of CCs with *than* phrases, the variability or lack thereof in the value contributed by the *than* phrase is the prime dividing line when it comes to native speaker judgements. The following, in addition to those above, are some of the examples that I asked about or that were suggested to me as felicitous when talking with speakers. Examples (125) (David Dowty, p.c.) and (126) (Carl Pollard, p.c.) are naturally-occurring.

---

28 Here, let’s say that dishwashing and drying speeds are measured in dishes/minute and that the amount of water on the counter is measured in milliliters (it doesn't really matter what the units are).
(115) The taller John is than you think he is, the more money I win. (den Dikken 2005)

(116) The faster you run than your friend, the further away you’ll be when the bear attacks him.

(117) The happier you are than me, the more I’ll turn to you for support.

(118) The taller a man is than Mr.T, the more likely he is to be able to beat him at arm-wrestling.

(119) The smarter someone is than her friend, the less likely they are to have similar musical tastes.

(120) The older Jane is than her sibling, the lonelier she’ll be. Yes, but the older she is than the baby, the more helpful she can be in taking care of it.

(121) The faster I can stuff the envelopes than you can lick the stamps, the longer the coffee break I can take while you catch up with me.

(122) The smarter than your opponent you are, the better your chances.

(123) The faster you drive than the speed limit, the more likely the police are to stop you.

(124) The taller you are than 36", the more likely they’re going to let you on the ride.

(125) In many cases the older the appliance the more energy it uses than a newer model. 29

(126) The older a man is than his female partner, and the more money and gifts he gives her, the less likely they are to use condoms. 30

On the basis of these examples, there are a few things that need to be investigated. We need to account for the possibility of than phrases but also for the possibility that some speakers do not find these examples felicitous. We also need to determine whether these are a serious problem for Beck's analysis.

I propose that than phrases, when felicitous, are not licensed by the (overt) comparative morpheme in the CC clause. I believe that than phrases in the CC have the

29 http://morethanswitchingoffthelights.blogspot.com/

30 http://www.guttmacher.org/media/nr/2005/04/21/index.html
same status as the second *than* phrase in (127) or two of the three *than* phrases in (128), examples which, like those with *than* phrases in the CC, are judged by some speakers to be felicitous and others not.

(127) John is much taller than Mary than Bill is.
(128) John is taller than Mary than Bill is than Sue.

In each of these cases, we find more *than* phrases than comparative morphemes to license them. In (127), the intended meaning is that Bill and John are both taller than Mary, but because John is much taller than Bill, John's height difference with Mary is much greater than Bill's height difference with Mary. Similarly, in (128), the difference between John and Mary's heights is greater than the difference between Bill and Sue's heights. An analysis of these examples is beyond the scope of this thesis, but I hypothesize that however they are handled might be extended to the analysis of CCs with *than* phrases. This way, the variability in judgments is accounted for, and Beck's account can be maintained with respect to this particular point (because an overt *than* phrase would not be occupying the spot for the second argument of the comparative morpheme).

If, instead, we took the presence of *than* phrases to indicate that a difference of degrees (as in the case of dishwashing and drying speeds, above) was the default for the CC (as we will see that Brasoveanu 2009 does), we would need to explain why we do not often find examples of them (or at least how to account for the examples without them). This is extremely tricky, because a default needs to be assigned in the case where no *than* phrase appears. The best option is to assume that the value is filled in contextually, just
as it is in the case of so-called Discourse Comparatives (Hendriks and deHoop 2001), as in (129).

(129) John is six feet tall. LeBron James is taller.

In the second sentence of (129), there is a comparative without a than phrase, just as in most CCs. Nevertheless, you know who LeBron James is being compared to because of previous discourse. The speakers I consulted agreed that there was a significant difference between examples like (129) and CCs in that, with (129), they were searching for what to fill in as the comparison class, whereas with CCs, they were not. This difference would be a significant problem for any analysis of CCs that takes each CC clause to have than phrase, whether overt or covert.

3.3 Lin 2007

Lin 2007 is largely concerned with giving a compositional interpretation for Mandarin Chinese sentences that have a meaning similar to that of English CCs (I will call these 'Mandarin CCs' in order to make the discussion simpler despite their potential differences from English CCs). Lin shows how this same approach can be extended to English. In this section, I will focus on his analysis of English, as I have done elsewhere.

He states that the Mandarin CC is causal and conditional. He argues that they are not if-type conditionals; they are what he calls 'bare conditionals', which others have called correlatives. This is roughly the difference between if they went, they had fun and whoever went, they had fun, except that in Mandarin, these have different presuppositions in terms of whether anyone had to have gone (or done whatever is
described in the first clause). The *if* conditional does not presuppose that anyone went, but the bare conditional does. Lin presumably thinks that this presuppositional difference is related to the difference in syntactic form between the two kinds of conditionals because he goes on to say that the Mandarin CC has a semantic form more like the *if*-type conditional, but a syntactic form more like the bare conditional (as just mentioned).

Each clause in a Mandarin CC includes the word *yue*. Lin shows that *yue* can’t be like *the* because it can’t occur with another comparative item in its same clause; it is the comparative item. But then, if *yue* is like *more* and not *the*, Beck’s analysis can’t be straightforwardly applied. Also *yue* can have a *than* phrase, and accordingly, so can the Mandarin CC, so Lin argues that this is a second reason that Beck’s analysis cannot be directly applied. He proposes two lexical entries for *yue* depending upon whether it quantifies over situations alone or individuals in addition to situations, and these are the only possibilities that he mentions. The lexical entry for *yue* quantifying over both individuals and situations appears in (130), where *g* is variable of degree type. This entry is like Beck’s but the degree predicate takes an extra argument, and *more* itself takes more arguments in the form of individuals, degrees, and situations.

(130) \[ \lambda[[yue]] = \lambda x \lambda y \lambda g_1 \lambda g_2 \lambda s_1 \lambda s_2 [P(x)(g_1)(s_1) \land P(y)(g_2)(s_2) \land g_2 > g_1] \]

Lin analyzes the CC as a whole in a somewhat constructional way by assigning a meaning to a template composed of an implicit or explicit AQ along with the word *jiu*, meaning *then* that often connects clauses in conditionals. He does not say anything about
what this item means outside the CC, but within the CC, it is the locus of a causal operator \( R \) that remains undefined. In sentences without \( jiu \), the causal operator is contributed by a null lexical item. The former case, in the form of a template for the meaning of a CC, is illustrated in (131).

\[
\begin{align*}
\text{In (131), the } s \text{ variables are all situations, and the } \leq \text{ sign indicates that one situation is a minimal situation of the other larger situation. All of the } P(...) \ldots \text{ and } Q(...) \ldots \text{ and } d > d' \text{ terms are presumably contributed by the two } yue \text{'s, and the AQ binds variables and perhaps contributes the material conditional, so that leaves the requirement that that the situations in the antecedents be minimal with respect to the situations in the consequent, the } R(...) \text{ terms and the existential quantification in the consequent as the contribution of the construction and/or } jiu. \\
\text{In extending his theory to English CCs, he gives the following analysis, including the Logical Form in (134), the denotations of lexical items in (132), and the sentence-level meaning in (133) for the sentence } \text{the busier John is, the happier I am.}
\end{align*}
\]

\[
\begin{align*}
\text{(131)}
\end{align*}
\]

\[
\begin{align*}
\text{In (131), the } s \text{ variables are all situations, and the } \leq \text{ sign indicates that one situation is a minimal situation of the other larger situation. All of the } P(...) \ldots \text{ and } Q(...) \ldots \text{ and } d > d' \text{ terms are presumably contributed by the two } yue \text{'s, and the AQ binds variables and perhaps contributes the material conditional, so that leaves the requirement that that the situations in the antecedents be minimal with respect to the situations in the consequent, the } R(...) \text{ terms and the existential quantification in the consequent as the contribution of the construction and/or } jiu. \\
\text{In extending his theory to English CCs, he gives the following analysis, including the Logical Form in (134), the denotations of lexical items in (132), and the sentence-level meaning in (133) for the sentence } \text{the busier John is, the happier I am.}
\end{align*}
\]

\[
\begin{align*}
\text{(132)}
\end{align*}
\]

\[
\begin{align*}
\text{(133)}
\end{align*}
\]

\[
\begin{align*}
\text{(134)}
\end{align*}
\]
One of the most surprising aspects of this analysis is the denotation of *the* according to Lin. For Beck, *the* and *three inches* or other measure phrases compete for the same spot. The meaning of *the* feeds one argument to the comparative twice with two different variables, explaining a lack of *than* phrases. Lin proposes that *more* have the same denotation as *yue*, then suggests that *the* compete with a *than* phrase (rather than a measure phrase, as in Beck) for its spot. For him, *the* picks up a copy of the subject of the clause for its denotation, which explains why its denotation is *john'* in (132). The way in which he accounts for the possibility of quantification over individuals v. situations (or stages of an individual) is as follows. If you have an indefinite NP in subject position in a CC clause, then *the*, when it copies that variable with its condition, for example *x|student'(x)*, could result in an exact copy (*x|student'(x)*) or a copy that is an
alphabetic variant of the original, as in $y|\text{student}'(y)$. In the former case, quantification takes place over situations/stages, and in the latter, individuals.

Lin's main focus and contribution is picking up where Beck left off in the analysis of CC examples with indefinites that could receive an individual or time/situation/stages of an individual reading. The same reading possibilities arise in Mandarin CCs with bare NPs in subject position. For Beck, the preferred reading for a sentence like the slimier an attorney is, the more successful he is is the one quantifying over individuals, not times or stages, which is the opposite of what Lin predicts for Mandarin. Beck just disregards what she considered the 'generic' reading (the one quantifying over times but also containing an individual variable that needs to be bound elsewhere), and that's where Lin's work fills in a gap. His explanation is partially in terms of processing, following Partee 1986 in suggesting that we try lower-type interpretations before higher-type ones. In this case, Lin says that indefinites are individual-type ($e$-type) pronouns in the case of the quantification over times readings and that they are properties ($<e,t>$) in the case of quantification over individuals. This amounts to the difference between saying that an attorney is $x|\text{attorney}'(x)$ (where this is just an individual variable with a condition on it) or attorney$'(x)$. Then, your processing constraint tells you to try to make the $e$-type pronoun work before trying the property-type denotation.

There is a second dimension that also influences the kind of quantification that is possible in these cases. Lin observes that when a Mandarin CC clause contains an individual-level predicate, it can only have the interpretation in which you range over individuals. With a stage-level predicate, the interpretation in which you range over times is preferred. He proposes that stage-level predicates can take their $e$-type indefinite
arguments via function application and that there is a second rule of property conjunction
for getting the property-type indefinite to combine with an individual-level predicate. By
comparison, in Lin's analysis of English, we saw that he used the and what it copied to
determine which reading would result. For Mandarin, if the pronoun combines via
functional application, it does not get counted twice, leading to an interpretation
quantifying over times, and if it combines via property conjunction, it's value does get
entered twice to yield quantification over individuals. In other words, Lin assumes that
the stage-level v. individual-level predicate distinction is responsible for different
readings in Mandarin but not English.

Independently, however, I had noticed a way in which a similar distinction was
relevant in English. In this chapter as in the previous literature, the focus has been on the
CC and how it compared to the conditional. But there are a variety of kinds of sentences
that have meanings similar to the CC that are not considered in any detail, including
(135), which has a meaning comparable to (136).

(135) As he gets older, he gets taller.
(136) The older he gets, the taller he gets.

Just as with the conditional, these seem equivalent at first glance. However, I noticed a
difference in terms of quantificational possibilities such that as... can only be understood
as quantifying over times, while a CC has a wide range. They appear similar in (135)
and (136) because I chose a CC that also quantified over times. If I choose instead an
individual-level predicate that cannot be construed as changing over time, the as...
construction becomes infelicitous, unlike the CC.
(137) #As a semanticist's birthday is earlier in the year, her husband is taller.

(138) The earlier in the year a semanticist's birthday occurs, the taller her husband.

Thus, a pattern that Lin noted for Mandarin CCs is useful in distinguishing similar meanings between CCs and other kinds of English sentences. I now turn to discussing a few aspects of Lin's theory that are problematic.

As with Beck, one of the problems for Lin is the conditionality of the analysis, including the assumption of implicit universal force. There is a further problem in that Lin assumes that causality is a part of the CC's meaning, but it is not clear whether he intends for that part of his analysis of Mandarin CCs to extend to English. There are a number of issues that are theoretical rather than empirical, all of which fall under the heading of being ad hoc or unmotivated. Certainly, positing a lexical entry for *more* that is different from the one used in other comparative sentences is a disadvantage as compared to Beck's analysis. Similarly, positing different pronouns, different rules of combination, and a processing constraint all toward the same effect seems inefficient. Finally, leaving some of the meaty questions about how the meaning of the CC is composed to a general template misses the full extent of compositionality that is possible.

The main empirical problems for Lin are pointed out by Liu 2008 and noticed independently by myself. One of the most obvious flaws of Lin's analysis is that it does not account for (B5) – the mandatory fronting in each CC clause – predicting instead that the *the more* phrases in each be in situ. A second problem relates to which meaning is considered the default. As noted above, Lin is hypothesizing that you always try the e-type meaning first (comparison over times only) unless you have bump it up to the
property-type indefinite, in which case you compare individuals. On the basis of my knowledge of the English CC data, I was surprised that this would be a processing effect rather than a context effect in the Mandarin CC. For example, different prior discourses in (139a) and (139b) lead to different interpretations of another speaker's response in (139c) in keeping with what it takes to make that statement relevant to the topic of discussion.

(139) a. Context biasing times: My son is an attorney, and I'm proud of him, but I fear that his firm is changing him and that his is becoming less respectable in his practices. The seem more and more slimy when I hear about the victories.
   b. Context biasing individuals: I need to find an attorney to represent me in a lawsuit, but I just find them so slimy. I want someone who will get the job done quickly and not make my skin crawl.
   c. Well, you know, the slimier an attorney is, the more efficient he is.

Though Liu does not provide an example showing the influence of prior discourse, he does give the example in (140), which translates as *the smaller the balloon, the harder it is for it to fly*. For this sentence, he says, the default is quantification over individuals, contra what Lin's processing strategy would predict.

(140) Qiqiu yue xiao, yue bu rongyi fei-qi'ai.
      Balloon more small more not easy fly-up

Liu includes many other criticisms of Lin 2007, but most others are specific to the analysis of Mandarin. Liu 2008 naturally puts forth his own analysis of Mandarin CCs, but I will not describe it in any detail here as it makes the same fatal assumptions as the other theories that CCs in Mandarin and English require implicit AQs, etc. without introducing any new benchmarks or raising issues for the analysis of English that need to
be discussed here. Though Brasoveanu 2009 also gives an analysis that assumes
universal quantification (via maximality of a \textit{wh}-operator), it raises significantly different
issues and will be addressed next.

\textbf{3.4 Brasoveanu 2009}

In this section, I once again divide the description of the analysis from the discussion of
its strengths and weaknesses.

\textbf{3.4.1 Brasoveanu's Analysis}

Every contemporary theory of the semantics of CCs that we have seen so far
gives them the meaning of a conditional. Brasoveanu is the first to explicitly deny that
all CCs are conditionals. However, this is primarily because he extends the category of
CCs to include all sentences with correlative clauses that begin with the Romanian
equivalent of \textit{how much}, as we will see presently. After forming this unified category of
CCs, he divides them into two groups: conditional and non-conditional. Not surprisingly,
the new additions to the group are the non-conditional ones, and the sentences with
meanings that parallel CCs in English are the conditionals. Thus, even though
Brasoveanu is the first to suggest that some CCs are not conditional, he still believes that
all English examples would be conditional, which is not a change from past approaches.
What Brasoveanu means by conditionality is somewhat different from what others mean,
though. For him, \textit{conditional} is shorthand for 'universally quantifies over something in
the first clause that is existentially bound in the second' and \textit{non-conditional} is
'existentially quantifies over something in both clauses,' as far as I can tell from his
discussion of various phenomena.
As we will see below, one difference between Brasoveanu and the other approaches reviewed so far is that he proposes what he calls a *relational* analysis, which is similar to what Beck would have called a *functional* analysis except that, obviously, his is relational rather than functional. Based on Beck's criticism of any functional approach, we know that this is an important difference and that Brasoveanu will be able to avoid the pitfalls of a functional approach by advocating for a relational one (since we determined that monotonic correlation was necessarily relational rather than functional).

What the relational approach really boils down to for Brasoveanu, however, is the difference between quantifying over worlds, times or individuals and quantifying over degree differences. He proposes that the CC asserts that a particular kind of relationship (monotonic correlation is a good candidate) holds between two sets of degree differences. This is very different from Beck 1997, for example, where the content of the relation is built on the basis of the denotation of its lexical items, particularly the interaction between *the* and *-er*. What we will see, however, is that Brasoveanu also builds a monotonic correlation on the basis of lexical items, rendering the content of the relation useless to the semantics of CCs in the cases where the relation is one of monotonic correlation.

Another difference between Brasoveanu and past approaches is that he is the first to distinguish the meanings of the two *the's* that occur in the CC. The one that occurs in C1, I will call *the1*, and the one that occurs in C2, I will call *the2*. For Brasoveanu, the quantificational force in the CC comes from *the1* as in the proposal I made in 3.2.2.2.

Turning now to more specifics of his analysis, Brasoveanu further divides the conditional and non-conditional categories described above into equative correlatives and
comparative correlatives, maintaining that all four categories should be given a unified analysis. The difference between equative and comparative correlatives for Brasoveanu does not have to do with their semantics, which is a sense of potential confusion. Instead, comparative correlatives are those that contain a comparative morpheme, and equatives are those that do not (but still begin with the relativizer how much). Examples from Romanian for each of these categories appear in (141)-(144), along with his English translations. As with Lin 2007, understanding the motivation behind Brasoveanu's analysis for the language of interest to him will help us to understand his analysis as it extends to English (which, like Lin, he explicitly claims that it does).

(141) **Non-conditional Comparative Correlative:**

    Cu cât e mai înalt fratele decât sora, (tot) cu aștăzi e mai înalt tatăl decât mama.
    With how much is more tall brother the than sister the, (also) with that much is more tall father the than mother the

    ‘The brother is taller than the sister by a certain amount and the father is taller than the mother by the same amount.’

(142) **Non-conditional Equative Correlative:**

    Pe cât e Irina de frumoasă, (tot) pe aștăzi e de deșteaptă.
    PE how much is Irina DE beautiful, (also) PE that much is DE smart

    ‘Irina is beautiful to a certain, significant extent and she is smart to the same, equally significant extent.’

(143) **Conditional Comparative Correlative:**

    Cu cât e un avocat mai agresiv, cu aștăzi e mai eficient.
    With how much is a lawyer more aggressive, with that much is more efficient

    ‘The more aggressive a lawyer is, the more efficient s/he is.’
One thing that is immediately obvious in terms of the glosses is that both of the equative correlatives contain particles *pe* and *de* that are not defined. Brasoveanu never mentions these and clearly considers *pe cît* and *cu cît* to be identical in that, when he gives lexical items for these sentences, he labels them as *cu cît*. It is also obvious that both of the comparative correlatives contain the comparative morpheme *mai*. They both permit *than* phrases, though only the non-conditional is shown with one here. The conditional comparative correlative examples like (143) that contain *than* phrases in Romanian are not unlike those we saw in the last subsection of the discussion on Beck. Brasoveanu gives an example that translates as *the greater a natural number is than another, the greater its square is than the square of the other one*. We also see that both of the non-conditionals can appear with the word *tot* ('also'). In this case, their meanings require that the relation between the degrees in both clauses be one of identity. Brasoveanu posits that the identity relation be a presupposition of *tot*.

In terms of the sentence-level meanings of these sentences, (143) and (144) have roughly the same meaning according to Brasoveanu. I will assume that the only difference is whether the CC denotes a relation or a stronger linear function (in the case of the equative). The meaning of (142) suggests that if the CC is true, Irina must be both beautiful and intelligent. He notes that this is not required of non-conditional equative correlatives in general, giving an example that means *the older the tree, the thicker its*
Brasoveanu argues that one reason examples like (141) and (142) can be said to not be conditional is because they require the existence of differentials (degree differences), which in turn require the existence of events or individuals that have the properties in question to a certain degree that can be used to compute a differential, unlike the antecedent of a conditional. Indeed, due to its hypothetical nature, the existence of a height increase in (145) is in question, whereas it must exist in (146).

(145) If Dad is taller than Alex by some amount, then I'm taller than Mom by that much.
(146) However much Dad is taller than Alex, I'm taller than Mom by that much, too.

Brasoveanu shows that the same is true for conditional CCs in the past tense:

(147) If he drove faster, then he got there sooner.
(148) The faster he drove, the sooner he got there.

Here, the CC implicates the existence of multiple events that meet the criteria of the restriction in a way that the conditional doesn't. We'll see that this plays a very important role in Brasoveanu's analysis. He states that a lack of hypotheticality on the part of all CCs is related to the fact that they quantify over differentials and not worlds. What would be quantification over times or individuals is also quantification over differentials in this analysis. On the difference between the time and individual quantification difference for (143) (or its parallels, discussed by Beck and Lin), Brasoveanu says only that “these two interpretations are not necessarily two distinct readings, since [quantification over times] is ultimately just a refinement of [quantification over...
individuals] that examines the aggressiveness and efficiency of lawyers at various times as opposed to a single, contextually salient temporal interval” (2009: 3).

Brasoveanu's argument is that it is crucial to have the anaphora be between differential degrees in the two CC clauses because it is possible for \textit{tot} to occur with conditional CCs (in addition to its presence in the non-conditionals as we saw above). In that case, the differentials must be the same across clauses (the same amount, etc.). Thus, there is a restricted relation (co-reference) between degree differences, so if the degree differences are not available for quantification (as they are not in Beck's analysis), it would be difficult to account for those examples. It's not identical, but in English, too, it is possible to add a restriction to a CC, as in \textit{the faster we drive, the faster the car behind us drives, to the exact same degree!} This is a reason to prefer a more complicated form for the comparative in which it is \([y - x = d \text{ and } d > 0]\) rather than simply \(x>y\). Beck does use the more complicated meaning in her analysis, but obviously stops short of quantifying over the degree differences.

The way that these differentials are obtained varies among (141)-(144). In the examples with comparatives and \textit{than} phrases, it is straightforward: you subtract the height, etc. of the entity in the \textit{than} phrase from the height, etc. of the clausal subject. In the case of the non-conditional equative, Brasoveanu assumes that you retrieve a standard value for the adjective (or other degree predicate) in question and subtract that from the subject's value for the predicate. In conditional CCs without \textit{than} phrases (for the attorney example), you subtract the degree of sliminess of an attorney from the degree of sliminess for any other attorney who is more slimy than he is (in the antecedent). This is, at least in terms of its result, the same as what Beck does; it gives
the sense of using a single degree predicate twice with different variables each time in order to derive a set of pairs that differ by some degree (where, for Brasoveanu, this degree of difference is also recorded for each pair).

Despite the differences we've seen between the different categories of CCs, the reason that Brasoveanu gives for wanting to give them a unified analysis is that the same morphology (cît 'how much', atît 'that much') is used in both constructions in Romanian. He says they also have the same syntax except that the conditional CCs don't always have a than phrase. He adopts Beck's syntax for the English and German CC in his analysis of Romanian, which was designed to extend to correlatives in von Fintel 1994.

Moving now to the actual formal theory, Brasoveanu gives an analysis in a dynamic semantic framework. In a static semantics, the order of clauses is not important; each just contributes a proposition to the common ground. In a dynamic semantics, clause order makes a difference; here, because the meaning of a sentence changes an information state (like a Common Ground, if we simplify) to include new entities in the discourse, etc. that can be referred to by the next sentence. In the case of the CC, the degree differentials and salient entities (like those slimy lawyers) from the first clause need to be available in the interpretation of the second clause. There are many technical details that I would like to try to avoid introducing here as they do not come up in the discussion later, so the following sketches the most important points.

A Discourse Representation Structure (DRS) is a sentence-level meaning in this dynamic system. It takes an information state (like a Common Ground that also tracks the assignment of values to variables, etc.) and returns another information state that only differs from the input state in that it may include a new Discourse Referent (DR) or a
new condition. Conditions are like a predicate would be in a static semantics: a constraint on the entities in question. A DR is similar to what a noun phrase would be in a static semantics (further description below). A DRS is judged to be 'true' if such an output information state exists. In order to combine DRSs, you use Dynamic conjunction, which takes two info states as argument and asserts that there is a third info state that is both the output of the first DRS and the input to the second DRS. The process of altering an input state to yield an output state is called an update.

To understand what Discourse Referents are for Brasoveanu, we must first understand stacks. Brasoveanu introduces stacks into the natural language ontology, which he says are basically sequences of individuals. Empty positions in the stack all have a dummy variable $e$, which is the universal falsifier (makes any sentence false). There is a $\text{lng}$ function for the length of a stack, determined as $n+1$ where $n$ is the number of positions to the right of (or below, depending on the way the stack is depicted) an empty element. You add 1 to $n$ because the first element is in the 0th position. The update process never gets rid of anything already in the stack but rather just adds a value to the first empty position. Brasoveanu then adds stacks for other kinds of variables besides individuals, such as differentials (the dynamic equivalents of degree intervals). Then he introduces discourse referents (DRs) for things of these various types, each a projection function over a kind of stack (so individual DRs are functions over the individual stack, etc.). DRs introduce variables into the next available empty position in a stack. Though I have represented this as multiple stacks, Brasoveanu actually sets this up as one stack with sub-stacks for different types of entities, but nothing in our discussion here hinges on this difference. Now that we understand more about DRs, we
can understand conditions better. They take DRs of the kind appropriate to what their arguments would be in a static system and return the set of info states where all values of the DRs in question verify the static version of that predicate. In other words, if the static predicate $\text{walk}'$ takes an individual argument $\text{john}'$, then the condition $\text{WALK}'$ takes a DR with individual values, $\text{JOHN}'$ and restricts the info state such that only the individuals in that DR that are true of the static predicate $\text{walk}'$ remain consistent with the Common Ground (roughly).

Next, Brasoveanu defines $\textbf{max}$ as an operator on DRs. It takes a DRS and returns a DRS, namely the DRS that takes an information state and returns one that is the maximal set that stores all the stacks that meet the conditions in that DRS. He then encodes the $\textbf{max}$ operator in the meaning of the $1$, guaranteeing universal quantification over degree differences. These degree differences will necessarily all be positive because they are sets of natural numbers according to Brasoveanu's definition. They will also be necessarily non-zero because of the contribution of the comparative morpheme, which includes this restriction. Based on the definition of a differential as an interval, the representation for someone's height ends up being a set of degrees such that an individual is at least as tall as that degree. The difference between two heights (sets of degrees) is just another set of degrees: the set difference between the other sets. Then he defines $\text{tall}$ as a function from dynamic intervals to a function from dynamic individuals to dynamic truth values (updates, DRS's). As noted above, he assumes Beck's syntax, including the idea that $\textit{more}$ and the $\textit{than}$ phrase form a constituent. He also assumes that each comparative morpheme involves two instances of quantifying-in (corresponding to $\lambda$-abstraction), per Schwarzschild 2008. This includes an instance for each (dynamic)
degree-predicate argument. Finally, these are the target meanings for the comparative version of the non-conditional (141) and conditional (143) above (these are the only examples that have corresponding dynamic derivations).

\[ \exists D_{\gamma}(D \neq \emptyset \land D = \{d_\gamma: tall(x, d)\} \setminus \{d'_\gamma: tall(x', d')\} \land \exists D'_{\gamma}(D' \neq \emptyset \land D' = \{d_\gamma: tall(y, d)\} \setminus \{d'_\gamma: tall(y', d')\} \land D = D') \]

\[ \forall x \forall y \exists D(lw(x) \land lw(y) \land D \neq \emptyset \land D = \{d: aggr(x, d)\} \setminus \{d': aggr(y, d')\} \rightarrow \exists D'(D' \neq \emptyset \land D' = \{d: eff(x, d)\} \setminus \{d': eff(y, d')\}) \]

Here, \(D\) and \(D'\) are intervals (sets of degrees), \(d\) and \(d'\) are individual degrees, and \(\setminus\) is set difference. We see the main difference is in terms of existential v. universal quantification over differential intervals in the first clause along with a corresponding difference in conjunction v. material implication. Brasoveanu calls this the 'variable conditionality' problem for his approach. We also see that (149), because it has \(\text{tot}\), has the equality relation specified. In (150), it does not look like any relation has been specified, but that is because none is needed – this achieves the effect of monotonic correlation already. That is because for any positive non-zero degree difference, there exists another positive non-zero degree difference for the same individuals (explaining how this happens is presumably the reason the dynamic analysis was developed).

Brasoveanu says “the correspondence between the differential intervals \(D\) and \(D'\) is implicitly contributed by the relative scope of the two quantifiers \(\forall D \ [D \neq \emptyset \rightarrow \exists D' \ [D' \neq \emptyset \ldots]]\): any non-empty aggressiveness interval \(D\) induced by a pair of lawyers \(x\) and \(y\) is (somehow) related to a non-empty efficiency interval \(D'\) induced by the same pair of lawyers” (2009:3).
Brasoveanu says that the morpheme meaning *how much* in Romanian is a degree proform that is anaphoric to degree differentials. This analysis of *how much* unifies it with other proforms used in correlative clauses (equating individuals, etc. using other *wh*-words). He says that *how much* is an indefinite that introduces a non-empty interval, which in dynamic terms makes it an indefinite that introduces a differential-type DR.

It has been shown for both *wh* constructions in English and for correlative constructions cross-linguistically that universality/maximality is involved in the following way – when one thing is described by the relative clause, it must be unique, and when there is more than one, it must be maximal/universal (Srivastav 1991, inter alia). Thus, Brasoveanu claims that the conditional comparative correlatives are like the plural cases, which is why he choose to assign the $\textbf{max}$ operator to the meaning of that lexical item. In terms of getting existential quantification for non-conditional correlatives and universal quantification for conditional correlatives, he assumes that when you have a singleton, $\textbf{max}$ knows to has it be existentially bound, and when you have two or more, it knows to have the universal sense.

The Romanian equivalent of *that much*, also *the2*, contains a free variable over relations in order to pick up from context (e.g. the presupposition of *tot*) which relation is salient. In the case of monotonic correlation, which we saw is already present in the sentence, he assumes that this relational variable gets contextually filled by the “always available” cartesian product on differentials (so $R(D,D')$ becomes $D\times D'$). This makes no truth-conditional difference and so Brasoveanu ignores it throughout the paper. *The2* also retrieves the interval that *how much* created and relates it via its relation $R$ to a new interval that it introduces.
Now, to tackle the variable conditionality problem, Brasoveanu adds something new to the ontology by introducing plural information states, which are sets of stacks. In addition to the values in stacks, Brasoveanu assumes that we can make reference to stack structures as well, and specifically, relations between values in the various stacks. This is useful because it allows us to say, for example, that the 20th value in the differential stack is a differential that represents a degree of something for the individual in the 20th value in the individual stack. In the CC, then, for a stack in which all of the entries meet the condition of being slimy attorneys, there will be associated stacks representing their sliminess differential and their efficiency differential. The rows in the plural information state will represent one of these individuals and their two values, etc. Then, referring to the structure of that row across clauses gives you access to the values for each stack and guarantees that you are retrieving values for the same individual in both clauses. This, in combination with the different contribution of max according to whether the set of differentials is a singleton or not, accounts for the variable conditionality.

Brasoveanu calls the second problem the compositionality problem for CCs without than phrases. In these sentences, as in Beck's analysis, you want some of the content of each clause to be used twice and other parts to be used once. In most understandings of what it means to be compositional, each item must be used once and only once in the composition process, so this is puzzling. Brasoveanu's way of handling this issue is as follows. He introduces an operator over stacks that removes the first item in the stack (this is presumably the one added most recently in order for this to make sense). He marks this operator as a superscripted e-1. He then generalizes this operator to DRSs as well, in which case it removes the first item in every stack in an information
structure. There is a difference in the way he describes the operation when it comes to DRSs, however. He doesn't use the word *remove*. He says that the first element in each stack will be treated as if it didn't exist, suggesting that it is still 'there' somehow. Then, he introduces another set of lexical items for *the1* and *the2* that are superscripted with this operator for use in the case of conditional comparative correlatives without *than* phrases. Essentially, this *e-1* operator lets the first update take place and then lets the second update in a clause occupy the same stack positions as the last update for the sake of comparison between them. Then it reattaches the one that it temporarily removed so that they can actually be compared. Brasoveanu compares this to what happens in modal subordination. He takes the meaning of the antecedent and consequent each to be used twice and not once to update the context, so two clauses result in four recorded context changes. With modal operators, too, the world variables they contribute are treated as if they were the actual world (displacing it temporarily) and then replaced on the stack.

In summary, Brasoveanu works hard to unify a set of heterogeneous data that occur with the same morphemes. He claims that the variables quantified over in any CC are degree differences rather than individuals, times or worlds. And he proposes a dynamic semantic analysis to account for the donkey anaphora and anaphora between differentials across the clauses of the CC.

### 3.4.2 Discussion

One of the things that is really nice about Brasoveanu 2009 is that, even though it requires a number of new tools, most are independently motivated. So, for example, the *e-1* operator could be used in an account of modal subordination; plural information
states and reference to their structure as well as their values does double-duty to account for donkey anaphora outside the CC; the max operator is already present in other wh-words in correlative, etc. A further advantage is that Brasoveanu relies on the overt forms in the sentence in deriving his sentence-level meanings. He strongly critiques Beck's approach related to this point. Brasoveanu accuses Beck's analysis of being noncompositional and claims that it is related to the fact that:

“most of the work in deriving the correct truth conditions is done not by the meanings of the lexical items and the syntax-based way they are put together, but by postulating covert pairs of variables and a variety of covert operators inserted in particular structural positions that are suitably coindexed with both the covert variables and the overtly present indefinites and pronouns” (2009:16).

By this, he is presumably referring to the covert pairs of variables that are introduced as part of the meaning postulate for the, the covert λ-abstraction that is type-driven, and the covert adverb of quantification following what is proposed in the literature for conditionals.

On the other hand, Brasoveanu does not even attempt to handle the interaction of CCs with AQs, which would be a challenge for him. A solution along the lines of the one I posited in 3.2.2.2 is impossible for his analysis for two reasons: universal quantificational force would not explain the presence of less stringent AQs like usually, and the times when sometimes is felicitous with CCs depended upon whether the AQ and the CC were quantifying over the same thing. Because, for Brasoveanu, quantification always takes place over differentials, there would always be a mismatch in quantification between the AQ and CC, leaving it mysterious that the AQ would appear sometimes and
not others. Also, though he criticizes Beck for using covert pairs of variables to account for what he does using his $e$-$l$ operator, his operator comes at the expense of positing lexical ambiguity for the, how much, etc. depending upon whether the operator is part of its meaning. This casts doubt on the clear superiority of one approach as compared to the other. The fact that Brasoveanu posits two versions of the, one for when there is a than phrase and one for when there is no than phrase, raises an interesting question. Why unify equative and comparative CCs but not those with and without than phrases? Especially given that the only justification given for a unified analysis is similar morphology and syntactic forms, which all of these share, one division seems just as arbitrary as another.

I was also surprised that Brasoveanu did not appeal to the denotation of morphemes that were different between certain subtypes of sentences, such as pe glossed as PE and de glossed as DE as loci of meaning differences between the different kinds of sentences. At the very least, the presence of PE and DE in one and not the other needs to be explained (e.g. are they in complementary distribution with the comparative morpheme(s)? if so, why? what do they do?).

In the same way that it is strange to have a lexical entry for cu cît applying to a sentence containing pe cît, another facet of this meaning that needs to be explained is why every entry for cu cît, the, etc. requires a comparative morpheme as its first argument. Half of the sentence types that Brasoveanu is accounting for are missing overt comparative morphemes, so it seems strange that he would level this criticism about having null morphemes against Beck and then do something similar in his analysis. He does not mention anything about this, so it is not clear that he intends for there to be a
null comparative, but either way, an explanation for what happens when no overt comparative morpheme is present is needed.

Though the use of independently-motivated pieces of analysis is one of the strengths of the theory (and may work for Romanian CCs), it also makes the theory rigid in a way that prevents easy reanalysis in the face of problematic data. So, for example, the fact that Romainia *how much* in CCs is unified with other correlative proforms is nice, but it leads to a number of false predictions. One is, of course, that it predicts universal force for the CC. Whereas Beck could revise her analysis easily to match my new data showing that the CC has a correlative force similar to *most*, Brasoveanu does not have that option since *wh*-correlatives always show maximality or uniqueness but never *most*-like force. Also, the relation between the clauses in all of the other *wh*-correlatives is one of co-reference, which is only true of one of the three types of CCs discussed here. Brasoveanu argues that non-conditional equative correlatives like (142) are also an instance of equating the values in both clauses, but the degrees in each clause are not even on the same scale; a degree of intelligence cannot technically be the same as a degree of beauty, surely.

A final problem for this analysis comes from a technical issue. Brasoveanu's way of taking differentials is in terms of set differences, not the *order* of set differences, which is what he would need in order to make 10-9 and 2-1 the same. As it stands, the difference in 10-9 is 10 because that's the only degree that's not in both. And the difference between 2-1 is 2 for the same reason. Thus, a sentence like (141) would come out false if the son were 6', the daughter 5'9", the father 5'10" and the mother 5'7" even though the difference between the son and daughter's height is the same as the difference
between the father and mother's height. Though you could change what is quantified over from differentials to orders of differentials, you would need a new type in the ontology, a new kind of stack, etc.

3.5 Conclusion

Before I conclude this chapter, there is one more set of data that I particularly want to lay out that relates to definites and indefinites in CCs. When C1 and/or C2 contain verbs (i.e. the gapped constituent that the more... is fronted from), definites and indefinites seem to behave basically like they do elsewhere in English. In example (151), the definite noun phrase in (151a) refers to a unique salient attorney, as we would expect on the basis of the meaning of the. Similarly, the indefinite NP in (151b), not a specific indefinite, makes a statement about any attorney.

(151) a. The better dressed the attorney is, the more efficient he'll be.  
    b. The better dressed an attorney is, the more efficient he'll be.

I have noticed, however, that when a verb (as simple as the copula!) is missing, both receive a non-specific interpretation:

(152) a. The better dressed the attorney, the more efficient he'll be.  
    b. The better dressed an attorney, the more efficient he'll be.

(153) a. The smarter the dog, the easier it is to train.  
    b. The smarter a dog, the easier it is to train.  
    c. #The smarter dogs, the easier they are to train.

Even more bizarre, we can see in (153c) that it is not the case that any old generic can fill this spot since a bare plural makes the same sentence infelicitous. I have no account of
these phenomena at this time and reserve them for future research.

In this chapter, we have seen early descriptions of the semantics of the CC that brought to light many similarities between CCs and conditionals. We saw that Beck 1997 took this idea to its logical conclusion, giving an account that was especially insightful in terms of how a regular comparative morpheme could be used to derive the monotonic correlation in the CCs truth conditions. We saw subsequent recent analyses due to Lin 2007 and Brasoveanu 2009, each of which suggests new patterns and parallels with other phenomena. In addition to these analyses, I presented several new data sets and arguments that the quantificational force in the CC is proportional and inherent to the lexical entry for the *the* in the subordinate clause. The data included evidence from entailment patterns with NPIs and Google counts of various AQs as they occur with CCs and conditionals. I have shown that the similarities between CCs and conditionals can be attributed to the fact that both share a more basic connection in the form of Generalized Quantification, and I have further shown that there are more differences between them than was previously believed. In short, then, we have seen that past approaches to CCs cannot account for the non-universality of CCs in a way that explains their difference from conditionals and that proposing an analysis in which *the* contains an operator with correlational force effectively accounts for the full range of semantic benchmarks.

### 3.6 Benchmarks for the CC

(B30) *Future tense:* CCs and RCCs do not occur with future *will* in their antecedents.
(B31) *NPIs in the CC:* Negative Polarity Items can occur in the antecedent but not consequent clause of the CC and RCC.
(B32) *Quantificational Variability:* CCs and RCCs have meanings that can correlate degrees with respect to a number of different natural language types, including worlds, times and individuals.
(B33) *Donkey anaphora*: The CC and RCC license donkey anaphora between their clauses.

(B34) *Overt AQs*: The CC can appear with an overt Adverb of Quantification.

(B35) *Narrow scope negation in CCs*: CC clauses can contain narrow scope negation where it is felicitous with respect to the comparative.

(B36) *Quantificational Force in the CC*: The (unmodified) CC's quantificational force is neither universal nor generic but rather patterns with *most* in terms of native speaker judgements of truth in proportional contexts.

(B37) *Indefeasible Force*: The correlational force of the CC cannot be altered by an AQ.
In this chapter, I introduce the framework that I will use to analyze the CC and then give its formal analysis. The framework is called Pheno/Tecto Distinguished Categorial Grammar (PTDCG). I will begin by discussing aspects of syntax, semantics and the syntax-semantics interface in PTDCG informally and situating it with respect to other frameworks in 4.1. I will then give the specifics of the framework in these same divisions (syntax, semantics, syntax-semantics interface) as well as examples of how simple sentences are analyzed in PTDCG in 4.2. Finally, I will analyze the CC and RCC within PTDCG, though I will wait until the next chapter to compare these analyses to previous theories and the benchmarks laid out in the last two chapters. This section, 4.3, will focus on how the theory actually works before turning to the predictions it makes.

4.1 Situating PTDCG with respect to other frameworks

In this section, I begin by describing characteristics of theories of natural language syntax, and I divide the syntactic component of PTDCG into two separate parts: the tectogrammar and the phenogrammar. This is the sense in which PTDCG is a kind of Categorial Grammar in which the pheno- and tectogrammatical components are distinguished, as per its name. I choose the simplest syntax-semantics interface for PTDCG, which it shares with other categorial grammars; in this case, it is a simple function between tectogrammar and semantic form. This semantic form is static, as
compared to dynamic (defined in the last chapter). However, insofar as a compositional, truth-conditional dynamic semantics will need to work in tandem with a static system, PTDCG is a crucial component of a dynamic framework, Dynamic Categorial Grammar, which is under development but will not be discussed at any length in this thesis.

Natural language grammars have at least three parts in order to treat three empirical domains: what we hear (the sounds of language), what we take these sounds to mean (interpretation), and how we can put sound strings together (sentence structure, word order). Syntax is responsible for answering the last of these questions. It is a part of natural language that mediates between sound strings and their interpretations. Most theories assume that the root of syntax is constituent structure, where constituency is determined on the basis of tests such as substitution, ellipsis, topicalization, clefting, coordination, etc. Once constituency is determined, a set of grammatical categories is defined (noun phrase, preposition, verb, etc.) in order for any constituent to have a grammatical category and any non-constituent not to. Then, grammatical rules describing what kinds of categories can combine are written in order to create bigger constituents from smaller ones (or from individual words). For example, the fact that (1) but not (2) is grammatical shows, according to the ellipsis test (in which did can be replaced for a constituent), that wrote a speech is a constituent but wrote a is not.

Assuming that a is a determiner (Det), speech is a common noun (N), and wrote is a verb (V), we determine the category of wrote a speech, i.e. verb phrase (VP) and do not give a category to wrote a. Then, (3) but not (4) would be a possible syntactic rule on the basis of the constituency revealed by (1) v. (2).
(1) The candidate wrote a speech, and the campaign writer did too.

(2) *The candidate wrote a speech, and the campaign writer did speech too.

(3) V Det N → VP

(4) V Det → ?

So far, these grammatical rules say something about what kinds of categories are constituents, but the order in which they are pronounced in the sentence, i.e. word order or constituent/phrase order, is another question. If we have a rule for forming sentences in English that says that you will always have a noun phrase followed by a verb phrase, in order to account for the grammaticality of *He walks* but not *Walks he*, then you might have trouble accounting for English questions in which the verb can come first (e.g. *Did he?*). In other words, we could interpret rules like (3) as saying that not only do those categories form a constituent, but they must appear in that same order. However, as the example of question formation shows, that kind of approach would run into problems. But then, if we do not assume that (3) indicates the order in which the strings associated with these categories are pronounced, there must be something else that does the job of telling us what the word order is for a given sentence (or a possible sentence of a language). As we will see below, some syntactic frameworks try to maintain that grammatical rules like (3) do determine word order but that other processes can change that order, while others try to define separate rules for word order and constituent structure, among other strategies.

So far, we have seen how one might figure out the orders in which sound strings can appear (as well as a more abstract concept of how they are grouped). The other part
of the puzzle, then, is how this relates to interpretation. For that, a syntax-semantics interface must be defined, which can take a number of forms. One example might be that every syntactic combination rule like (3) could have a corresponding semantic rule such as (5) that tells you how to interpret a structure of that kind. A rule like this assumes that semantics is compositional just like syntax is; that words have meanings and that they are built into larger meanings.

(5) If the meaning of V is A, the meaning of Det is B, and the meaning of N is C, then

\[ A \circ B \circ C \rightarrow D, \text{ where } D \text{ is “}B+C\text{ is the person, place or thing that }A\text{ is done to”}\]

Just as with (3), this is in no way sophisticated enough to represent any current theory of natural language, but it gets at some of the issues that any theory must deal with. One that we see in (5) is that we have to understand what it means for \(B\) to get added to \(C\). We need to understand the kind of system that underlies the different components of the grammar, and we need to be able to apply this system to much more complicated patterns in natural language. Usually, following Montague 1973, inter alia, the system of interpretation is taken to be a lambda calculus (whether it is the Intensional Logic of Montague or Church's 1940 simply-typed lambda calculus). This means, in part, that some meanings are functions that take other meanings as arguments to return another meaning. Returning to (5), then, if \(A\) is a function that takes person, place or thing arguments and \(B\) is a function that maps a kind of thing to an actual person place or thing, then the rule could be rewritten so that \(D = A(B(C))\). The point here is that in addition to the constituent structure and linear order that were discussed above, the
syntax also includes information about argument structure that relates to its interface with semantics.

Different syntactic frameworks take different approaches to modeling constituency, argument structure and word order. In order to tease apart the role played by various parts of each syntactic framework, I will appeal in my discussion to a distinction made by Curry 1961, which he called *tectogrammatics* and *phenogrammatics*. As the Greek origins of these words suggest, tectogrammatics is the part of the grammar that models how expressions are built, while phenogrammatics is the part that models how the expressions appear, i.e. the order in which they appear in spoken/written/signed sentences. I will follow Dowty 1982a and subsequent literature in calling these two components the tectogrammar and the phenogrammar. In many contemporary frameworks, these two aspects of the grammar are at least partially conflated, but they will be distinguished in PTDCG. We now look at two examples of how syntactic frameworks divide the labor of accounting for tecto- and phenogrammatical components before turning to the specifics of the setup of PTDCG.

In Transformational Grammar (TG), beginning with Chomsky (1955), there are a basic set of syntactic categories, a set of base rules, and a set of transformational rules. The syntactic categories and grammatical rules work together to model constituency in a way that is similar to the discussion of (1)-(4), above. TG assumes that the grammatical rules indicate both tecto- and phenogrammatical processes, meaning that they tell you how to build syntactic strings *and* the order in which they are pronounced (the other part of the phenogrammar, namely how individual words are pronounced, is part of the phonological component of the grammar). The transformational rules account for
question formation and other cases that would seem to contradict the order inherent in the rules. These rules 'transform' the order of words or morphemes in a sentence without altering the 'underlying' constituent structure. In this sense, then, grammatical rules have a tecto and pheno component, while transformational rules have only a pheno component (they leave the tectogrammatical structure intact). As TG-style theories developed, the difference between 'Deep Structure' and 'Surface Structure' became prominent, where the deep structure was the one that resulted from the base rules and the surface structure was the result after the transformational rules had applied. This fits somewhat well with Curry's own description of tectogrammar as what was underlying and phenogrammar as the concrete manifestation of the sentence (it's surface form).

In Categorial Grammar (CG), syntactic structure comes more from the grammatical categories than the grammatical rules because the categories are complex: they have argument structure and thus say what kinds of categories are arguments of a particular word, etc. So whereas *give* is a verb of category V in transformational grammar (among other frameworks), its category in CG would be \((\text{NP}\backslash S)/\text{PP}/\text{NP}\), which says that *give* is the kind of thing that produces a sentence when given two NP arguments and one PP argument. So instead of having a V that can become a sentence according to various grammatical rules, the \((\text{NP}\backslash S)/\text{PP}/\text{NP}\) category tells you how to combine *give* with other things to get a sentence. In the CG setup, then, there are only a few grammatical rules, such as function application, that are not specific to particular category names. So TG has fewer possible syntactic categories and many grammatical rules, while CG has many possible syntactic categories and few grammatical rules in order to try to accomplish the same goal. Given that in CG, a syntactic category is
specified for every lexical item and the fact that these categories do the work of grammatical rules in TG, this is the sense in which CG has been called a 'radically lexical' framework (Karttunen 1989).

The lexical specification of categories and few rules of combination is general to all forms of CG. One important line along which different CG theories can be divided is whether there is a tecto/pheno distinction present in their syntax, which is the feature I will focus on here. In the most established forms of CG – Combinatory Categorial Grammar (CCG) (Ades and Steedman 1982, Steedman 1985, inter alia) and Type-Logical Grammar (TLG) (Morrill 1994, Moortgat 1997, inter alia) – tecto and phenogrammatical components are conflated in the specification of syntactic categories (again, the other part of the phenogrammar, the phonological form of individual words, is a separate component). Specifically, if we look again at the category for give in one of these frameworks, ((NP\S)/PP)/NP, we see that the slashes separating the various categories are directional. A left-leaning slash (A\B) means both that the argument this category is looking for is a category-A argument (tectogrammar) and that this category-A argument will appear to the left of the category-A\B functor (phenogrammar).31 A right-leaning slash (B/A) has the same tectogrammatical component as the left-leaning slash but a different phenogrammatical interpretation, namely that the category-A argument will appear to the right of the functor.

More recently, categorial grammars that separate phenogrammatical and tectogrammatical components have been proposed. Beginning with Oehrle 1994, not only were semantic forms terms of the lambda calculus, but phonological forms were as well. The exact interpretation of slashes actually differs between CCG and TLG as well as between practitioners of these frameworks, but the differences aren't such that it affects the general point that I am making that slashes with directionality conflate tecto- and phenogrammatical components.
well. This means that in addition to the particular phonological strings that had always been a part of categorial grammars, it was now possible for the phonological form to be a function from strings to strings, etc. where the phonological form of larger phrases is built from other phonological forms. Following Oehrle, de Groote's 2001 Abstract Categorial Grammar (ACG), Muskens' 2003, 2007 Lambda Grammar, and Pollard's 2008 Convergent Grammar share this property. My primary reason for using a very similar but not identical framework in this thesis is as follows. As we have seen in past chapters, theories of donkey anaphora and comparison are relevant to the analysis of CCs in that CCs incorporate both elements. Recently, as suggested in Barker (2007), some occurrences of comparison are instances of parasitic scope (discussed further below), which is difficult to account for in most grammatical theories (he proposes a grammar that includes continuations to get around this). Similarly, dynamic theories of meaning are needed to account for donkey anaphora, among other phenomena. One reason to prefer PTDCG to its predecessors, then, is that it can account for parasitic scope (as we will see), and it can form the basis of a dynamic theory that will be able to account for donkey anaphora.

In conclusion, we have seen in this section that syntactic frameworks can vary in how many components of syntax they make reference to, how syntactic categories are defined, how grammatical rules are determined, etc. PTDCG's syntax, in keeping with the tradition of CG, will involve few rules, lexically-specified syntactic categories that contain subcategorization information, a rigorous basis in logic and a simple syntax-semantics interface in the form of a function between the tectogrammar and the interpretation, as we will see in the next section. Unlike some kinds of CG (CCG, TLG)
and like others (ACG, Convergent Grammar, Lambda Grammar), PTDCG divides the syntactic component of grammar in two, separating the tectogrammar and the phenogrammar. This move enables simpler analyses of complex phenomena such as parasitic scope, as we will see in future sections.

4.2 Pheno/Tecto Differentiated Categorial Grammar

Linguistic expressions such as lexical items are represented in PTDCG as ordered triples of the components discussed above: pheno, tecto, and semantics. These linguistic expressions are axioms in the proof theory of PTDCG. Each component is underpinned by a particular logic, consistent with other kinds of Categorial Grammar. Simply-typed lambda calculi underlie the phenogrammatic and semantic components, and linear logic underlies the tectogrammatic component. I will say more about what this means below. For now, I want to focus on the implications of treating grammar as a proof system. In this way of doing things, the proofs are derivations of sentences. The would-be sentence is a triple, and the question is whether that triple is deducible from other triples (for our purposes, lexical items, including hypothetical ones) via the rules of deduction for that system (for our purposes, there will be two inference rules, defined below). This means that in derivations of sentences in PTDCG, each stage is a step in a proof, one that does not involve a change in the actual objects of representation used in that proof. This is typical of categorial grammars, but it differs from other derivational theories of syntax. In other theories, a series of steps applies to objects that have been altered, reflecting the history of changes (often via phrase structural rules). In this sense, there is an actual change of representation between steps that is unlike what we see in PTDCG.
Returning now to our three components, each has its own type system. In addition, the phenogrammar and semantics also have terms. The types are what is used by a proof theory to determine whether an expression is well-formed (provable). The terms are used by another algebra, such as a model theory, to indicate, e.g. whether a well-formed proof (sentence) is true or false in that model. Thus, we need to define the types for each component to serve as formulas for our proof theory. For each component below, I provide the basic types (formulas) as well as a rule of recursion that lets us build additional formulas using the connectives of that logic. For the phenogrammatic component, we have only a single basic type, str. Then the rule in (6b) gives the possibility of generating types such as \( \text{str} \rightarrow \text{str} \):

(6) **Phenogrammar:**
   a. Basic Types: str
   b. Rules: Any basic formula is a formula, and if A and B are formulas, so is \( A \rightarrow B \).

(7) **Tectogrammar:**
   a. Basic Types: N, NP, Deg, PP(T), PP(W), S
   b. Rules: Any basic formula is a formula, and if A and B are formulas, so is \( A \rightsquigarrow B \).

(8) **Semantics:**
   a. Basic Types: e, d, i, s, t
   b. Rules: Any basic formula is a formula, and if A and B are formulas, so is \( A \rightarrow B \).\(^{32}\)

In (7), we see that the basic types for the tectogrammar include the noun, noun phrase, degree phrase and sentence. I distinguish two categories of prepositional phrase (PP) – those that are temporal phrases and those that are 'world' or modal phrases. An example of the former would be \( \text{at 10 pm} \) and an example of the latter would be \( \text{in a world where racism did not exist} \). I return to this in 4.3. The connective in (7b) is called linear

\(^{32}\) Even though a function from individuals to truth values should be written \( e \rightarrow t \) according to this rule, I will write any \( A \rightarrow B \) as \( <A,B> \) for this component in keeping with the tradition in Montague Semantics.
implication, also known informally as lollipop. It is a more restricted form of implication that is used in linear logic (as compared to implication used in other logics). Between (7a) and (7b), we can derive types such as Deg→NP→S, which would be the type of a verb that takes two arguments: first, a degree phrase, and second, a noun phrase. In (8), we see that the basic types for the semantics are individuals ($e$), degrees ($d$), times ($i$), possible worlds ($s$), and truth values ($t$). As with the other components, (8b) uses the basic types in (8a) to create more complex types, such as <$e,<d,<s,t>>$ or $<<e,t>,<<e,t>,t>>$. I will abbreviate these terms to $<e,<d,st>>$ and $<et,<et,t>>$, removing the innermost sets of angled brackets and separating comma when they contain only a single implication throughout the rest of this section.

Now, we have seen these three type systems, but how are they combined into a single proof system for PTDCG? This is based on the fact that a function can be defined from tecto types to semantic types and another from tecto types to pheno types. This means that given a tecto type, both the pheno type and the semantic type are predictable. That this is possible should be clear from the fact that (6)-(8) have very similar rules of recursion. So, for example, we can define one function $sem$ and another $phon$ that takes a tecto type and returns a semantic or pheno type. We then define the correspondences for the basic types and for the rules, as in (9).\textsuperscript{33}

\textsuperscript{33} Technically, $sem(S)$ could also be $st$ or $t$, but this makes $sem$ not be a function; one way that the literature (cf. Dowty 1982b) has handled this is to distinguish syntactic categories $S$ v. $S'$, etc. where each has a different one of the possible output types. My analysis is not influenced by this choice.
This definition guarantees that anything that is tectogrammatically well-formed will also be phenogrammatically and semantically well-typed. Thus, using the tectogrammatical form as the glue, rather than defining inference rules separately for each component, we can give a single rule of each kind that spans all three grammatical components.

As mentioned earlier, the phenogrammar and semantics have terms as well as types. The basic forms for each are too numerous to list, but the phenogrammar terms denote the pronunciations themselves (the sound strings that are in fact of type str), and the semantic terms denote the meanings themselves (or, the representations of those meanings, depending on your view). In other words, the lexical entry for a word like dog would include the following: str (pheno type), [dog] (pheno term), N (tecto type), et (semantic type), and dog' (semantic term). Throughout this thesis, x’ will represent the meaning of x (but x will not always be represented by x’, as in the case of variables and constants). There is no difference between dog' and λx.dog'(x); the latter just has its argument(s) spelled out. When the latter version is used, the semantic type information becomes superfluous. Thus, in order to simplify the way that rules, proofs, etc. are stated, I will omit the pheno and semantic types in the diagrams of this thesis, but try to indicate in the terms what the types would have conveyed.
We are nearly ready to state the two inference rules that will be used in the creation of our proofs, but a few more details about the particular proof system being used are in order here. It is a theory of natural deduction largely in the Gentzen sequent style, though I will depart from this style in certain ways. Sequents are pairs of contexts (sets of formulas) and statements (formulas). Let's let uppercase Greek letters $\Gamma$ and $\Delta$ be metavariables ranging over contexts and uppercase Roman letters A, B, C be metavariables ranging over statements. The way in which a statement is notated as being provable (or deducible or derivable) from its context is by placing a turnstile ($\vdash$) between them. Thus, $\Gamma \vdash A$ means that $A$ is provable in (or given) the context $\Gamma$. In the case of most of our lexical items (axioms), they are provable in the null context. If no context is required, we write simply $\vdash A$. Also, when there are multiple elements in the context or multiple formulas that are part of the statement, I separate them with commas, just as in typical lists of sets.

We are now ready to introduce the rules that will determine when one proof is is provable on the basis of one or more existing proofs. There are only two rules, which correspond to the well-known proof theoretic rules of Modus Ponens and Hypothetical Proof, given in their simplest form in (10) and (11). Here, the letters represent propositions in the simplest possible statement logic. However, given the added complexity of our system (multiple components and higher order logics), the axioms and rules for our system are slightly more complicated, so it is best to use these simpler forms to get the general idea.

(10) Modus Ponens: 

$$
\begin{array}{c}
B \rightarrow A \\
B \\
\hline
A \\
\end{array}
\rightarrow E
$$
(11) Hypothetical Proof:

\[
\begin{array}{c}
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots \\
B \\
\hline
A \rightarrow B
\end{array}
\rightarrow I^n
\]

The schema in (10) and (11) are written as proof trees where the conclusion of the inference rule appears at the bottom and the premises used to derive the result are on top. This is the opposite of the tradition in generative grammar, and this is one respect in which I will depart from the tradition of sequent style natural deduction in the derivations below, by inverting the order in which the steps are presented and ending with the resulting proof at the top of the tree (which nevertheless remains only a representation and not an object of the ontology itself, as noted earlier). Here, the horizontal line that separates the root (mother) node and its daughters is notated with the name of the rule schema that was instantiated. In other words, modus ponens is another name for implication elimination (\(\rightarrow E\)), and hypothetical proof is another name for implication introduction (\(\rightarrow I\)). Having a pair of rules that introduce and eliminate the same connective (here, implication) is common to natural deduction proof systems, and it is another similarity that the PTDCG proof system shares with others. It is easy to see why these rules are called \textit{introduction} and \textit{elimination} in that in each case, either the daughters or the mother but not both contain the connective, so in one case, it disappears, and in the other, it appears after having not been present before. In the rule for modus ponens, you have a function and its argument, and the result is what that function yields upon taking said argument. The rule for hypothetical proof is a bit trickier. In it, some
formulas among others (indicated by the dots) is (hypothetically) assumed within the larger proof whose conclusion is $B$. Then the inference rule states that it is permissible in this case to conclude that the other premises are enough to prove $A \rightarrow B$ with $A$ removed. Intuitively, this makes sense because you can always add $A$ back in to yield $B$ by using the rule of modus ponens. This rule additionally contains brackets to indicate that $A$ is hypothetical and matching superscripts on the hypothesized premise and $\rightarrow I$ rule so that in a case where multiple premises have been hypothesized, you can be sure of which have been discharged (in our system, all hypothesized premises must be discharged at some point).

The inference rules in PTDCG take triples of phenogrammar terms ($\phi$), tectogrammar types ($\tau$) and semantic terms ($\sigma$), separated by colons, i.e. $\phi : \tau : \sigma$. Each rule is titled to show the correspondence between modus ponens and hypothetical proof as we have seen them above and the rules of Merge and Move in mainstream generative grammar (Chomsky 1995). Also, since it is the tecto terms that are the most important for PTDCG's proof theory, we see that implication has been replaced with linear implication (the lollipop) in terms of the connective that is being eliminated or introduced.

(12) Modus Ponens, AKA Merge, AKA $\exists$ Elimination ($\exists E$):

\[
\Gamma, \Delta \vdash f(a) : B : g(e)
\]

\[
\Gamma \vdash f : A \exists B : g \\
\Delta \vdash a : A : e
\]
(13) Hypothetical Proof, AKA Move, AKA \rightarrow Introduction (\rightarrow I):

\[ \Gamma \vdash \lambda p.f : A \rightarrow B : \lambda x.g \]

\[ \Gamma, p : A : x \vdash f : B : g \]

In (12) and (13), \( f \) and \( a \) are pheno terms, \( p \) is a pheno variable, \( A \) and \( B \) are tecto types, \( g \) and \( e \) are semantic terms, and \( x \) is a semantic variable. In (12), all of the components to the right of the turnstile in the left-hand daughter are implicative formuli, and all of the components to the right of the turnstile in the right-hand daughter are potential arguments of those formuli. Thus, the mother contains the result of each components' function application. In (13), I have misaligned the mother and daughter proofs to show the difference in placement of the elements with respect to the turnstile. In essence, there is a hypothetical triple in the context (in addition to whatever triples are in the context \( \Gamma' \)) of the daughter that does not appear in the context of the mother. Instead, those hypothetical components have been abstracted out of the components of the statement to the right of the turnstile, creating a functional argument slot in each of the three components.

In applying this system to account for natural language data, we will have two kinds of axioms/triples: the lexical entries already discussed above and a set of hypothetical triples analogous to traces in other syntactic frameworks. From this point on, rather than writing sound strings as bracketed terms of the International Phonetic Alphabet, I will simply write the word in capital letters, so the sound string for \textit{dog} is \texttt{DOG}. I will further use the '+' for string concatenation, so that a phrase like \textit{a dog} is represented as \texttt{A+DOG}. Now, (14) and (15) represent each kind of triple:
(14) ⊢ FIDO : NP : f (f is a constant of type e that rigidly designates Fido)

(15) p : NP : x ⊢ p : NP : x

The entry in (15) is a product of an axiom schema A ⊢ A for hypothetical traces whereby you can prove A on the basis of having A in the context. These are the hypotheses that will eventually be discharged by lollipop introduction. As we will see in the next section, these traces will help us analyze unbounded dependencies.

As an intermediate summary, this section has so far introduced the PTDCG framework in which the phenogrammar, which includes at least the phonological representations of words and their linear order, and the tectogrammar, which includes information about what expressions can combine and the result of that combination, are separate components of the grammar. We have seen that the phenogrammar parallels the semantics in that both are lambda calculi and the relationship between the semantic calculus and the model theory for interpretation is analogous to that between the phenogrammatical calculus and the algebra of phonological forms. We have defined the basic types for each component as well as inference rules and one axiom schema for triples of phenogrammar, tectogrammar and semantic components. What follows are a number of examples of how various kinds of sentences are treated in PTDCG, each of which is relevant to the analysis of the CC and RCC.

4.2.1 PTDCG Analysis of Simple Sentences

The goal of this subsection is to analyze the sentence in (16) in PTDCG. As we will see, a simple sentence such as (16) can be analyzed via the lexical items and
without any hypothetical elements.

(16) Jen loves Ben.

First, we need the lexical entries for *jen, loves*, and *ben* as in (17).

(17) a. ⊢ JEN : NP : j
    b. ⊢ BEN : NP : b
    c. ⊢ λp.λq.q+LOVES+p : NP→NP→S : λx.λy.love'(x)(y)

The proper names need no explanation; they are simple constants in terms of their phenogrammatic and semantic forms. The entry for *loves* is extensional – in the analysis of CCs, I will intensionalize verbs, giving them time and world arguments in the semantics (as indicated in the type given for the *sem* function in the last section), but for the simplest kind of example, that is not needed, so I have excluded it here. We see that *loves* takes two NP arguments according to the tecto type. The sound string corresponding to the first NP is placed after the sound string LOVES, and the second one, before it, in accordance with the fact that it takes its object first and its subject second (along with the fact that English is SVO). The semantics tells us that the two individuals that correspond to the NP arguments will be semantic arguments of the two-place function *love’*, which will return 'true' just in case the individual in the *y* position loves the individual in the *x* position. Next, we look at how the sentence is proved according to our inference rules in the proof tree in (18).
Here there are two steps: one that combines *loves* and *Ben*, and one that combines *loves Ben* and *Jen*. Each is an instance of Modus Ponens. That does not need to be labeled explicitly because anytime two nodes combine as one, it can only be due to Modus Ponens. Similarly, if there is a step in which one leaf goes to a node with no other branches, it must be an instance of Hypothetical Proof. The form of each node is the final form for that node, meaning that any beta reduction has been carried out. Thus, 
\[ \lambda p \lambda q.q+LOVES+p (\text{BEN}) = \lambda q.q+LOVES+\text{BEN}, \]
which is the phenogrammatical form at the *loves Ben* node. As predicted, *loves* takes its arguments (in each of the three components) one at a time and gives back the desired word order, syntactic type, and meaning for the target sentence. In the next sections, we will see additional complexities in the forms that sentences can take.

4.2.2 PTDCG Analysis of Quantifier Scope Possibilities

This section shows how sentences with multiple quantifiers, each of which takes scope, can be handled. Specifically, it shows how the variety of possible readings is derived in the same vein as Quantifier Raising, Quantifier Lowering, or Quantifying-In in other theories. The basic kind of sentence is as in (19).
(19) Everyone loves someone.

This sentence has two prominent readings: one when *someone* takes wider scope than *everyone* and the other when *everyone* takes wider scope than *someone*. Any theory needs to show how both readings can be derived. Let us start with the lexical items, which are the same regardless of reading (*loves* is not listed because it does not differ from the entry in the last subsection):

(20) a. *someone*: \( \lambda p.p(\text{SOMEONE}) : (\text{NP} \; \text{S}) \; \text{S} : \lambda P. \exists x. \text{person}'(x) \land P(x) \)

b. *everyone*: \( \lambda q.q(\text{EVERYONE}) : (\text{NP} \; \text{S}) \; \text{S} : \lambda Q. \forall y. \text{person}'(y) \to Q(y) \)

In both cases, because *someone* and *everyone* are both generalized quantifiers; they take a predicate (NP\(\to\)S) argument in the tectogrammar. This argument, being a formula itself, will have a phenogrammatical component that is a function from strings to strings. The phenogrammatical component of these GQs takes that function as argument and puts the string that represents the phonological form of the GQ in as the argument of that function (Oehrle 1994). This means that even though *everyone* is the function that takes a predicate as argument, it is able to put its own form into the place in that predicate's phenogrammatical structure where a simple NP would have gone. The semantics of these GQs takes a property and either existentially or universally quantifies over individuals that are both people (the contribution of *-body*) and possess the property in question. The connective \( \to \) \( v. \land \) is also dependent upon which quantifier it is.

Now, if these lexical items all combined in the order they occur in the surface string via \( \to E \) (which wouldn't be well-formed anyway), there would only be a single
scope possibility in (19); namely, the one in which *everyone* outscopes *someone*, since it would combine last/'most recently'. Instead, we posit hypothetical traces for each argument of the verb, then abstract over these variables in a particular order that differs depending on what reading we are deriving. This corresponds to the order in which the quantifiers will be 'lowered' or quantified into those arguments. We see both possibilities in the proofs in (21) and (22) below. I begin with the reading that would be unexpected on the basis of surface word order. As we can see from the sentence-level semantic term, this is the reading that is paraphrasable as 'there is some person such that every person loves that person.'
(21) Existential Wide-Scope

\[ \vdash \text{EVERYONE+LOVES+SOMEONE} : S : \exists x.\text{pers'}(x) \land \forall y.\text{pers'}(y) \rightarrow \text{love'}(x)(y) \]

\[ \vdash \lambda p. p(\text{SOMEONE}) : (\text{NP} \rightarrow S) \rightarrow S : \lambda p. \exists x.\text{pers'}(x) \land P(x) \]

\[ \vdash \lambda t. \text{EVERYONE+LOVES+t} : \text{NP} \rightarrow S : \lambda y'. \forall y.\text{pers'}(y) \rightarrow \text{love'}(y')(y) \]

\[ t : \text{NP} \vdash \text{EVERYONE+LOVES+t} : S : \forall y.\text{pers'}(y) \rightarrow \text{love'}(y')(y) \]

\[ \vdash \lambda q. q(\text{EVERYONE}) : (\text{NP} \rightarrow S) \rightarrow S : \lambda Q. \forall y.\text{pers'}(y) \rightarrow Q(y) \]

\[ t : \text{NP} \vdash \lambda s. s + \text{LOVES+t} : \text{NP} \rightarrow S : \lambda z. \text{love'}(y')(z) \]

\[ t : \text{NP} \vdash \lambda s. s + \text{LOVES+t} : \text{NP} \rightarrow S : \lambda s. \text{LOVES+t} : \text{NP} \rightarrow S : \lambda z. \text{love'}(y')(z) \]

Combining with a trace rather than a lexical item, and the same is true for its subject argument as well. Again, we recognize these as traces because they enter something into the context (before the turnstile) that will need to be discharged later. And in fact, the subject trace gets discharged immediately. This is the first instance of hypothetical proof in action, and we see that all of the variables and categories introduced by the trace are abstracted over in that step. Thus, for instance, love'(y')(z) becomes \( \lambda z. \text{love'}(y')(z) \). This creates a predicate that is the right type to be the argument of \textit{everyone} in the next step.
Because we abstracted over the subject trace, *everyone* will get put into the phenogrammatical structure in the position of the subject, as intended. The result of this step would be a sentence if only there weren't a hypothetical premise still in the context (this is indicated by the fact that what follows the turnstile is of tecto type $S$). Thus, the next step is discharging the object trace of *loves*. Again, this amounts to abstraction in each component. In the final step, *someone* takes the newly-minted NP-$\rightarrow$S/predicate as argument to return a sentence with no hypothetical premises, where *someone* has been 'lowered' into the position following LOVES in the phenogram.

(22) Universal Wide-Scope

\[
\begin{align*}
\vdash \text{EVERYONE+LOVES+SOMEONE} : S : \forall y.\text{pers'}(y) & \rightarrow \exists x.\text{pers'}(x) \land \text{love'}(x)(y) \\
& \vdash \lambda q. q(\text{EVERYONE}) : (\text{NP} \rightarrow S) \rightarrow S : \lambda Q. \forall y.\text{pers'}(y) \rightarrow Q(y) \\
& \vdash \lambda s. s+\text{LOVES+SOMEONE} : \text{NP} \rightarrow S : \lambda z. \exists x.\text{pers'}(x) \land \text{love'}(x)(z) \\
& \quad s : \text{NP} : z \vdash s+\text{LOVES+SOMEONE} : S : \exists x.\text{pers'}(x) \land \text{love'}(x)(z) \\
& \quad \vdash \lambda p. p(\text{SOMEONE}) : (\text{NP} \rightarrow S) \rightarrow S : \lambda P. \exists x.\text{pers'}(x) \land P(x) \\
& \quad s : \text{NP} : z \vdash \lambda t. s+\text{LOVES+t} : \text{NP} \rightarrow S : \lambda y. \text{love'}(y)(z) \\
& \quad \quad t : \text{NP} : y, s : \text{NP} : z \vdash s+\text{LOVES+t} : S : \text{love'}(y)(z) \\
& \quad \vdash \lambda q. q : \text{NP} : z \vdash \text{NP} : z \\
& \quad \quad t : \text{NP} : y \vdash \lambda q. q : \text{NP} \rightarrow S : \lambda a. \text{love'}(y)(b) \\
& \quad \quad \vdash \lambda p. \lambda q. q + \text{LOVES+p} : \text{NP} \rightarrow \text{NP} \rightarrow S : \lambda a. \lambda b. \text{love'}(a)(b) \\
& \quad \quad t : \text{NP} : y \vdash t : \text{NP} : y
\end{align*}
\]
In the derivation of the second reading, the first two steps are the same. *Loves* takes two traces as arguments, simultaneously adding two triples to its hypothetical context. The next step, however, differs from the earlier derivation in terms of which trace is removed from the context via hypothetical proof. This time, it is the object that is abstracted over first. Because the word order is the same in the two readings (it is a single sentence), *someone* must combine next given that it should appear phenogrammatically where the object trace is. In other words, the change in which trace is pulled out of the context first determines which quantifier will combine first. Thus, *someone* combines with the predicate formed on the basis of the first hypothetical proof step, leading $\exists$ to be the innermost quantifier, whereas $\forall$ was the innermost quantifier in the last derivation.

Here, after *someone* combines, we have another hypothetical proof step, this time drawing the subject trace out of the context. The result is the argument of *everyone* in the last step, leading to a reading in which the universal quantifier takes wider scope than the existential quantifier, predicting that the sentence's meaning on this reading is that each person has a person that they love, where the people they love are potentially different people.

### 4.2.3 PTDCG Analysis of Comparatives and Parasitic Scope

The last section has shown how differential scoping is handled in PTDCG. This section will treat a different kind of scope relationship, called parasitic scope (Barker 2007), in which a particular element creates two scope positions where only one previously existed. This is closely related to the notion of 'tucking in' in mainstream
generative grammar. Further explanation will be given below, but what is crucial in this section is that this kind of relationship exists in some comparative sentences. Thus, this section lays out an argument for what kind of comparative we find in the CC and how it should be treated in PTDCG that preserves many of the insights from the literature on both comparison and parasitic scope. In other words, an analysis of parasitic scope is necessary in order to analyze phrasal comparison (in order to preserve its parallels with other natural language phenomena that are analyzed in terms parasitic scope), and an analysis of phrasal comparison is necessary for a non-ad hoc analysis of the comparative in the CC. I begin with some background on comparatives.

Comparative sentences are identified on the basis of containing a comparative morpheme: *less, more,* or *-er.* To make things simpler, this thesis will only include examples with *more.* Within the general category, Kennedy 1999, inter alia, divides comparative sentences (with or without measure phrases such as *10 inches*) into three types: *predicative comparatives,* *attributive comparatives,* and *comparative nominals.* In the following examples, (23) is predicative, (24) is attributive and (25) is nominal.

(23) a. Alex is taller than 5' 6".
   b. Alex is taller than Colleen.
   c. Alex is taller than Elizabeth thought he was.
   d. The table is as long as it is wide.

(24) a. Randy has longer legs than Pam.
   b. Pooh found sweeter honey than Rabbit did.

(25) a. More gifts were given to Crista than Larry.
   b. There are more cookies than there are cakes.

All three kinds of comparatives appear in the CC:
(26) a. The taller Alex is, the better his chances are. (predicative)
   b. The longer someone's legs, the greater their stride. (attributive)
   c. The more cookies there are, the more milk there is. (nominal)

As in the literature, the predicative comparative will be my focus. Within the predicative
category, there is further subcategorization based on the kind of constituent that we find
as the complement of than or as, following Bresnan 1975 and subsequent work. Some
complements are clausal; these are the ones where the complement of than is a clause or
reduced clause, as in (27) or (28). Some complements are non-clausal phrases; these are
the ones where the complement of than is a non-clausal constituent, whether it is a
degree phrase as in (29) or a noun phrase as in (30). I will call this complement the than
phrase regardless of whether it is clausal, as I have throughout the last two chapters
(though it is often called the comparative clause in the literature, cf. Kennedy 1999).

   Clausal comparatives have two subtypes, called comparative deletion and
comparative subdeletion, exemplified by the following.

(27) Alex is taller than Elizabeth thought he was. (Comparative Deletion)
(28) The table is longer than it is wide. (Comparative Subdeletion)

In comparative subdeletion examples like (28), the gap is only a degree argument, where
a measure phrase would normally go before an adjective, as in it is 3 feet wide. In
comparative deletion examples such as (27), the complement is still clausal (as indicated
by the finite verb in the complement of than), but it is missing a gradable adjective in
addition to a measure phrase/degree adverbial (as in Elizabeth thought he was 5'7" tall).
Phrasal comparatives such as (30) and Simple Comparatives such as (29) have a
complement that is a simple phrase, but in (29), it is a degree phrase (Deg), and in (30), it is a noun phrase. In both of these cases, there is no finite verb to indicate clausality.

(29) Mo eats more than 2 hamburgers. (Simple Comparative)
(30) Mo eats more than Jo. (Phrasal Comparative)

Finally, I will make a further division within the category of phrasal comparatives between examples like the one in (30) and examples such as (31), which I will call associate-remnant comparatives, inspired by an idea in Pollard 2008. Though I assign examples like (31) to their own category, I consider them a subtype of phrasal comparatives insofar as they have only an NP following than and, according to both Pollard and myself, they should be able to be analyzed with the same lexical entry for more that is used in (30). I discuss their differences in just a moment.

(31) Bo owes Mo more than Jo. (Associate-Remnant Comparative)

Another similarity between (30) and (31) is that in each case, there are two phrases, optionally bearing focal pitch accents that intuitively denote the two things being compared. One of them, called the remnant, is the comparee and is in the than phrase. The other, called the associate, is the comparer and is usually the main clause subject (in a predicative comparative). So, in (30), Mo is the associate and Jo is the remnant. The difference between (30) and (31) has to do with whether there is a unique associate-remnant possibility in the sentence. So (31), unlike (30), has two NPs that are possible associates (Jo is the remnant), and depending on which is intonationally prominent, different readings emerge as primary. In one reading, Bo owes money to Mo and Jo, but
he owes more to Mo than he does to Jo. In this case, Mo is the associate and Jo, the remnant. In the second case, Bo and Jo both owe money to Mo and Bo owes more than Mo owes. In this case, Bo is the associate and Jo is the remnant.

In addition to the fact that each kind of comparative sentence has a comparative morpheme, they also share a common meaning component (presumably due to that morpheme) in that each sentence expresses a comparison between two degrees belonging to the same scale (e.g. in (31), the scale whose members are monetary values). Take the following examples of sentences and their meanings:

(32) Evan is taller than 6'.  
Evan's height > 6'

(33) Evan is taller than William.  
Evan's height > William's height

(34) Evan is taller than William is.  
Evan's height > William's height

(35) Evan is taller than William is wide.  
Evan's height > William's width

Even if you don't actually calculate the exact degree of Evan's height, the point here is that what is common to the comparative is a sense of one object or individual having a property that is greater than that of another object or individual (or, for less, less than).

The same is true for predicative, attributive, and nominal comparatives, as in (36)-(38).

(36) Evan is taller than William.  
Evan's height > William's height

(37) Evan has a faster drive than William.  
Speed of Evan's drive > Speed of William's drive

(38) More three-pointers were scored by William than Evan.  
Number of 3-pt shots scored by William > Number of 3-pt shots scored by Evan

There are a number of theories of comparison on the market, including von Stechow 1984, Kennedy 1999, Heim 2000, 2006, Schwarzschild and Wilkinson 2002 and
Schwarzschild 2004, just to name a few. All of the theories mentioned here agree that comparatives can only be formed with gradable adjectives (those that are not all-or-nothing, such as *fake* or *dead*) and that gradable adjectives involve reference to degrees in the ontology of natural language meaning. They differ with respect to whether (and to what extent) these degrees are actually extents (degree intervals), how those intervals are referenced, whether aspects of scale structure are also ontological objects, and what kind of operator exists in the comparative. I will abstract away from these issues here given that the general goal of this thesis is merely to use an analysis of comparison in order to show that the same lexical item can be used in the composition of the CC and RCC. On the basis of the similarity in the meanings of various kinds of comparative sentences that we saw above, it may seem that these are all derived using the same form and meaning for the comparative morpheme. However, as the literature listed above has demonstrated, multiple lexical entries for *more* are necessary given that the number and kind of arguments *more* takes, etc. differs between these constructions. They have also been shown to have different patterns with respect to scope possibilities, etc.

Nevertheless, authors disagree about how many of these constructions can be unified. For example, some authors such as Heim believe that comparative deletion and phrasal comparison can be unified, while others such as Kennedy argue that they must be different. Again, I sidestep these debates here and leave for future research a discussion of how comparison in the CC contributes to these issues. I assume, following Kennedy 1999 and Pollard 2008 and p.c, both of whom discuss associate-remnant comparatives directly, that these comparatives feature the same comparative morpheme as other phrasal comparatives. Because Pollard's work is already developed within a formal
framework similar to the one used here, I will take it as my point of departure. Thus, I hypothesize that the *more* in the CCs treated in this thesis are instances of phrasal comparison, which must also be able to account for associate-remnant comparison.

Motivating the kind(s) of comparative at work in the CC or RCC is made difficult by the fact that they do not have *than* phrases, since this is the typical criteria according to which the various categories are distinguished. However, on the basis of (32)-(35), one can see that there are some differences in the semantic constituent to the right of the greater-than sign. For example, instances of comparative subdeletion will always compare objects with respect to two different gradable properties: here, height v. width. This is potentially because #Evan is taller than William is tall sounds strange because it could be more efficiently stated as the fully-deleted comparative in (34) (though note that that same reasoning does not explain why (34) would both be felicitous given the presence of the simpler (33)). Regardless of reason, only phrasal comparatives and comparative deletion compare two individuals, etc. with respect to the same gradable predicate. As in Beck's analysis, we see that this is the kind of comparison that is required in the CC, where different individuals, times, or worlds are compared with respect to a single predicate, which is the only predicate in the clause since there is no *than* phrase. One thing that makes this interesting is that, if you were to assume that the form of the single predicate phrase in a CC clause were copied such that it was indicative of what would appear in a *than* phrase, given that many CC clauses considered here are 'missing' only their degree argument, then from that perspective, they would be most like instances of comparative subdeletion. However, as mentioned before, I consider this an impossibility given that any theory of comparatives in the CC that would be extended to
comparative subdeletion would incorrectly predict that the predicate would be the same on both sides of '<'.

Because I choose the phrasal comparative as the model for the comparative in the CC and because the phrasal comparative extends to the analysis of associate-remnant comparatives, the final thing that needs to be discussed before giving the analysis is the notion of parasitic scope. Parasitic scope, in the sense of Barker 2007, is when the scope of one operator in a sentence depends upon the scope of another in that the scope target for the former does not exist until the latter has taken scope (though this is stated in terms of a process with various steps, the scoping of both can occur together; it's just that the scope of one is the required by the scope of the other (exemplified further below)). Barker illustrates his analysis for sentences with an 'internal' reading of same (the reading in (39) where each person had the same waiter as one another and we may not even know who that waiter was).

(39) The same waiter served everyone.

The problem is that, assuming same has the meaning of a choice function (returning the single individual who was everyone's waiter in (39)), the universal quantifier has narrow scope with respect to the choice function. But normally, if you quantified it into the object position of served, it would be expected to be wide scope (or at least have that possibility). Since only the narrow-scope reading is possible, another solution is required. Barker claims that the scope of same is parasitic in that it 'hijacks' (Barker's term) the scope of everyone, “intervening between the quantifier and what would otherwise be its semantic argument” (429). Basically, here is how this works. Everyone
has as its scope target a property of type \(<e,t>\). \textit{Served}, like other verbs with two arguments like \textit{loves} is of type \(<e,et>\) but could combine first with traces, then abstract over the subject position and next take its subject NP argument (or be taken by it and then abstract over the object position) to yield an \(<e,t>\) in order to be the scope of \textit{everyone}, as we saw in the examples in the last section. Instead, \textit{same} is going to take the would-be sentence with two gaps, i.e. \textit{served} with trace arguments that are then abstracted over, and return a property. The property that it returns will require a particular kind of individual argument: a plural individual. This accounts for the fact that (40) but not (41) is felicitous.

(40) The same waiter served Alex and Colleen.

(41) #The same waiter served Alex.

In other words, \textit{the same waiter} is of type \(<<e,et>,et>\) instead of the typical quantifier type \(<et,t>\). We understand that the type returned needs to be \(et\) to require the plural individual, but this also relates to why it must take a doubly-gapped \(<e,et>\) argument as well; it is quantifying itself into one of those individual positions, as a regular quantifier would, but the other is the position that needs to be filled by a plural entity but is not yet specified as requiring the plurality of that entity (which is then what the result of \textit{same} does). To put it another way, \textit{the same waiter} takes an extra individual argument position and returns something that also has an extra individual argument position, changing it only in that it now requires a plural entity that the choice function can distribute over.

Though we might try to appeal to Cooper Storage in analyzing parasitic scope, Barker shows that it won't generalize from binding in-situ operators to this kind of
problem. This is largely the motivation for using PTDCG rather than a categorial grammar in which word order is not separated from the tectogrammatical component.

Carl Pollard, p.c., reformulated Barker's analysis of parasitic scope in a way that simplifies it while maintaining its essential insights. Both Barker and Pollard have also suggested that the analysis of parasitic scope with *same* should be extended to other comparative-type phenomena, such as internal readings of superlatives and associate-remnant comparatives. This is why it forms the basis of the analysis of comparatives given here.

We are now ready to look at an example of a comparative sentence. First, we need the lexical entries used for the following two sentences, as in (44) and (45).

(42) Mo eats more than Jo.

(43) Bo owes Mo more than Jo.

(44) a. Bo: ⊢ BO : NP : b
   b. Jo: ⊢ JO : NP : j
   c. Mo: ⊢ MO : NP : m
   d. eats: ⊢ \( \lambda z.\lambda u.\lambda EATS+z : \text{Deg-}\rightarrow(\text{NP-}\rightarrow S) : \lambda d.\lambda x.\text{eat}'(d)(x) \)
   e. owes: ⊢ \( \lambda z.\lambda u.\lambda s.\lambda OWES+z+u : \text{NP-}\rightarrow(\text{Deg-}\rightarrow(\text{NP-}\rightarrow S)) : \lambda x.\lambda d.\lambda y.\text{owe}'(x)(d)(y) \)
   f. than: ⊢ \( \lambda s.\text{THAN}+s : \text{NP-}\rightarrow\text{Th[NP]} : \lambda x.x \)

There is nothing surprising about the first three lexical entries; they are all proper names like ones we have seen above. (44d) and (44e) each contain verbs that take a subject and a degree argument, and *owes* takes an additional NP object argument as well. \( Z, u, \) and \( s \) are variables over strings, \( x \) and \( y \) are variables over semantic entities of type \( e \), and \( d \) is a variable over semantic entities of type \( d \). When verbs occur with measure phrases or other degree terms such as *very, few, how much/many*, etc., I assume that they take degree
arguments, which is uncontroversial. The entry for than shows that its NP argument will appear to its right in the phenogrammar, and it is treated as an identity function in the semantics. In the tecto term, we see something new, namely a $Th[NP]$. I take than to contribute a syntactic feature ($than$-marking) to the category that it takes as argument. This is needed so that other lexical items, such as more can require a constituent that is marked by than. This, too, is uncontroversial. In (44f), I have instantiated than as it appears in the phrasal comparative when it takes NP $e$-type arguments. In reality, than could have a more general entry, such as $\vdash \lambda \alpha. THAN^+ : X \rightarrow Th[X] : \lambda \beta. \beta$ where $\alpha$ is a variable over strings or functions on strings, $X$ is a variable over syntactic types such as NP, PP, Deg $\rightarrow$ S, etc., and $\beta$ is a variable over semantic variables of any corresponding type. We now turn our attention to the most important and most complicated lexical item, the comparative morpheme.

(45) more: $\vdash \lambda r. \lambda s. \lambda t. (MORE)(s)+t : (\text{Deg} \rightarrow NP \rightarrow S) \rightarrow (NP \rightarrow Th[NP] \rightarrow S) : \lambda G. \lambda x. \lambda y. more'(G)(x)(y)$

As in the lexical entries for someone and everyone, above, we see a variable over a functional type $r$ in the pheno term for more. Here, $r$ is a function from strings to strings. $S$ and $t$ are variables over strings. In the semantic term, $G$ is a function of type $<d,et>$, so it takes a degree and an individual to return a truth condition. $X$ and $y$ are variables over individuals and $d$ is a variable over degrees, as will be true throughout this chapter. As for the tecto type, more's first argument is a would-be sentence with two gaps, just like the first argument of same as we saw above. I will call this the double abstract for the remainder of this chapter, based on the fact that there are
two gaps, which represent two lambda-abstractions in the first argument in both the pheno and tecto terms (as we have just seen). Once more takes this double abstract with one Deg gap and one NP gap, it returns something that is looking for two NPs, one of which is marked by than. These arguments correspond to the associate and the remnant in the comparison. The double abstract is just a would-be regular sentence without a comparative, so it does not have a second argument position for whatever is in the than phrase. However, just as in other instances of parasitic scope, this scope position is created as a result of another operator, here, more. When you take a look at the lexical entry for more holistically, in each component there are three arguments: one complex argument with two things missing and two simple arguments following it. In the double abstract, one of the things missing corresponds to a degree, which is what more quantifies into (or is 'lowered' into, etc.), since it is a degree operator. The other thing that is missing from the double abstract matches the kind of thing that more takes as its next two arguments, and more essentially uses the information from the double abstract (such as what predicate forms the basis for the comparison, etc.) and uses it twice – once with each of the subsequent arguments. Based on this description, where what is important is that one of the gaps in the double abstract match the kind of thing that more take next, one can see that just as with than, this type could be generalized for comparisons of other kinds. In addition to the lexical entry, however, we need a meaning postulate for more in order to understand when a sentence with more' and its arguments is judged to be true:

(46) \([\text{more}](G)(x)(y) \text{ is True iff } \max(\lambda d. G(d)(x)) > \max(\lambda d. G(d)(y))\)

34 Here, \(\max\) is a maximality operator that takes a set of degrees and returns the maximal element in that
This is where we finally see how the pattern in comparative meaning that we saw earlier in this section is encoded. *Max* is an operator of type $<dt,d>$ that takes a set of degrees (or degree intervals, depending on how you define $d$) and returns a single degree or interval that is maximal in that set. Remember from above that $G$ is type $<d,et>$, however, which is not able to combine with *max*. This is why *more*'s meaning postulate specifies that it take a degree argument $d$ that is then abstracted over (i.e. $\lambda d.G(d)$) before it takes its individual argument (since it couldn't take its individual argument first).

Within each *max* operator, you will have the set of degrees to which $x$ or $y$ is $G$. Then the sentence is true if the associate's maximal degree is larger than the remnant's.

Turning now to the proof of (42), we have the derivation in (47).

---

set. To be more precise, this is actually a least upper bound (in order to avoid the problem of a scale without a unique maximum), but *max* is in keeping with what the literature uses, e.g. Kennedy 1999. 221
First, as in the case of quantifier scope in the last section, the proof starts (from the bottom) with the verb taking traces for both of its arguments. Next comes two hypothetical proof steps that abstract over those traces (pulling them out of the context) in the opposite order that they were taken as arguments. In other words, the result after the end of the fourth step is an alphabetic variant of the lexical entry for *eats*. This probably seems superfluous since we could have just started with the entry for the verb and skipped those four steps, but as we will see in the next derivation, doing that would
not generalize to associate-remnant comparatives or even to other phrasal comparatives that have more than two arguments. So in (47), those first steps are included because (42) and (43) are parallel in their analysis. Returning to the analysis in (48), the alphabetic variant of *eats* is a double abstract, so it can combine with *more*. The result puts MORE into the first argument slot in the pheno term of the double abstract. This corresponds to the degree position syntactically and semantically. With respect to the meaning postulate for *more* at this step, we have \( \text{more}'(\lambda d. \lambda x. \text{eat}'(x)(d)) \), which is equivalent to \( \lambda y. \lambda y'. \text{max}(\lambda d'. [\lambda d. \lambda x. \text{eat}'(x)(d)](d')(y)) > \text{max}(\lambda d'. [\lambda d. \lambda x. \text{eat}'(x)(d)](d')(y')) \), which, when beta reduced (lambda converted) is equivalent to \( \lambda y. \lambda y'. \text{max}(\lambda d'. \text{eat}'(d')(y)) > \text{max}(\lambda d'. \text{eat}'(d')(y')) \). Thus, we see that *more* takes the double abstract and converts it into something that simply needs the two individuals being compared to yield the correct meaning. This is exactly what the next two steps provide, as we will see momentarily.

The other potentially tricky part of the step where *more* takes its first argument is the lambda conversion in the phenogrammar. The pheno term of the double abstract, \( \lambda q. \lambda p. p+EATS+q \), goes in for \( r \) in \( \lambda r. \lambda s. \lambda t. r(\text{MORE})(s)+t \). This yields \( \lambda s. \lambda t. [\lambda q. \lambda p. p+EATS+q](\text{MORE})(s)+t \), which in turn becomes \( \lambda s. \lambda t. s+EATS+\text{MORE}+t \).

Finally, the rest of the derivation proceeds as expected: *mo* gets filled in as the subject, *jo* becomes a *than*-marked NP and goes into object position. Next, we turn to the analysis of the associate-remnant example from (43), which has the derivation in (48).
As discussed above, this sentence has two possible readings, and (48) obviously only represents one of them. This is the proof of the reading in which Bo owes Mo more than Bo owes Jo. I will discuss but not give the full derivation for the other reading. The first few steps are crucial for distinguishing these readings, which is in turn what distinguishes associate-remnant examples from simpler phrasal comparatives. The first argument of
owes is the argument that represents who someone owes something to. We have a choice between using a trace as this object or using Mo as the object. This choice is made based on the desired position of the associate and remnant in the comparative sentence (even though at this step in the derivation it just looks like we are building up a regular sentence with a degree argument). For example, in the paraphrase of the meaning of the reading derived in (48) (Bo owes Mo more than Bo owes Jo), Bo appears twice, Mo is the associate, and Jo is the remnant. Here, the comparison is between the two objects of the owing, whereas in the other reading (Bo owes Mo more than Jo owes Mo), the comparison is between the two subjects and Mo gets repeated twice along with owes. When the comparison is between objects, we use a trace in object position, and when the comparison is between subjects, we use a trace in subject position instead. When the trace is in object position, as it is in (48), then the subject NP is not a trace and Bo gets filled in for the subject NP. In the other reading, Mo gets filled in for the object NP right away and a trace is used in subject position (and Bo gets filled in after the double abstract combines with the comparative). This way of doing things assures a couple of things. First, you will only ever have two gaps, as required by the comparative morpheme: one will be a degree, and the other will be (in this case) an individual. It also guarantees that the individual gap will be in the position where the comparison will take place. Finally, it assures that any other argument that is neither the associate nor the remnant (in this reading, Bo) will already be incorporated within the double abstract, meaning that it will be repeated on both sides of the '>' in the meaning of the comparative sentence, which is what the truth conditions for this reading require (as indicated by the paraphrase discussed earlier). This also explains why the earlier derivation for a simpler phrasal
comparative in (47) used traces and hypothetical proof rather than just having more combine with owes right away. We see in (48) that that would not work because Bo is not the first argument of owes, so in order to get a double abstract where only the subject argument is filled in, it is necessary in this framework for the other arguments to be taken in order as traces, then abstracted over once the subject has combined. Once the double abstract has been formed, the rest of the derivation in (48) parallels the one in (47) in combining with more and taking the remnant and associate as arguments. Based on the final semantic term \( \text{more}'(\lambda d.\lambda x.\text{owe}'(x)(d)(b))(m)(j) \), we see that the meaning of the sentence on this reading is \( \max(\lambda d'.\text{owe}'(m)(d')(b)) > \max(\lambda d'.\text{more}'(j)(d')(b)) \).

This section has described natural language comparison and parasitic scope phenomena. A PTDCG theory that extends to cover both phenomena was illustrated and explained. As we will see in 4.3, this theory can be seamlessly incorporated into the analysis of the CC and RCC, but first, there is one more building block that is needed.

### 4.2.4 PTDCG Analysis of Unbounded Dependency

As we saw in Chapter 2, the long distance dependency in each CC clause between the comparative morpheme (and often an adjective, etc. that pied-pipes with the comparative morpheme) and the gap in the clause that follows it is one of the most noted features of the CC. That chapter also noted a number of other UDCs such as relative clauses. This section is devoted to the analysis of UDCs outside the CC in order to see what the key pieces of any UDC analysis are that can then be incorporated into a CC or RCC analysis. Taking the relative clause as my example, I analyze the sentence in (49).

(49) Jen loves the book that Ben wrote.
The new lexical entries required for the analysis of this sentence are listed in (50) and include *the*, *book*, *that*, and *wrote*. Beginning with *the*, we see that it combines with a noun to form a noun phrase and that semantically, it combines with a property to return an individual. The iota in the semantic form is a variable binder that requires that there be a single unique element $x$ in the set whose description follows. Because nouns, like noun phrases, are simple strings, the pheno term for *the* simply takes the sound string corresponding to the noun and places it after THE. The meanings of *book* and *wrote* are self-explanatory, and *wrote* mirrors other two-place verbs we've seen before. That leaves *that*, which is the lexical item whose presence is indicative of a UDC. Not surprisingly, then, particular aspects of its lexical entry are what are crucial for the analysis of the UDC in this sentence. Here, $l$ is a variable over functions from strings to strings, and $q$, like $p$ is always a variable over strings. The innovation in the phenogrammatical term is that the function $l$ receives an argument $e$, which is not a variable; it is the empty sound string. In other words, silence is fed into the argument position of this function. The tecto and semantic components of *the* are straightforward; it takes two properties $(\text{NP} \rightarrow \text{S} \& \text{N})$ and its semantic form conjoins them and predicates each of the same individual.
The proof tree for (49) as in (51) demonstrates the viability of these lexical definitions in arriving at the intended word order, category, and meaning for this sentence. First, we see that a trace appears in the embedded clause where the gap is.
Then, the subject can combine with the verb (as would any other argument of the verb). Next, a hypothetical proof step removes the trace from the context and creates an abstract with a single gap where the extracted argument would have been. This NP-project S is the kind of thing that *that* takes as argument to return a noun modifier. The pheno term of the NP-project S is a function from strings to strings, and the string that it takes as argument gets placed on the right periphery of the phrase. Because *that* feeds it the null string, we expect there to be nothing to the right of the verb in that phrase. After taking its first argument, *that* is a function from strings to strings that adds its next argument to its left. This argument it takes is the string corresponding to a noun, but once that whole thing is the argument of *the*, we see that what would have been in the object position of the embedded clause (*the book*) appears to the left of *that* rather than the right of *wrote*. In other words, it is primarily the fact that the pheno term of *that* feeds an empty string into the function that is its first argument and then returns a function that concatenates a future argument string to its left that handles UDCs. The fact that the dependency is unbounded, i.e. that the gap can be embedded arbitrarily deeply can be treated in this analysis because hypothetical reasoning can cross any number of sentence boundaries. Once a trace is introduced, it can stay in the context as long as desired, creating an NP-project S from an S whenever it is required by a relative pronoun or other 'movement trigger' in a UDC. Of particular note is the fact that this analysis accounts for cases in which the gap is phrase-medial (i.e. not on the right periphery) to begin with, which is a problem for many kinds of categorial grammar (cf. Moortgat 1988). This is because *that* feeds the null string to whatever position was occupied by a trace, regardless of where it appeared in the embedded phrase.
Returning now to the final steps of the derivation, the *that* clause combines with *book* to form another noun with a more complex meaning than just *book*. Nevertheless, it is just a single property, which is the right kind of argument for *the*. After *the* combines with *book that Ben wrote*, the semantic term denotes the unique individual that is a book written by Ben. Thus, it is an individual-type noun phrase whose pheno term is a single string, making it just like any other object NP argument of a verb like *loves*. The rest of the derivation is just like that of a simple sentence.

### 4.3 PTDCG Analysis of the CC and RCC

In this section, I finally give an analysis of the CC in PTDCG. I draw on the analyses of comparison, unbounded dependencies and quantifier scoping that we have seen in the last several sections in creating this analysis. Given the complexity of the derivation, I will break the components down one by one before giving the full analysis. Because the overall proof system in PTDCG is based on the tecto types, I begin with the syntactic categories of each expression in the CC in (52) as in (53). Because the CC clause does not resemble any other syntactic constituent in English, I add *CC* to the set of basic syntactic categories to indicate a matrix CC clause, which you can see in the category specification of both *the's* in (53). I assume that *sem(CC) = <x,xt>* where *x* is a variable over semantic types. In other words, it corresponds to a set of pairs of some kind, whether they be pairs of individuals, occasions, times, situations, etc.

(52) The more a dog eats, the more it drinks.
The more a dog eats, the more it drinks.

At the topmost level in (53), the CC as a whole is a sentence. Below that, the subordinate clause takes the matrix clause as argument to return a sentence. Below that, the subordinate clause and matrix clause are built up in parallel with all of the same tecto types except for the difference in the rightmost (output) type at each step, since the subordinate clause is building to a CC-S, while the matrix clause is building to a CC. In other words, each (X-CC-S) type in the left clause matches a (X-CC) type in the right clause, where X is any string of tecto types. Looking at each clause more closely, we see that the2 takes three arguments: more, a generalized quantifier, and a double abstract. The1 takes the same three kinds of arguments before taking the matrix clause as its final argument. I will call the generalized quantifier argument the pseudo-associate since its
semantic role is analogous to that of the associate of an associate-remnant comparative sentence. As we will see in the pheno term derivation, the ends up taking the pseudo-associate and quantifying it into the non-degree gap in the double abstract just as more would do, but the argument structure as reflected by the tectogrammatical types is different. Because this component of grammar does not tell us anything about word/phrase order or interpretation, this is all there is to say about (53). Note, however, that in this derivation, as in all the others in this chapter, what is important at each Modus Ponens step is which two formulas are combining and their relative order in the tree representation is of no consequence. In other words, wherever one constituent is on the left and the other on the right, they could just as easily have been reversed because with Modus Ponens there is never any question which of the two premisses is the major one and which is the minor one.

One thing that is missing in (53), however, is the part of the derivation that stems from eats and from drinks. In this derivation, it seems that this could represent just the lexical entry for that verb (since, indeed, both are lexically specified as type Deg-NP-S). As in the example of phrasal comparatives in the last section, however, this is a red herring and would not be true of a verb that took more than two arguments. Since I have demonstrated this in a previous section, I will continue to leave out the sub-derivation that creates the double abstract in all of the CC clause and sentence derivations, but I do illustrate it in (54) below with all three components listed together.
As in previous sections, we see here that drinks combines with a degree trace and an additional trace in the position of its associate. It would additionally combine with any other arguments that are neither the degree nor associate trace at this stage, and none of those would be traces. Once a sentence has been formed (though not one that is capable of standing alone given that the hypotheses in its context have not been discharged), we begin the two hypothetical proof steps to discharge the traces from the context. There are always exactly two of these steps because we have only two traces in the construction of the sentence, which is turn true because we are deriving a double abstract (otherwise, the types won't be able to combine with more or the or anything else that requires a double abstract as one of its arguments). Of these two hypothetical proof steps, the degree trace will always be discharged last, again because of the specification of more, the, etc. that it be a function from degrees to some other type of phrase that is indicative of what the associate is. Now that we have seen the derivation of the double abstract, I return to the analysis of the CC as a whole. Next up: the semantic component. I have divided the
semantic analysis across two derivations. The first is the derivation of the meaning of (52) as a whole and the subderivation of the first CC clause, as in (55). The second is the subderivation of the second CC clause, as in (57).

(55) The more a dog eats, the more it drinks.
\[ \text{CORREL}(\lambda a. \lambda b. \exists c. a=c \land \text{dog}'(c) \land \exists h. b=h \land \text{dog}'(h) \land \text{more}'(\lambda d. \lambda x. \text{eats}'(d)(x))(a)(b)) \]
\[ (\lambda x. \lambda y. \exists z. x=z \land j=z \land \exists v. g=v \land y=v \land \text{more}'(\lambda d. \lambda a. \text{drinks}'(d)(a))(x)(y)) \]

The more a dog eats
\[ \lambda R. \text{CORREL}(\lambda a. \lambda b. \exists c. a=c \land \text{dog}'(c) \land \exists h. b=h \land \text{dog}'(h) \land \text{more}'(\lambda d. \lambda x. \text{eats}'(d)(x))(a)(b))(R) \]

The more a dog
\[ \lambda G. \lambda R. \text{CORREL}(\lambda a. \lambda b. \exists c. a=c \land \text{dog}'(c) \land \exists h. b=h \land \text{dog}'(h) \land \text{more}'(G)(a)(b))(R) \]

The more
\[ \lambda P. \exists c. P(c) \land \text{dog}'(c) \]

The more
\[ \lambda H. \lambda x. \lambda y. \text{more}'(H)(x)(y) \]

First things first: the meaning of the sentence as a whole is rendered as the existence of a correlation between two sets of individuals. These sets are \( \lambda a. \lambda b. \exists c. a=c \land \text{dog}'(c) \land \exists h. b=h \land \text{dog}'(h) \land \text{more}'(\lambda d. \lambda x. \text{eats}'(d)(x))(a)(b) \) and \( \lambda x. \lambda y. \exists z. x=z \land j=z \land \exists v. g=v \land \text{more}'(\lambda d. \lambda a. \text{drinks}'(d)(a))(x)(y) \).
\[ y = v \land \text{more}'(\lambda d.\lambda a.\text{drinks}'(d)(a))(x)(y) \]. If we use the meaning postulate for \text{more}, these translate into \[ \lambda a.\lambda b. \exists c. a = c \land \text{dog}'(c) \land \exists h. b = h \land \text{dog}'(h) \land [\text{max}(\lambda d.\text{eats}'(d)(a)) > \text{max}(\lambda d.\text{eats}'(d)(a))], \] and \[ \forall x. \forall y. (\exists z. x = z \land j = z \land \exists v. g = v \land y = v \land [\text{max}(\lambda d.\text{drinks}'(d)(x)) > \text{max}(\lambda d.\text{drinks}'(d)(x))], \] respectively. In the first case, this gives us the set of pairs that are dogs where one dog eats more than the other. In the second case, this gives us the set of pairs of individuals where one drinks more than the other. For a reason that I will explain more later, the individuals in the second clause refer back to the individuals in the first clause, further restricting the domain of drinking individuals to drinking dogs. Thus, the meaning of the sentence is taken to be the existence of a correlation between pairs of dogs with a difference in how much they eat and pairs of dogs with a difference in how much they drink. Thus, eating increases and drinking increases are correlated across dogs. And indeed, this is the meaning we want for the sentence \textit{the more a dog eats, the more it drinks} on the reading that compares different dogs to one another.

You may be wondering at this point why all of the equalities between individual-type variables are needed to derive this result. It is because the pseudo-associate (\textit{a dog}, \textit{it}) has a variable that is already bound by an existential quantifier. If that variable is equated with another (free) variable, that free variable can be bound by something else; in this case, the lambda. Additionally, by carefully choosing the variables that the existentially-bound variables will be equated with to be those that are also substituted as the final two arguments of \textit{more}, this guarantees that the individuals whose eating or drinking amount is being compared will also be those that are asserted to exist by the pseudo-associate. Thus, one of the primary things that \textit{the} is doing semantically is tying together the individuals quantified over in its various arguments. \textit{The} also \lambda\text{-abstracts
over the individual variables that it used in this process to create a set of individuals with
the property included in the GQ/pseudo-associate and the relationship described by more.

To see exactly how the does this, look at the bottom-most entry in the proof tree. Here,
Φ is a variable of type \(<d,et>,<e,et>\) (here, more), \(B\) is a variable of type \(<et,t>\) (the
pseudo-associate), and \(G\) is a variable of type \(<d,et>\) (the double abstract). \(R\) is a
variable of type \(<e,et>\) (the meaning of the “main” clause), and it is the argument that
will be filled by the second CC clause. The \(R\) variable stands in for the second set of
pairs that is an argument of CORREL and thus gets filled in on the extreme right of the
expression. All of the other arguments of the are crucial to building the first set of pairs
that is an argument of CORREL, and we will see in the next derivation that this process
is identical to what happens with the in building the meaning of the second CC clause to
be a set of pairs (except that the pseudo-associate is a pronoun instead of an indefinite).

The's first argument is more, which goes in for \(\Phi\). There is nothing particularly
surprising about this or the fact that \(G\), the double abstract, is assigned by the to be the
first argument of \(\Phi\). So \(\Phi(G)(a)(b)\) is just giving more exactly what it would have taken
as argument had \(\Phi\) taken \(G\) as argument directly rather than via the. If this were a regular
comparative sentence, \(a\) and \(b\) would be constants rather than variables, but, as in Beck's
analysis, it is crucial to deriving the correct meaning for the CC that these be variables.

Moving to the next argument, which is the pseudo-associate, one thing to note is that
when the meaning of a dog gets taken as argument by the more, it gets put in two places
in the semantic representation (given that there are two \(P\)'s in that expression). This, too,
is necessary in order for a single indefinite to introduce two discourse referents (to use
the terminology of a dynamic theory of meaning). Here, that is modeled by having the
indefinite introduce two existential entailments. The existential binding itself is already part of the meaning of *a dog*, but the fact that *a dog* is plugged in for $\mathcal{P}$ means that two existentials will be present in the semantic representation of the restriction on this set of pairs. As you can see, I have used different variables (c versus h) in each instance of that predicate. These are alphabetic variants of one another, which makes this a licit change that I thought necessary in order to not confuse the reader into thinking that these would be co-referential. Speaking of co-reference, however, one of the clever parts of the meaning of *the* is that it enforces the co-referencing of *a* and *c*, etc. by filling something particular in for the argument of the GQ; here, *a dog*. So, as we saw in the section on quantifier scoping, above, *a dog* would normally take a property like *walks* to return a truth value. It is an advantage of the present analysis that the same lexical entry for the indefinite article that was used there is also used here without modification. In other words, even though *a dog* does not take a predicate argument in the CC and needs to be instead co-referenced with the variable arguments of *more*, etc., all of this can be done through the lexical specification of the meaning of *more* rather than the creation of a separate lexical entry for *a*. *The* gives a particular kind of property as the argument of $\mathcal{P}$. That property is a two-place predicate (with one argument already taken) that I have called $Eq$, for 'equate'. Its meaning postulate appears in (56).

\[ (56) \quad [[Eq]](x)(y) \text{ is True iff } x=y \]

On the basis of (56), then, $Eq(a)$ in (55) is a property that will equate *a* and whatever argument it takes next. In this case, that argument is the variable that is existentially bound in *a dog*. This produces the desired effect. Looking now at the meaning of the
second CC clause, we see that there are two equalities per instance of existential quantification. One of these (per instance of \( \exists \)) is a result of the effect of \( Eq \), just as in the first CC clause. The second in each is part of the meaning of the pronoun \( it \). We'll see this in more detail in the derivation in (57), but I mention it here in order to explain how anaphora is handled in the static system in PTDCG. Following Roberts 2004, pronouns are treated as definite, but not as definite descriptions. Equating the existentially bound variable in the pronoun with a free variable (this is the equality that is encoded in the meaning of the pronoun) creates a familiarity presupposition for that pronoun that there must be a familiar individual referent in the discourse context that this free variable is meant to be associated with. Then, the existential entailment of \( a \ dog \) has introduced just such a kind of individual-type discourse referent, so the free variable in the denotation of the pronoun is associated with the existentially-bound variable in \( a \ dog \). Following the suggestion of Chierchia 1995, inter alia, this is guaranteed on the basis of the conservativity of the operator CORREL as discussed in the last chapter. The conservativity of the operator requires that the objects of predication in one clause be the objects of predication in the next, which, in combination with the existential entailments in the first clause and the familiarity presupposition of the pronouns in the second, produces the correct intended meaning.

There is no doubt that an existential-binding theory of the meaning of indefinites runs into problems with respect to anaphora, just as unselective binding theories and other newer e-type theories do also. There is also no doubt that a dynamic theory of meaning is necessary to successfully capture the full generality of anaphora in discourse. The point here, however, is that any working theory of anaphora can be plugged into this
analysis in place of the way I am treating it here. This goes back to the general point that the CC's analysis is largely a combination of analyses of other phenomena.

Turning now to the analysis of the matrix CC clause, we see that it parallels the development of the subordinate clause. Related to the last discussion of the analysis of pronouns and anaphora, we see that it has an equation with a random free variable \( j \) in place of the kind of restriction that one might normally find in an NP formed on the basis of a noun. In other words, instead of something like 'dog', it has \( z=j \). Both quantificational NPs share the lambda abstraction over a property-type variable that has the existentially-bound variable as its argument.

\[
(57) \quad \text{The more it drinks (semantics)}
\]

\[
\lambda x. \lambda y. \exists z. x = z \land j = z \land \exists v. g = v \land v = y \land \text{more}'(\lambda d. \lambda a. \text{drinks}'(d)(a))(x)(y)
\]

\[
\lambda G. \lambda x. \lambda y. \exists z. x = z \land j = z \land \exists v. g = v \land v = y \land \text{more}'(G)(x)(y)
\]

\[
\lambda G. \lambda x. \lambda y. \lambda P. \exists z. P(z) \land j = z
\]

\[
\lambda P. \lambda G. \lambda x. \lambda y. \text{more}'(H)(x)(y)
\]

\[
\lambda \Phi. \lambda P. \lambda G. \lambda x. \lambda y. \Phi(G)(x)(y)
\]

Everything else in (57) is simply a matter of \( \lambda \)-conversion. \textit{The(2)} takes \textit{more}, the pseudo-associate and the double abstract one at a time to yield a set of pairs of
individuals with a degree difference in terms of how much they drink. If this were a
typical comparative sentence, the top-most line in the derivation would be a point at
which two individual arguments could be taken to form a sentence. This is where the
semantic component works in tandem with the tecto type, however, in that the constituent
that is at the top of this tree is of type CC, which is a basic type that does not take further
arguments. For that reason, no than phrase (licensed by this comparative) is possible,
just as in Beck 1997.

Moving on, we have the phenogrammatical component's derivation in (58).

(58)

The more a dog eats, the more it drinks
THE+MORE+A+DOG+EATS+THE+MORE+IT+DRINKS

The more a dog eats
λs'.THE+MORE+A+DOG+EATS+s'

eats
λq.λp.p+EATS+q

The more a dog
λr'.λs'.THE+MORE+(r'(e))(A+DOG)+s'

a dog
λk.k(A+DOG)

more
λr.λs.λt.r(MORE)(s)+t

The more
λv.λr'.λs'.THE+MORE+v(r'(e))+s'

the(2)
λu.λv.λr'.THE+u(iso)(e)(e)+v(r'(e))+s'

The(1)
λu.λv.λr'.λs'.THE+u(iso)(e)(e)+v(r'(e))+s'

the more
λv.λr'.THE+MORE+v(r'(e))

more
λr.λs.λt.r(MORE)(s)+t

it
λk.k(IT)

the more it drinks
THE+MORE+IT+DRINKS

drinks
λq.λp.p+DRINKS+q
In this proof, the top-most levels are the easiest to understand. Obviously, at the last step, we have the correct word order for the CC sentence we are proving. Below that, the matrix clause is a fully-formed string, and the subordinate clause is a function from strings to strings that puts its argument string on the right of the already-existing string that corresponds to the pronunciation of the first CC clause. The steps below that are more complex. The next step down combines the double abstract with *the more a dog* or *the more it*. In each case, the double abstract goes in for \( r' \) and has the empty string as its first argument, which is the place in the phenogrammar that a degree argument would have gone. This is related to the analysis of fronting that we saw in the last subsection.

The null string is substituted for where the degree argument would have been in situ, and *more* (which quantifies into the degree argument) appears to the left of the clause instead. The second argument of \( r' \) is whatever the pronunciation of the pseudo-associate for that clause is; here, either *a dog* or *it*. In this example, the individual argument in the double abstract was the subject, to the left of the verb. Thus, the pronunciation of the pseudo-associate gets filled into that argument position, generating the correct pronunciation for the clause where *a dog* precedes *eats*, etc. Another step down and we see the combination of the pseudo-associate with *the more*. The pseudo-associate goes in for the variable \( v \), which is a function that takes a function from strings to strings and returns a string. In this case, the function it takes from strings to strings is \( r'(e) \). As we know, \( r' \) is a function that takes two strings to return a string, which is why it has already taken its first argument before serving as the argument to \( v \), since \( v \) requires a function on a single string. Because the pseudo-associate, like other GQs we've seen, takes a function and
then substitutes its own pronunciation as the argument to that function, the same is true here. $R'(e)$ needs a second argument, and the pseudo-associate puts the string that is its pronunciation in as that argument, turning $v(r'(e))$ into $r'(e)(A+DOG)$ or $r'(e)(A+IT)$. In the bottom-most step, the combines with more. This is a bit tricky. In the example of long-distance dependency that we saw earlier, the fronted element (book) had a pheno term that was a simple string. In this example, more is fronted, which is a complex function that takes a two-place function on strings and then two strings as arguments and puts its own pronunciation in as the first argument in the two-place function. In the example of fronting from the last section, the clause with a gap was fed the empty string for where the fronted argument would have gone. This makes sense given that the whole point of fronting is that the would-be object is no longer there. Similarly, then, it is intuitively the case that if the fronted element were something that took strings or functions on strings around it as arguments, those would need to be filled it with empty strings as well since it now appears to their left rather than in situ. However, one of the arguments of more is not just a string; it is a function, so filling in the empty string there will not work. The purpose of putting in the empty string is to isolate the constituent in question from what would have otherwise been a dependency on another string. Thus, I create the operator iso as in (59) to 'isolate' its first argument, MORE.

(59) $iso_1(x)(y)....(x_{32}) = x$

The way that I have defined iso makes it a polyvalent function, capable of taking any number of pheno arguments. The point is that it always returns exclusively the first argument. In the case of iso in (58), it is a two-place function, equivalent to $\lambda x.\lambda y.x$. The
subscript 1 on the operator in (59) is there to indicate that it is part of a larger set of possible functions, each of which isolates a different argument. For example, \texttt{iso}_2 would isolate the second of its arguments. Returning to the bottom-most steps of the proof in (58), we see that \texttt{the} puts the two-argument version of \texttt{iso}_1 in as the first argument of \texttt{u}, which gets filled by the pheno term for \texttt{more}. Because the pheno specification of \texttt{more} is such that it puts \texttt{MORE} into the first argument position of that function, \texttt{iso} ends up isolating \texttt{MORE} in the position directly to the right of \texttt{the}. Technically, it ends up turning $\lambda r. \lambda s. \lambda t. r(MORE)(s) + t$ into $\lambda s. \lambda t. \texttt{MORE} + t$. This is because it does not affect the other arguments of the pheno term of \texttt{more}, only the arguments of \texttt{r} (which are its arguments). Given that \texttt{s} was a non-initial argument of \texttt{r}, though, we see that it has been removed from the latter portion of the term so that it is lambda bound but there is no occurrence of the variable in the formula. Nevertheless, the pheno term of \texttt{more} has two remaining arguments, which is why \texttt{the} further specifies that it take two more empty strings as argument. In the case of \texttt{s}, obviously this doesn't matter, but in the case of \texttt{t}, it does in fact remove a position that a string might have otherwise occupied. The substitution of these empty strings is actually not related to the fronting of \texttt{more}. This time, the empty strings are substituted in order for the pheno term to match the tecto type in being fully saturated. In other words, even though the final semantic term for \texttt{e.g.} the matrix clause is a function on pairs of individuals, the tecto type and pheno term are not functions. This accounts for the data in that those unfilled semantic arguments remain in the final sentence-level meaning of the CC and thus should never be able to be filled, nor should their corresponding pheno terms. So instead, those pheno locations are eliminated. This concludes the explanation of the derivation in (58).
The next big question now that we have seen the compositional analysis of the CC for each component according to the same rules and a simple pheno-tecto-semantics interface is how the corresponding RCC in (60) should be treated. Here, I only show the first clause of the RCC since that is the one that differs substantially from what we have seen above. This clause combines with a the1-initial subordinate CC clause just the way the matrix CC clause above did. One exception is whereas the1 takes its last pheno argument and concatenates it to its right, this time the1 will concatenate the pheno string associated with the RCC clause it takes as its last argument to its left. The only other exception is that a dog and it have switched clauses so that the indefinite is always introduced first (this doesn't affect anything).

(60) A dog drinks more, the more it eats.
The first thing to notice is that the RCC clause is treated as having a silent *the* (*the3*) out front. This is necessary in order for the clause to have the tecto type CC, which is the required argument of a *the1*-initial clause. It is also necessary because *the* takes the pseudo-associate as a separate argument from the double abstract and, in terms of their semantics, combines them into a different representation than what would result from the double abstract combining with the pseudo-associate independently of *the*. The next thing to notice is that *the3* combines with the exact same arguments as *the2* in the exact same order. This might be surprising were it not for the fact that the order of arguments does not correspond to word or phrase order in pronunciation in PTDCG. In fact, there is
absolutely no difference between the matrix clause in a CC and the matrix clause in the 
RCC with respect to the tecto derivation or the semantic derivation. The only difference 
is in the phenogrammar. Further, the pheno terms are identical for all of the lexical items 
that these two clauses share. That means that the only difference is in the pheno 
specification of the3 as compared to the2. The most obvious difference between these is 
that the3 does not contain the string THE anywhere in it, as well it shouldn't since no the 
is pronounced in the surface string. Looking more closely at the difference, we're 
comparing $\lambda u. \lambda v. \lambda r. v(\lambda s.u(r)(s)(e))$ to $\lambda u. \lambda v. \lambda r'. T H E + u(\text{iso})(e)(e) + v(r'(e))$. They 
obviously take the same three arguments, but there are major differences. First, $r/r'$ (the 
double abstract argument) does not take an empty string as argument in the3, nor does $u$ 
(the comparative argument) take iso. These are related. It is because the3, unlike the1 or 
the2, is not a movement trigger. In other words, there is no long-distance dependency in 
the matrix of an RCC, and these two empty/isolating elements were responsible for 
encoding that long-distance dependency in the same way as it is handled elsewhere in 
PTDCG. Thus, they are no longer present given a lack of long-distance dependency.

Next, even though it may not look like it at first glance, $v(\lambda s.u(r)(s)(e))$ and 
$v(r'(e))$ share the fact that they put the pronunciation of the pseudo-associate into the 
place where the associate would have been pronounced in a regular sentence. The 
double associate, as we are well aware by now, has degree and individual gaps for this 
example in the semantics. Those correspond to positions before and after the 
pronunciation of the verb in this example. What I am saying is that despite other 
differences between the parts of the pheno terms for the3 and the2 that begin with $v$, both 
manage to place A+DOG or IT in the spot in the double abstract's pronunciation where
the pronunciation of its individual-type argument would have been. The difference is that \textit{the3} also makes sure that the pronunciation of \textit{more} gets put into the spot in the double abstract's pronunciation where the pronunciation of its degree-type argument would have been. Here's how this happens. First, \textit{more} combines with \textit{the3} and it requires three arguments, all of which are provided. \textit{R} is function that takes \textit{MORE} as its first argument and \textit{s} as its second argument, where \textit{s} is also the second argument of \textit{more}'s pheno term. Finally, as in \textit{the1} and \textit{the2}, \textit{the3} fills in the empty string for the last argument of \textit{more}'s pheno term since the specification of two variables for \textit{more} in the semantics means that no \textit{than} phrase is possible in association with this comparative in the phenogrammar. The reason that \textit{the3} includes a \textit{\lambda s} before the \textit{u(r)(s)(e)} term is because \textit{u(r)(s)(e)} is a single string in terms of its pheno type, but \textit{v} (the pheno term of the pseudo-associate) requires an argument that is a function from strings to strings. We know that \textit{s} holds the place of what would be the pronunciation of the associate within a double abstract, which is the position in which we want the pronunciation of the pseudo-abstract to appear. Thus, \textit{the3} requires abstraction over \textit{s}, which makes this the right kind of argument for \textit{v}. Then, given that \textit{v}'s argument is \textit{\lambda k.k(A+DOG)}, it takes \textit{u(r)(s)(e)} and puts \textit{A+DOG} in for \textit{s}, just as we hoped. Finally, the double abstract's pheno term goes in for \textit{r} to yield our desired result in which nothing is fronted, MORE is in the position of a degree word and \textit{A+DOG} is in the position of the subject (associate).

The final derivation in this section is an example of a sentence that quantifies over something other than individuals. There are many interesting questions within the study of tense, aspect, and modality that relate to time and world arguments of verbs or verb phrases (or even sentences). Because it would be too much to try to tackle these
issues in this thesis, I will restrict myself to examples in which the non-individual arguments are additional arguments of the verb and have some concrete pronunciation associated with them. To this end, I will analyze the example in (62).

(62) The more Mo eats on an occasion, the more Jo drinks then/on that occasion.

This is a case in which quantification is over occasions or events, so the sentence-level meaning of the sentence is a correlation between pairs of events in which there is a difference in how much Mo eats and pairs of those same events in which there is a difference in how much Jo drinks (perhaps Mo and Jo are a couple, Mo has been gaining a lot of weight and Jo is dismayed about it and drowns her sorrows). Quantification over events is distinct from quantification over times. Compare (62) to (63).

(63) The later it is, the less humid it is.

(64) The hotter it was, the more tired Uli was.

Example (64) is Beck's illustration of quantification over times, the only one she gives. We can see that this is truly a case of quantification over occasions rather than times. Like Beck, however, events aren't specified as part of the inventory of basic types that I defined above. Since I am not treating multiple different kinds of examples here, I'll just use the $i$ type and its variables to indicate occasions here even though I recognize the difference. Similarly, (65) is an example that more clearly quantifies over possible worlds, while (66) is Beck's example of quantification over possible worlds.

(65) The larger the gravitational constant, the more I would weigh.
(66) The slimier an attorney, the more efficient he is.

In addition, (67) and (68) show that quantification is not limited to basic types, however many there may be. Quantification over propositions \(<s,t>\) as in (67) and properties \(<e,t>\) as in (68) show that the range is wider than previously expected.

(67) The more preposterous a rumor is, the more that people believe it.

(68) The more things you can do, the better your chances of winning.

These examples are relevant for the general point that an analysis of CCs must be able to be easily extended to a wide variety of types. For each of the three types noted by Beck, she posits different, lexically ambiguous the's. As noted above in this section, however, the use of variables over types in lexical specifications of these entries make this analysis extendable to any type (for example, more can be specified as

\[(\text{Deg} \rightarrow X \rightarrow S) \rightarrow (X \rightarrow \text{Th}[X] \rightarrow S)\]

so that it can be extended beyond NP gaps without requiring another lexical item). For cases in which a subset of possible types could be quantified over, we can create names for those sets. For example, one possibility suggested on the basis of Lewis' 1975 notion of 'cases' is a variable c that can range over individuals, times, and worlds, but perhaps not higher types. This could lead to a re-evaluation of the basic type system and sem values from earlier sections. I originally separated PP(T) and PP(W) in section 4.2, but prepositional phrases can also denote individuals, such as on the farm, etc. Rather than positing multiple different basic syntactic types that are all prepositional phrases, being able to make reference to a fixed set of types as a single entity enables a single PP category such that \(\text{sem}(\text{PP}) = c\) or
whatever the relevant subset is, which preserves the ability of \textit{sem} to be a function (which having \textit{sem}(PP) = i or s wouldn't) while not proliferating syntactic types. The same is true for NPs and other common verbal arguments. The goal of the derivation of the reading in (62), then, is just to demonstrate the feasibility of extending the analysis of CCs and RCCs where the pseudo-associate is a GQ over individuals (as in the examples in this section so far) to other types for a concrete example.

One reason that I began with the example of quantification over individuals rather than times has to do with the fact that every sentence requires a subject. The clause that \textit{more} is extracted from in the CC is no exception. The only time that a CC clause seems to appear without a subject is when its subject has been pied-piped with the comparative morpheme, as in \textit{the more people that come to the party, the more ice we'll need}, in which case the subject is present, but at a distance. Thus, just as the associate-remnant comparative was slightly more complex than the phrasal comparative because there was an extra argument, so too is the CC quantifying over times or occasions, etc. slightly more complicated than the one quantifying over individuals because there is always an individual in the form of a subject in addition to the quantifier over times, etc. that is the pseudo-associate. Thus, I begin with an example of the subderivation of a double abstract for the example in (62) as in (69).
Unlike the examples above, here, the lexical specification for *drinks* is not an alphabetic variant of the topmost element of the proof tree. It begins the same way as always by taking a degree trace, but it next takes a concrete individual rather than a trace as its subject. That leaves the PP/time argument of the verb, which, in order for the abstract to have two gaps, must be a trace as well. After that, we have a sentence, at which point we bring the time argument out of the context before the degree trace, which comes out last. Up to this point, any time a $t$ variable has been used, it has been a phenogrammatical variable over strings. Now, it will additionally be used as a variable over times when it
occurs in the semantic component (and will continue to be a variable over strings when it appears in the phenogrammatical component). In addition, I use arabic numerals as variables over times, as we see in the derivations of each clause of the CC in (62), the first of which is shown in (70).

\[(70)\quad \vdash \lambda s'. \text{THE+MORE+MO+EATS+ON+AN+OCCASION} + s':\]
\[\text{CC} \rightarrow S: \lambda R. \text{CORREL}(\lambda 1. \lambda 2. \exists t. t=1 \wedge \exists t'. t'=2 \wedge \text{more'}(\lambda d. \lambda t. \text{eats'}(d)(m)(t))(1)(2))(R)\]
\[\vdash \lambda q. \lambda p. \text{MO+EATS} + q+p: \]
\[\text{Deg} \rightarrow \text{PP} \rightarrow S: \lambda d. \lambda t. \text{eats'}(d)(m)(t)\]
\[\vdash \lambda r'. \lambda s'. \text{THE+MORE}+(r'(e))(\text{ON+AN+OCCASION}) + s': \]
\[(\text{Deg} \rightarrow \text{PP} \rightarrow S) \rightarrow (\text{CC} \rightarrow S): \lambda G. \lambda R. \text{CORREL}(\lambda 1. \lambda 2. \exists t. t=1 \wedge \exists t'. t'=2 \wedge \text{more'}(G)(1)(2))(R)\]

on an occasion: \[\vdash \lambda k. k(\text{ON+AN+OCCASION}): \]
\[(\text{PP} \rightarrow S) \rightarrow S: \lambda P. \exists t. P(t)\]
\[\vdash \lambda v. \lambda r'. \lambda s'. \text{THE+MORE}+v(r'(e)) + s': \]
\[(\text{Deg} \rightarrow \text{PP} \rightarrow S) \rightarrow ((\text{Deg} \rightarrow \text{PP} \rightarrow S) \rightarrow (\text{CC} \rightarrow S)): \lambda \Phi. \lambda G. \lambda R. \text{CORREL}(\lambda 1. \lambda 2. P(\text{Eq}(1)) \wedge P(\text{Eq}(2)) \wedge \text{more'}(G)(1)(2))(R)\]

More: \[\vdash \lambda r'. \lambda s'. \lambda t'. r'(\text{MORE})(s') + t': \]
\[(\text{Deg} \rightarrow \text{PP} \rightarrow S) \rightarrow (\text{PP} \rightarrow \text{Th}[\text{PP}] \rightarrow S): \lambda H. \lambda t. \lambda t'. \text{more'}(H)(t)(t')\]

The(1): \[\vdash \lambda u. \lambda v. \lambda r'. \lambda s'. \text{THE}+u(\text{iso})(e)(e)+v(r'(e)) + s': \]
\[((\text{Deg} \rightarrow \text{PP} \rightarrow S) \rightarrow (\text{PP} \rightarrow \text{Th}[\text{PP}] \rightarrow S)) \rightarrow (((\text{PP} \rightarrow S) \rightarrow S) \rightarrow ((\text{Deg} \rightarrow \text{PP} \rightarrow S) \rightarrow (\text{CC} \rightarrow S))): \lambda \Phi. \lambda \Phi. \lambda G. \lambda R. \text{CORREL}(\lambda 1. \lambda 2. P(\text{Eq}(1)) \wedge P(\text{Eq}(2)) \wedge \Phi(G)(1)(2))(R)\]

The biggest difference between this derivation and the ones we saw for the CC and RCC with quantification over individuals relates to the difference we already saw in (69).
Namely, there is an extra argument of the verb within the double abstract. Otherwise, things are very similar. The same arguments in the same order apply, with PP and i filled in for NP and e. As you can see in the specification for *on an occasion*, I deemed the expression to have no content apart from predicing that whatever property it takes as argument occurred at time $t$ and that such a time exists (here, this time or occasion could be a single unit, an interval, etc. and it would make no difference to the analysis).

One of the nicest aspects of this analysis is that the phenogrammatical component needs no change at all despite the fact that *on an occasion* occurs on the opposite side of the verb from the subject, which was the position of the pseudo-associate in the earlier example. The way this system is set up, however, lexical entries like *more* and *the* are designed to work with the specification of the double abstract. Regardless of where the non-degree gap occurs in the double abstract, it will always be a function from strings to strings to strings. Because of that, the pheno terms of the other lexical entries can specify simply that their pronunciations should be taken as argument by the function that is the pheno term of the double abstract, and that pronunciation will automatically appear in the right position regardless of which side of the verb, etc. that happens to be on.

The final derivation shows the formation of the second CC clause for (62), where we see the same thing as in (70) and parallels to the matrix clause of the earlier CC. We also see that the same relationship exists between the temporal indefinite in the subordinate clause and the temporal pronoun *then* in the matrix clause as the relationship between *a dog* and *it*. Specifically, we see that *then*, like *it* includes an extra equality between the bound variable and a free variable, creating a familiarity presupposition.
These past three proofs that, when put together, show the complete analysis for (62), demonstrate the flexibility of this analysis of CCs. One final note includes extension to examples in which more than one indefinite forms part of the restriction of quantification. For example, in \textit{the faster a dog chases a cat, the more likely it is to catch it}, the pairs that form each set argument of CORREL are not merely pairs of individual dogs, they are pairs of pairs of a dog and a cat. I will not go into this kind of example in any detail, but the system that I have described here could be extended to such cases with a minimum of additional assumptions. If we assume that it is possible to
have a product type as the type of a single gap along with projection functions that can
isolate the various components of that product type, then it would be possible for the
phenogrammatical component to have the first and second projections of that product
type (here, *a dog* and *a cat*, respectively) quantified into different places in order to
derive this kind of example in which the multiple indefinites are non-contiguous. Once
again, this is an advantage over a system like Beck's in which each additional indefinite
would require a new lexical item for *the*, or, in fact, three, since there would be one for
individuals, one for times, and one for worlds.

### 4.4 Conclusion

This chapter has described, in a fully-specified formal system that makes concrete
predictions, a new analysis of CCs and RCCs. I described a new framework (PTDCG),
situated it with respect to other frameworks that handle syntax, semantics and the syntax-
semantics interface, and showed how it would handle a variety of standard (quantifier
scope, long-distance dependency) and not-so-standard (parasitic scope) phenomena.
This analysis demonstrated the possibility of giving a compositional analysis for the CC
that is largely composed with the same tools used elsewhere in natural language. Finally,
I demonstrated the flexibility of the analysis in handling a variety of kinds of CCs, both
in terms of extension of the analysis to the RCC and in terms of its extension to CCs in
which quantification occurs over a non-individual type.
Chapter 5: Discussion and Conclusion

In this chapter, I return to the benchmarks from chapters 2 and 3 and determine which of these the analysis developed in chapter 4 accounts for. I also discuss further issues for future work before concluding. Not surprisingly, we will see that the theory proposed in this dissertation accounts for more benchmarks than any previous theory, whether syntactic or semantic, and additionally is the only theory to account for some of the benchmarks.

5.1 Benchmarks Redux

I begin this section with the most basic set of benchmarks that describe simple aspects of the form of CCs:

(B2) *CC as S:* CCs are of the same syntactic category as other sentences of English.
(B3) *CC clauses don’t stand alone:* whenever one *the more* constituent occurs, a second *the more* constituent must follow to form a sentence.
(B4) *Mandatory The:* *The* is required phrase-initially in both main CC constituents.
(B6) *No the in the 1st clause of the RCC:* In the RCC, no *the* can appear initially in the first main constituent.
(B11) *Obligatory Comparative:* Each CC clause requires a comparative form.
(B10) *Comparative Variety in the CC:* Any form of the comparative can occur in the CC.

Each of these is met by the proposed theory within PTDCG. Because the tecto type of *the1*, which appears in both CCs and RCCs, ends with CC→S, it is necessarily the case that both CCs and RCCs as a whole will be sentences, satisfying (B2: CC as S). Also, since subordinate CC clauses build to CC→S and matrix CC and RCC clauses build to
CC, neither of these are simple sentence types that would be able to stand alone, satisfying (B3: CC clauses don't stand alone). In the theory proposed in this dissertation, the lexical entry for the *the* that begins each clause does much of the heavy lifting in the overall analysis of CCs and RCCs. Each *the* is the primary functor in its clause, taking other elements as arguments to return a CC or CC- S clause. Thus, without *the*, it is not possible to create a CC, in keeping with (B4: Mandatory *the*). As noted in the discussion of the analysis of the RCC example in the last chapter, the *the* that begins an RCC clause does not contribute a pronunciation of its own in the phenogrammatical term; it merely arranges the pheno terms of its arguments, sometimes specifying particular arguments or abstractions, but never contributing any string other than the empty string. For this reason, an RCC clause will never contain a pronunciation of *the*, which is what (B6) requires. As for the last pair of benchmarks in this set, a comparative morpheme is required as the first argument of *the*, so since *the* exists in some form in each clause, a comparative morpheme will also always occur (in order for the sentence to be well-formed). Though I restricted the demonstration of how my analysis works to examples with *more*, I assume that *-er* and *less* are also specified as having the tecto type \(((\text{Deg} \rightarrow X \rightarrow S) \rightarrow (X \rightarrow \text{Th}[X] \rightarrow S))\), though in the case of *-er*, the pheno term may be different since *-er* will appear on the opposite side of the adjective, adverb, etc. from *more or less*, which I will not worry about here. Thus, since the proof is driven by the combination of tecto types, each of the three lexical entries for *the* in this analysis will require something of type \(((\text{Deg} \rightarrow X \rightarrow S) \rightarrow (X \rightarrow \text{Th}[X] \rightarrow S))\), meaning that a comparative morpheme will appear but that which one appears is indeed flexible, meeting both the (B10: Comparative Variety in the CC) and (B11: Obligatory Comparative) benchmarks.
The next pair of benchmarks are related to the existence of the comparative in the CC, and these are the ones that Beck 1997 best accounts for:

(B17) Than-phrase Restriction: CCs cannot occur with than phrases in either clause.
(B20) No MP in the CC: Overt measure phrases are not possible in (R)CC clauses.

As we saw in more detail in chapter 3, there are actually some exceptions to (B17), but as I argued there, these than phrases are on par with other places in English where we find a than phrase not licensed by a comparative morpheme (or rather, a second than phrase licensed by a comparative morpheme). Thus, the than phrases that are possible for many speakers are not the same as the ones that would (but can't) occur in a CC or RCC clause. These are predicted not to occur in my analysis in basically the same way as in Beck's: both the associate and remnant positions in the comparative are left as variables that are quantified over to create the monotonic correlation that characterizes the CC's meaning. Thus, a than phrase would eliminate the remnant variable and not derive a CC. As for (B20: No MP in the CC), the meaning for the comparative morpheme upon which I built my analysis was the meaning of the comparative in an example of a simple phrasal comparative without a measure phrase. Assuming that a measure phrase is a further argument of the comparative, then, my analysis does explicitly block this possibility. If a measure phrase were instead deemed to be a modifier of more, etc. which took something of more's tecto type to return something of the same type, then I would not be able to explain their absence. Beck accounts for (B20: No MP in the CC) by assuming that the is like how much and occupies the position of the would-be degree difference in the CC, thus preventing a measure phrase from occurring there. From my perspective, this would
be an interesting analysis for a language in which how much and that much are used in place of the1 and the2 (and many such languages exist), but since these do not occur in English CCs and because English CCs differ in many ways from how much correlative clauses in English, this seems to me to be a cross-linguistic assimilation that leads to problems. Instead, I prefer to posit that languages’ lexical items and their combinations in certain constructions may vary in small ways, such as having the English CC built on the basis of the comparative without a differential argument while another language may have a similar construction built instead on the comparative with a differential argument. Brasoveanu, for Romanian, argues for the necessity of an analysis that includes a differential argument that is not filled by an item parallel to the on the basis of the fact that you need to be able to make reference to these differentials in order to equate them in non-conditional CCs. He also points out that explicit measure phrases are possible in Romanian despite the fact that its morphology introducing each clause is clearly an example of a how much/that much system. English, however, does not ever require co-reference or a proportional relationship between values in its clauses, so this is a difference between these languages that should be able to be explained on the basis of the difference in lexical items (and which ones are used in the CC) in the two languages. For English, then, I believe that the comparative without the differential is the right choice, which differentiates my analysis from Beck’s or Brasoveanu’s and explains (B20: No MP in the CC). I should also point out that the theory developed here can be very easily extended to comparatives with a differential argument in order to analyze languages such as Romanian, so, in a sense, the general theory developed here can be extended cross-linguistically, but not the specifics, nor would we want them to be given
the difference in behavior of CCs cross-linguistically.

The next set of benchmarks has to do with the behavior of long-distance dependency in the CC and RCC:

(B1) **CC as UDC**: Each CC clause contains an unbounded dependency and is therefore subject to island constraints such as the Complex Noun Phrase Constraint, the Empty Category Principle, and the Strong and Weak *wh*-island constraints.

(B1.1) **Major Category Connectivity**: The syntactic category of the filler and gap must match.

(B1.2) **Number Connectivity**: The number of the filler and gap must match.

(B1.3) **Idiom chunk Connectivity**: idiom restrictions from the gap constrain the identity of the filler.

(B5) **Mandatory Fronting**: the comparative (and whatever part of the filler it brings with it) cannot appear in situ in either CC main constituent.

(B18) **Lack of RCC 1st clause fronting**: The comparative is not fronted in the first main constituent of the RCC.

The combination of (B1) and (B5) require that (minimally) *the more* appear clause-initially, forming a long-distance dependency with a gap in a clause. In the CC, *the* will always appear initially with *more* following it based on the pheno term for *the*, which puts the string THE first, followed by the isolated term MORE (or LESS, etc.). Therefore, they will never appear in situ, satisfying (B5: Mandatory Fronting). There will also always be a unbounded long-distance dependency between *more*, etc. and the degree gap in a subsequent clause insofar as we saw in the last chapter that the same analysis used for other UDCs in PTDCG is extended to the analysis of the CC. This included a part of the lexical specification for the movement trigger (*that, what, the1, the2*, etc.) such that it feeds the gapped clause an empty string (and the fronted element as well, if applicable) and makes the pronunciation of the fronted string precede the pronunciation of the gapped clause. This meets (B1: CC as UDC). Then (B18: Yes-No
Questions) requires that the same not be true of the matrix RCC clause, which, as we saw in chapter 4, it is not in this theory. The only difference between the2 and the3 was their pheno terms, which is the component that controls fronting, and the primary difference (besides the presence of THE) between these was whether fronting had occurred. Thus, (B18) is met as well. As for (B1.1)-(B1.3), connectivity is achieved in the following way. The type for more and the type for the contain variables that range over a variety of tecto categories. The thing that determines what the category will be for the comparative morpheme in a simple comparative sentence is the double abstract that it takes as argument; its non-degree gap must be of the same category as all of the variables in the tecto type of more. The same is true of the in the CC. As we know, it takes more, then the pseudo-associate and then the double abstract. The category of the pseudo-associate is a type-raised version of the same category of the non-degree gap in the double abstract. 

*The* could be ((Deg¬PP¬S)¬((PP¬S)¬S))¬(((PP¬S)¬S)¬((Deg¬PP¬S)¬CC)), as in (4.71) (example (71) in the last chapter). Here, the category PP is filled in for each instance of the variable X in the tecto type of the. Beginning from the right edge, we see that the PP category of the double abstract must match the non-S category of the pseudo associate as well as the three instances of the same category in the type for more. This illustrates the way in which type connectivity is guaranteed for (B1.1), and the same can be done for number, satisfying (B1.2). Idiom connectivity in (B1.3) is more complicated and depends upon a theory of idioms, which I have not developed at this time.

Another benchmark related to long-distance dependency is (B9: *That* in the CC). Just as in relative clause formation, where a *that* appears between a fronted element and the clause that (at some level of embedding) contains its gap, *that* in the CC appears
between fronted more (and whatever it pied-pipes) and the double-abstract that contains the degree gap that it would quantify into if it were in situ. Normally, that would be the lexical item that caused it not to occur in situ, which is not the case here since that is the job of the. However, since the is clause-initial, it does not intervene before the fronted constituent and its gap, which leads to the potential for ambiguity, and in fact, we do find ambiguous examples such as (1a) with each of its readings disambiguated using that in (1b) and (1c).

(B9) That in the CC: That can optionally appear before the gapped clause in a CC clause.

(1) a. The more people come, the better the profits will be.
   b. The more people that come, the better the profits will be.
   c. The more that people come, the better the profits will be.

If (1a) is in a context where the speaker is a symphony director and it is true both that more people on a given night and more repeat customers will help fund the symphony, then either of these readings should be felicitous. It seems like that is valuable in making clear which reading is intended. It is true that wh free relative clauses have the movement trigger as the initial item but do not allow an additional that, but these can never be ambiguous in the same way as the CC because they do not have the same flexibility of the more in terms of what tecto type can follow it, e.g. how many people v. how often people came. I propose that in all cases like (1c), there is something understood to be the noun, adjective, etc. that the that relative modifies. In (1c), it would be the word often, as in the more often that people come,... but it could also be something like the more (money) that an attorney makes,... If this is true, then given the fact that
CCs can occur without a main verb phrase, as in *the more people, the more problems*, we can treat any example with *that* as an instance of a true movement-trigger *that* in the clause. CCs with *that* are then cases of a relative clause within a CC clause that is similar to a relative clause in terms of its required long-distance dependency. A treatment along these lines is something that this theory would easily accommodate.

A final set of benchmarks directly dealing with long-distance dependency is the set related to negation in the CC and RCC:

(B13) *Negating a CC*: The only way to negate the sentence-level meaning of the CC is by saying *It is not the case that X* where X is a CC.
(B14) *Negating an RCC*: An RCC's sentence-level meaning can be negated in the same way as the CC's or by negating the main verb in its first (matrix) clause.
(B35) *Narrow scope negation in CCs*: CC clauses can contain narrow scope negation where it is felicitous with respect to the comparative.

The reason that I claim that these are related to the UDC status of *the1- and the2-initial clauses is because *wh*-relatives and other UDCs also distinguish between negation within a gapped clause whose gapped constituent has been fronted and sentential negation. For example, *whenever he didn't go, I went* and *it's not the case that whenever he went, I went* are not truth-conditionally equivalent. In a model where there are some times when the two people in question overlapped in being at the place in question, the first is false, but the second is true as long as there are also times when he was there but she wasn't. Similarly, what (B13: Negating a CC) and (B35: Narrow scope negation in CCs) are saying in combination is that there can be clause-internal or clause-external negation in a CC, but clause-internal negation is not ambiguous between narrow and wide (sentential) scope; it can only have narrow scope, as with this example of the *wh* free relative clause.
Thus, even though the analysis presented here does not include a specific detailed analysis of negation, if its behavior is conditioned on the basis of fronting behavior in a clause and this analysis handles fronting in the CC in an analogous way to the way that fronting is handled in *wh*-relatives (which we know it is), then the analysis should be able to account for these benchmarks. Similarly, because there is no fronting in the matrix RCC clause, clause-internal negation there should be able to take wide or narrow scope, just as in other non-UDCs, and we know that it does, as evidenced by (B14: Negating an RCC). The theory here would account for that as well given any theory of negation that handles the UDC/non-UDC difference without respect to the CC.

The next set of benchmarks relates to the CC's status as a subordinate-matrix construction. Analyses of all of these phenomena (tag questions, imperatives, etc.) are significant in their own right and beyond the scope of this thesis, but the point here is that all of these are general properties of subordinate-matrix clause constructions, of which the CC and RCC are but two examples. In the analysis proposed in the last chapter, the difference between the1 on the one hand and the2 or the3 on the other creates an asymmetry between the two CC or RCC clauses with respect to each of the three components. Thus, the subordinate-matrix relationship of the CC is clear and similar to the way that other subordinate-matrix clauses would be modeled in PTDCG. Thus, any theory of any of these phenomena that was in place for other analyses of subordinate-matrix clause sentences could be extended to this theory of CCs and RCCs.

(B8) *Subordinate Gapping*: The CC, like other subordinate-matrix structures, does not allow gapping between its clauses.

(B15) *Yes-No Questions*: Only the first clause of the RCC can be used to form *yes-no* questions.
Tag Questions: Tag questions can be formed on the basis of the second clause of a CC or first clause of an RCC but not their other clauses.

Subjunctive: Both CCs and RCCs are possible in subjunctive contexts, but the verb in the clause corresponding to what is semantically the consequent is the only one that receives subjunctive morphology reflecting this fact.

Imperatives: The clauses in the CC and RCC that correspond to semantic antecedents can take on an imperative verb form when in imperative contexts, while their other clause cannot.

Acc-Gerundives: The clauses in the CC and RCC that correspond to semantic antecedents can take on an accusative-gerundive form, while their other clause cannot.

The following pair of benchmarks, however, are slightly different from what you find in other subordinate-matrix constructions. I argue that this is only because many subordinate-matrix constructions do not also involve fronting, especially in their matrix clause (which, as we saw above, is the only one that can be used to form imperatives, etc. in the first place). In other examples of long-distance dependency, completely independently of subordination, we see the same pattern as in the matrix CC clause: an imperative is okay with a subject but not without, as in (2). Once again, because there is no fronting in the matrix of the RCC, it would be expected to show a different pattern, which it does.

Subjects with Imperatives: In order for a CC clause to take on an imperative form, it must have a subject.

Subjects with Acc-Gerundives: In order for a CC clause to take on an accusative-gerundive form, it must have a subject.

(2) a. I suggest that everyone keep your mouth shut about this, okay?
   b. *I suggest that keep your mouth shut about this, okay?

The constraints on imperatives and accusative-gerundive forms in the CC are a great example of what I have been trying to illustrate throughout this thesis. Here we see an
interaction of two regular natural language processes (UDCs and subordination) with respect to a single phenomenon (e.g. imperatives). Neither of the patterns we see here (requiring a subject, only being formed in a matrix clause) are surprising, but when they are combined in a single construction, the pattern may look odd when not broken down into its constituent parts. I hope to have shown, then, that because the CC is treated here as an amalgam of all of these processes, the amalgam of the constraints that apply to these processes will also apply to the CC. In other words, just as the CC is a combination of two UDCs and a subordinate-matrix construction, so too is its pattern with respect to a given phenomenon like imperative formation a combination of their patterns with respect to that phenomenon, where the CC shows both restrictions. Thus, this theory is uniquely set up so that any analysis imperatives, etc. that accounts for their formation in UDCs and subordinate-matrix constructions will be able to be extended to the CC, just as their basic analyses were extended to the analysis of the CC proposed in chapter 4.

The following benchmarks are related to UDCs in which an entire CC is part of the gapped sentence/embedded clause. As discussed in chapter 2, this flexible extraction pattern does in fact match the pattern found in the conditional but potentially differs from other subordinate-matrix constructions such as because X, Y. We saw there that this pattern also differs from clauses that are conjoined and do not have a conditional-like (cause-effect, etc.) meaning.

(B22) *C1 Extraction*: Extraction can occur from the first clause of the CC.

(B23) *C2 Extraction*: Extraction can occur from the second clause of the CC.

(B24) *C1 & C2 Extraction*: Extraction can occur from both clauses of the CC at once.
Given that this pattern differs within the class of subordinate-matrix constructions, it is reasonable to think that the flexibility related to extraction (or lack thereof) should not be attributed to the way that property is manifest in PTDCG. Instead, I suggest that the logical target for such differences is the specification of the words that are specific to each construction. In other words, I hypothesize that it is something about the entry for *because* as compared to the entries for *if* or *the* that make extractability from its clause less acceptable than extractability from theirs. The analysis developed here did not add any specific constraints of this kind to the meaning of any of the entries for *the*, making extraction from any clause possible. Also, *the* as a movement trigger was based on the analysis of a relative pronoun as a movement trigger, and those pronouns show the same kind of unboundedness (where extraction is possible from a clause that already contains a long-distance dependency itself). This is meant to demonstrate that the fronting behavior of *the* also does not prohibit extraction from either clause.

Because this is such an important issue in previous syntactic literature, I give a rough idea of the treatment of a sentence with extraction from a CC here. If the sentence we're interested in analyzing is (3a), the CC with the fronted element in situ is as in (3b). This is a case of quantification over occasions, so we already have traces in the formation of *you meet...* and *you like...* for times/occasions and degrees. In order to form (3a), there will additionally be an individual-type trace in the position of the object of *meet/like*. Because the *you meet...* clause must be a double abstract of the right kind in order to combine with *the more...*, only two of the three traces will be withdrawn from the context by the point of that combination. These will be the time and degree traces, leaving the
individual-type trace in the context at this stage. Then, after the would-be last step of the derivation of the CC, it will be a sentence with a hypothetical element still in its context. At this point, there is another step of hypothetical proof that will make the CC an NP→S instead of an S. This is then the right kind of argument for a relative pronoun like who, and the rest of the derivation proceeds just like any other relative-clause derivation. It is easy to see that this analysis works whether there is a single gap in either the first or second clause or whether, as in (3a), there are gaps in both clauses. In the first two cases, you have a single gap in the context that is discharged and then associated with the fronted element. The case of multiple gaps is trickier, only because it requires additional coreference between the gaps in each clause so that only one element is withdrawn rather than two distinct ones. Chapter 4 has shown the way in which I handle anaphora in this static system, but this is another case in which generalization to a dynamic system would also work (and presumably in a more precise and explanatory way).

(3) a. Jo is the person who the more you meet, the more you like.
   b. The more you meet that person/Jo, the more you like that person/Jo.

(4) a. The more a dog eats, the more it drinks.
   b. *Fido is a dog that the more eats, the more drinks.

One of the positive attributes to the analysis I have just sketched, in addition to the fact that it accounts for (B22)-(B24) (the possibility of extraction from either or both CC clauses) is that it also prohibits certain cases of extraction; namely, those in which the pseudo-associate would be extracted, as in (4a) v. (4b). I think that the reason we find this pattern is because, when something is extracted, it becomes specific. Here, the singular a dog with a restrictive relative clause means that there must be a single salient
dog with the property in the relative clause. However, this contrasts with the function of
the pseudo-associate to introduce a restriction on the variables that are quantified over to
create sets of pairs. The way that my analysis blocks the extraction of the pseudo-
associate is by making it a separate argument of *the*, or, more specifically, not having it
combine with the clause that becomes a double abstract before it becomes a double
abstract. The derivation of a CC will fail if there is no double abstract, e.g. if it were
only a single abstract or a triple abstract, and it will fail if there is not a quantifier of a
tecto and semantic type that matches the tecto/semantic type of the gap in the abstract
that is not a degree gap. Thus, if the pseudo-associate combines with the main verb
before the creation of the double abstract, it will not be abstracted over, and it will not be
available later to be the pseudo-associate argument, which will not lead to a well-formed
sentence. This analysis also correctly predicts that in a CC that is ambiguous between
various readings (where more than one thing could be understood as the pseudo-
associate), which potential pseudo-associate is extracted will disambiguate the sentence
such that only the pseudo-associate that is not extracted can actually serve in that role.
We see in (5) that (5a) could technically mean that for all/most (however many are
relevant for establishing a correlation) pairs of people where you meet one more often
than the other, you like that one better than the other. It could also mean that for
all/most/etc. pairs of times where you are meeting/seeing an individual and where you
will have met him/her more often at one time than another, you will like him/her better at
that time. Example (5b) has only the latter reading, as predicted by my analysis.

(5) a. The more you meet a person, the more you like him/her.
   b. Jo is a person who the more you meet, the more you like.
The final benchmark related to extraction (though not particularly related to those above) is (B29):

(B29) Preposition Pied-Piping: The CC does not permit prepositions to be extracted along with the more.

My analysis captures this benchmark via the phenogrammatical specification for the. In the other constructions we saw that do permit prepositional pied-piping (so much that, how much, etc), the morphology that characterizes them does not have to be clause initial. Because the is necessarily clause-initial (which is part of what this benchmark shows), the pheno term for the does not have any argument slots to the left of THE. This effectively prevents anything, including a preposition, from appearing prior to the.

The remaining benchmarks are all largely semantic. I begin with those that are semantic reflexes of the subordinate-matrix relationship in the CC and RCC.

(B30) Future tense: CCs and RCCs do not occur with future will in their antecedents. (B31) NPIs in the CC: Negative Polarity Items can occur in the antecedent but not consequent clause of the CC and RCC.

Just as with theories of tag questions, etc., theories of futurity and polarity are beyond the scope of this dissertation. Nevertheless, as before, there is no reason to think that my analysis would not be able to capture these patterns given any working theory of tense and polarity. In other words, even though I am arguing that CCs are not conditionals, my analysis still treats them as subordinate-matrix constructions with a generalized quantifier, which is what is relevant for these benchmarks. In other words, not all subordinate-matrix constructions show the pattern in (B30: Future Tense) and (B31: NPIs
in the CC) (e.g. because sentences), and not all sentences with GQs show this pattern (given the words antecedent and consequent, we know that it is not meant to apply to cases of conjunction, etc.), but as chapter 3 showed, they are general to those with both. Thus, it is not the hypotheticality of the conditional (or any other property unique to it) that causes the patterns in (B30: Future Tense) and (B31: NPIs in the CC). Instead, though I do not have a precise theory of these phenomena to work with, whatever it is must extend to any GQ-headed subordinate-matrix construction on the basis of these properties, both of which my analysis possesses. That means that it is possible for my non-conditional analysis to account for them. The same is true of the next pair of benchmarks.

(B33) Donkey anaphora: The CC and RCC license donkey anaphora between their clauses.
(B7) Backward Pronominalization: Pronominalization is possible in either direction in the CC (between clauses).

Co-reference in backward pronominalization or cataphora is licensed in English only when the pronoun the precedes the referent is in a left-subordinated clause (Cann and McPherson 1999). This is why, as the data before (B7: Backward Pronominalization) showed, the RCC or reversed conditional cannot license cataphora the way that a CC or if-initial conditional can. I've said before that the CC and RCC are examples of subordinate-matrix constructions, and that the first clause of the CC and second clause of the RCC are the subordinate clauses, but it's important to define what I mean by that in order to see whether this analysis truly accounts for (B7: Backward Pronominalization), among others. In PTDCG, I consider a clause subordinate if the following are met: (1) it
is clausal (there is a finite verb), (2) its lexical specification is such that it subcategorizes
for another clause (its tecto term requires something with a finite verb), (3) it is not of
tecto type $S$ without this clausal argument, but it is of type $S$ with it, and (4), these
clauses are joined by material implication or comparable (the subset relation) in the
semantic term that corresponds to the sentence-level meaning. This definition
circumscribes the set of constructions usually called 'subordinate' and rules out those that
are normally called 'coordinate' or 'paratactic.' We also see that the CC and RCC clauses
that are CC-$\triangleright$S meet these requirements. The last requirement is met by the meaning of
CORREL, which requires that some number of pairs in the first set also be in the second.
Thus, any constraint that depends upon subordination depends upon these properties in a
PTDCG analysis and will be able to be integrated into this analysis of CCs and RCCs.
Donkey anaphora is not limited to subordinate clauses, but it is limited to those clauses in
which there is a material implication co-occurring with a non-existential quantifier in
addition to the indefinite and pronoun. The analysis from chapter 4 of an example with
donkey anaphora and its subsequent explanation show how this analysis can and does
handle anaphora as per (B33: Donkey Anaphora).

Turning now to the benchmark in (B32: Quantificational Variability), I explicitly
demonstrated the way in which my analysis generalized to quantification over pairs of an
even wider variety of types than past analyses in chapter 4. This was handled by
lexically specifying that the and more had a tecto type with a variable in it that ranged
over a variety of types. The way the analysis was set up to allow traces of various types
depending upon the verb's arguments and for the pseudo-associate to be a quantifier of
any type that was a trace used to form a double abstract additionally assured the type-
flexibility in this analysis without requiring different lexical entries, different grammatical rules, different interface specifications or anything else, which is a major advantage of this kind of approach over an analysis in which each new type requires a set of new lexical entries for the.

(B32) *Quantificational Variability*: CCs and RCCs have meanings that can correlate degrees with respect to a number of different natural language types, including worlds, times and individuals.

The next benchmark is intriguing, since it is somewhat typical of idiomatic expressions, but less common among others, and not well-studied (other than the difference between phrasal comparatives and comparative deletion, which, as previous authors showed, is different from the pattern in CCs). My analysis is a move in the direction of actually explaining why this is true of the CC.

(B12) *Copula Omission*: a conjugated form of *to be* may be omitted at the end of a CC clause as long as the subject it would normally follow is not specific.

Now, technically, my analysis as it currently stands would predict the ungrammaticality of any sentence with an omitted copula because the analysis requires a double abstract, which in turn requires a finite verb. However, if the last argument of the were made optional, with its pheno place fed the null string, its tecto type eliminated from the subcategorization requirement of the, and an identity function of sorts used in the semantic term, a CC could still be derived under certain conditions, such as the non-specificity of the subject, so (B12: Copula Omission) would be the predicted result. This may seem like an ad hoc move in that it requires a change to the system to account for
one benchmark, but the striking thing is that the system is configured in such a way as to prove a CC with minimal modification as long as it is the double abstract as a whole that is missing. So, for example, it would not be possible to do this if the pseudo-associate were missing. This also means that the noun phrase that remains in the CC when the copula is deleted cannot be a subject in the double abstract that is not the pseudo-associate; it must be the pseudo-associate. And it is for that reason that it cannot be specific. Clearly, the pseudo-associate must contain an unbound variable in order to be equated with one of the lambda-bound variables that create the pair of things to be quantified over. That would not be possible with a specific referent. This sheds light on why a restriction on the omission of a copula would be related to the specificity of its would-be subject. Previous semantic analyses such as Beck's, because they fold the pseudo-associate into the double abstract (to characterize it in my terms) before the takes it as argument, cannot distinguish between the pseudo-associate and other noun phrases, whether they are specific or not. (B12: Copula Omission) is a real puzzle in that kind of setup, unlike here. Another note is that my analysis makes a finer-grained prediction than (B12) that seems to be borne out. If my analysis is right that what is crucial is that the noun phrase be the pseudo-associate (rather than just non-specific), then one would expect that in an example with is that is ambiguous between readings, that ambiguity would go away without is since the quantification would necessarily be over the only remaining NP. Take the example of (6a). As noted before, this is Beck's example of quantification over possible worlds, and it does have a reading in which pairs of possibilities are compared for the same attorney, but it can also compare attorneys. Example (6b), without the copula, can only be understood as comparing pairs of
attorneys, however, which supports my hypothesis.

(6) a. The slimier an attorney is, the more efficient s/he is.
    b. The slimier an attorney, the more efficient...

The next benchmark is one that my analysis cannot account for. In my theory, the RCC matrix clause only combines with the subordinate clause because it is of tecto type CC, and the only reason it is of tecto type CC is because it begins with a silent the. Other possible clauses that we find substituted for a matrix RCC clause are things like as he grows or his weight increases. I have no reason to believe that these are CC clauses, but there is no question that they do combine with subordinate CC clauses to form the same meanings as RCC sentences.

(B19) **RCC Flexibility:** The first clause of the RCC can take any syntactic form as long as it corresponds to the right kind of meaning (semantic term) to be combined with its second clause, whose syntactic form is fixed/does not show similar flexibility.

Basically, then, this benchmark suggests that it should be the semantic type and not the tecto type that restricts what kind of clauses combine in the RCC, unlike in the CC. However, this theory is designed such that the tecto types alone determine what combines, so it would have trouble explaining this kind of difference.

The next benchmark has been influential in the decision to treat CCs as conditionals, and the two that follow it are benchmarks that are noted for the first time in this thesis and are accounted for only by the present analysis.

(B34) **Overt AQs:** The CC can appear with an overt Adverb of Quantification.
(B37) **Indefeasible Force:** The correlational force of the CC cannot be altered by an AQ.
Though (B34) is a true statement, it was previously believed that any AQ would be possible with the CC. As we saw in the third chapter, that is not the case, and the analysis offered there showed that when an AQ occurs that would seem to conflict with the quantificational force inherent to the CC, that AQ is actually quantifying over a different domain than the domain of the the quantifier intrinsic to the CC. This relates to the final benchmark insofar as its explanation requires that the CC have a recognizable quantificational force to begin with.

(B36) **Quantificational Force in the CC**: The (unmodified) CC's quantificational force is neither universal nor generic but rather patterns with *most* in terms of native speaker judgements of truth in proportional contexts.

And indeed, as we saw in chapter 3, CCs and RCCs are judged to give rise to proportional readings without an overt proportional quantifier, which is unlike conditionals that only give rise to those readings when overt proportional quantifiers are present. That, combined with the fact that this proportional force in the CC is not eliminated when an overt AQ is present (those that are attested as co-occurring with the CC in the first place) made the case against treating CCs as conditionals. Thus, a proportional force in the form of CORREL was posited rather than the universal quantifier common to all other theories, and it was additionally posited as part of the lexical meaning of *the* rather than as a silent operator that was not directly a part of anything in the CC. So, with respect to (B36: Quantificational Force in the CC) and (B37: Indefeasible Force), the theory proposed in these pages is the only theory of CCs to
this point that can account for them.

5.2 Future Research

As indicated at various points throughout this dissertation, there are a number of aspects of the CC or constructions that share some aspect of meaning or form with the CC that remain to be investigated. Below, I divide these into three basic categories: aspects of the English CC that require further investigation, English sentences with meanings nearly identical to the CC’s meaning, and constructions in other languages with correlational meanings.

First, there are all of the extensions of the theory that I mentioned in previous chapters and above in the evaluation of benchmarks that require theories of other phenomena to be integrated into this analysis. These include theories of the morphologically-incorporated comparative (-er), nominal and other non-attributive forms of the comparative, extraposition, tag questions, etc. A deeper investigation of definites and indefinites in the CC and their relationship to clausality is also needed, as we saw in Chapter 3 with the presentation of data like (7) v. (8), where (8b) differs from all the other examples in not quantifying over different attorneys.

(7) a. The slimier an attorney, the more efficient.
    b. The slimier the attorney, the more efficient.

(8) a. The slimier an attorney is, the more efficient he is.
    b. The slimier the attorney is, the more efficient he is.

This is but one facet of domain restriction in the CC that requires further investigation. The analysis presented here permits any kind of variable to be quantified over as long as
it matches the kind of non-degree gap in the double abstract. If we assume that parts of the index (time or event, situation or possible world) are possible free variables in a double abstract, then there will be multiple possible interpretations for the CC in terms of quantification. As we have seen, multiple interpretations are often possible depending upon the context of utterance and other aspects of the CC itself, so I consider it an advantage of this approach that it does not attribute to the semantic composition something that is affected by other, pragmatic, factors. Nevertheless, a better theory would make clear exactly what those factors are and how they influence the selection of which of the possible variables is bound to create the double abstract as opposed to being fixed to the current time, world, etc. For example, consider (9).

(9) Context: A new special education teacher is describing her class, and you know he's been having discipline problems. He says:
   a. The angrier they are, the less well-behaved they are.
   b. The angrier they are, the less well-behaved they get.
   c. The angrier they get, the less well-behaved they get.

Here, get biases quantification over stages of an individual in (9b,c) rather than just the individuals, which is the default interpretation in this context for (9a). Comparing (9a) and (9b) is especially noteworthy given that their discourse contexts and first the clauses are identical, so the only difference is are v. get in the second clause. This means that factors outside the previous discourse and restrictor clause can restrict the domain of a CC. Now, this could be a semantic requirement of get, i.e. one that is part of the restriction that arises from its semantic specification in the lexicon, a presupposition, or an implicature of some kind. Following Roberts 1995 and 1996, inter alia, anaphora and
domain restriction can be treated presuppositionally. In this theory, felicitous interpretations of sentences (relative to a particular domain) are constrained by the presuppositions of lexical items and operators in that sentence (and surrounding discourse). The way this constraint works is roughly that if the accommodation of any of these presuppositions would contradict something in the common ground between a speaker and hearer, it cannot be entered, prohibiting that reading. If the accommodation of all these presuppositions is licit, then so is the interpretation of the sentence in which all of the accommodating restrictions has been made. This seems to work for most of the CC readings in context that I have tested, but there is much more work to be done in understanding what happens in the case of multiple, possibly competing, presuppositions in terms of their interaction with one another. In particular, theories of tense and modality in other kinds of subordinate-matrix constructions need to be applied to the CC.

There are other subtle differences in interpretation according to my judgement that need to be investigated empirically and, if found to be robust, explained. One such difference arises between the CC and RCC in (10).

(10) a. #? Semanticists' husbands are taller, the more prolific their wives.
   b. The more prolific the semanticist, the taller her husband.

I specifically chose a non-causal possible generalization since the difference of interest to me here has to do with an implication of causality. It seems to me that the RCC in (10a) has a stronger implicature of causation, which is what makes it sound less felicitous than (10b), where the implicature either doesn't exist/gets canceled or is weaker. It could be that the syntactic properties of the RCC v. CC lead to a difference between them in terms
of number (*their v. her*) in the translation of CCs like (10b) to RCCs and that that has to do with this difference, but regardless, this is an issue for further study.

Also related to empirical evidence, I would like to gather more data both via psycholinguistic experimentation and corpus methods to further support some of the evidence presented in this thesis, especially that in chapter 3 with respect to the interpretation of *the* as compared to quantification in bare conditionals and their co-occurrence with various adverbs of quantification.

Another kind of CC not explored elsewhere in the literature is one in which there are more than two *the* clauses along with the possible interpretations of such sentences. In many of these cases, two or more of the clauses are explicitly conjoined, as in (11), but this is not always true, as in the naturally-occurring example in (12).  

(11) a. The more articles you read, and the more papers you write, the more likely you are to get a job.
   b. The more articles you read, the more presentations you give, the more papers you write, the more likely you are to get a job.
   c. The kinder you are, the more friends you will have and the wider your support network.

(12) I should learn not to get my hopes up so high. The farther you fall, the deeper the bruises, the longer it takes to heal. (twitter 5/8/09)

There are a few different possible meanings (one is a special case of another) and structures (coordinated or not) for multiclausal sentences. For example, for 3 clauses X, Y and Z, a multiclausal sentence could assert a correlation between X & Y and Y & Z (of course, thereby correlating X & Z); it could assert that a combination of X & Y are

---

35 Other examples like (iv) *The more presentations the more students make, the more money the department spends on travel* may have a similar semantic effect to the sentences in (11) but not the same structure. Assuming *the* occurs in place of a degree gap, the first clause of (iv) is like the regular sentence (v) *3 more students made 2 more presentations (than last year)*, while (11a)'s is like the sentence (vi) *You read 9 more articles and wrote 2 more papers.*
correlated with Z, or that X is correlated with a combination of Y & Z, etc. Once again, pragmatics and real world knowledge plays a role in determining which interpretation(s) is/are salient. So in (11c), your kindness is correlated with both your number of friends and with your support network individually, while in (11a), it's not clear to me whether both paper writing and article reading are each correlated with job success or whether some (any?) combination of the two are such that, for example, 2 papers written and 3 articles read leads to less job success than 5 papers written and 1 article read, where number of articles read is not by itself correlated with job success. Either way, the analysis of the CC has to be able to accommodate various possible readings and determine how they arise compositionally.

The next broad category of future research pertains to other English sentences with CC-like correlational meanings as well as sentences that would seem to be violations of CC benchmarks that I have found in naturally-occurring speech. The first set includes examples like those in (13)-(15).

(13) As you eat more, you get fatter.
(14) As I build more muscle, I run faster and faster.
(15) If he steals more and more money, he's more and more likely to end up swimming with the fishes.

Already in chapter 3, I described a difference between examples like (13) and the CCs analyzed in this thesis in terms of the restriction to state-level predicates. Further work is needed to determine whether there are other distinctions or distinctions between CCs and other kinds of examples. In particular, this thesis has argued that CCs are not
conditionals despite many similarities, so the comparison with (15) is of interest. Also, the repetition of more in more and more is sure to be a challenge to analyze.

The second set of examples in (16)-(20) demonstrates a range of actual examples that seem to violate certain assumptions from the literature about the CC.

(16) I just figure I need to get a start on days coming after tomorrow. 24 hours at a time isn't gonna get it done, the farther this term progresses. (email 5/19/05)

(17) I'm kind of surprised, the more I think about it, that I've never had it called to my attention before. (email 5/28/09)

(18) The more that TV pundits reduce serious debates to silly arguments, big issues into sound bites, our citizens turn away. (Obama's State of the Union, 2010)

(19) Movie trailers are not something I typically think of as being on television commercials, but the more trailers that appeared during commercial breaks convinced me that movie trailers are a big part of children's advertising. (student's rough draft, winter 2009)

(20) The deeper into the tour the show gets, the more her mother notices how Camille seems to miss friends and relative back home less and less. (The Columbus Dispatch 5/18/2009)

Examples (16)-(19) are all cases in which one clause is clearly a CC clause, with the more... but the other clause is not. Nevertheless, I posit that the most obvious understanding of the meaning of these sentences is one in which a correlation is established, with the possible exception of (19). In (18), for example, Obama seems to be suggesting that increases in punditry lead to decreases in citizens' engagement, despite the fact that the last clause, in isolation, does not contain any elements of gradability, such as a comparative. In (19), on the other hand, I get the sense that the speaker is not necessarily more convinced by seeing more movie trailers, only that there was a tipping point during her viewing where she had seen enough to be convinced. One could say
that these are all examples of speech errors or cases where people get side-tracked while talking, forget that they started with a CC clause and then end with a non-CC clause, etc., but I find this unconvincing. In most of these cases, the intended meaning is clear and a correlation is expressed by another means. In example (20), no clause is missing a gradable element, but one clause has two separate indicators of a set of increases or decreases, both the more and less and less. I'm not sure how to interpret this, but I'm inclined to believe that being with the show longer is correlated with missing friends less and it is this correlation that her mother notices. A compositional derivation from the actual form of (20) might be that the longer the girl is in the show, the more obvious something becomes to her mother, and that something is that she is gradually getting used to not having her friends around. Clearly, though, examples like these need further investigation.

Finally, there is the issue of other languages. First, we have the languages that previous authors have studied, including Mandarin Chinese, German, and Romanian. Then, there are also languages that I began to study and noticed interesting differences from English with respect to CC-like sentences. These are Japanese, Hindi, and Serbo-Croatian. I posit that no single analysis of correlational meanings can be a cross-linguistic universal but that the simple rules of elimination and introduction in PTDCG in combination with the lexical specifications for the words present in a particular language's CCs will be able to derive the range of meanings for CCs in that language. In other words, just as particularities of comparatives, UDCs, etc. in English influence the behavior of English CCs, so too will particularities of phenomena in other languages (that may differ from the way the same phenomena behave in English) influence the
behavior of CCs in those languages. Detailed analyses of each language are required to support this hypothesis, but below are some of the intriguing particulars of Hindi, Serbo-Croatian and Japanese that drew me to them.

I first looked at Hindi because it is the language in which the most has been published on the syntax and semantics of correlative clauses (as in (21)), which have been argued to form the base of the Correlational Comparative cross-linguistically by den Dikken 2005. Remember from chapter 1 that a correlative clause is a kind of relative clause that is differentiated from other kinds of relative clauses on the basis of its syntactic form; it appears at the left edge of the sentence. There is potentially an information-theoretic difference as well in that sentences with correlative clauses have been argued to be Topic-Comment structures (with the correlative clause denoting the topic), c.f. Bittner (2001). The content of the correlative clause is thought to be linked to the matrix clause via a demonstrative. In the case of (21), the first clause jitna...hoon is the correlative clause, which is linked to the demonstrative utna in the main clause. An example of a Correlational Comparative in Hindi is the sentence in (22).

(21) Jitna kaam 'main sangeet par karti hoon, utna hi kaam main bhasha par karti hoon.
howmuch work 1sg music on todo-fem is, thatmuch also work 1s language on todo-fem is
'However much I work on music, I work that much on language'

(22) Jitni umar badhegi, utna gyan badhega.
how-much age increase-fem, that-much wisdom increase-masc
'The older you are, the wiser you are.'

(23) Jitni jaldi khao, utni jaldi pet bharega.
how-much fast eat that-much fast stomach full
'The faster we eat, the sooner we get full.' OR
'The speed at which I eat is the speed at which I get full'
Because Hindi has productive correlative clause formation, there are strings that are ambiguous between a CC interpretation and a regular correlative interpretation, as we see in (23) where two different translations are given. Matching scales (here, speed) are crucial to the correlative interpretation, which is what prevents (22) from having both readings. The possibility of both of these readings is problematic both for analyses of CCs such as den Dikken's and analyses of correlatives, such as Dayal's (1995, 1996) in that neither takes the other into account. Dayal proposes that correlative clauses are interpreted as plural definite descriptions of individuals following Jacobson (1993). Thus, for Dayal, sentences like (21) require *utra* to anaphorically link the plural definite (the denotation of the correlative clause *Jitna...hoon*) to the matrix clause interpretation. This is significantly different from CC meanings based on monotonic correlation, as was discussed with respect to Brasoveanu's analysis in chapter 3. Other interesting aspects of the Hindi data include:

- Frequent lack of a comparative marker (none of the examples here have one), and further, lack of something that hints of comparison in some cases (compare (22) with *increase* to (23) without), reminiscent of the kinds of examples we saw in (16)-(19).
- The possibility of having comparative and equative morphemes together in the same clause with a different meaning than the comparative alone (these data are not included because further empirical verification is required).
- Different possible readings of correlatives tied to generic v. specific tense (a distinction native in the Hindi tense system) and whether mismatches exist between generic tense and the availability of CC-type meanings.

Serbo-Croatian is an interesting case partly because it has two kinds of Correlational Comparatives (as evidenced by different morphological frames) with different interpretations, one of which has meanings that are a subset of the other's
possible meanings. The *sto/to* morpheme frame, as in (24), is more general, while the *koliko/toliko* morpheme frame, as in (25), is more restricted, being true only if the things being measured (here, your work hours and earnings) are *linearly* correlated (c.f. the discussion of how CCs in English are not linear/proportional in chapter 3).

(24) Što više radiš, više i zaradjuješ.
sto more work.2sg, more also earn.2sg
'the more you work, the more you earn'

(25) Koliko radiš, toliko i zaradjuješ.
how much work.2sg, that-much also earn.2sg
'the more you work, the more you earn'

Most of the time, a *koliko/toliko* sentence is less acceptable than its *sto/to* counterpart in being used to form Correlational Comparatives; see the acceptability of (26) as compared to (27), below. The preferred interpretation of *Koliko/toliko* sentences is as Dayal describes (co-reference), whereas *sto/to*'s preferred interpretation is similar to the interpretation of CCs in English. In her work, Dayal analyzes only correlatives that begin with *what* or *which* rather than those that begin with *how much*, but she posits that the theory can be extended. Further research is required to show what that extension would look like and how it fares empirically in both Hindi and Serbo-Croatian (since Dayal's analysis is claimed to apply to all correlatives cross-linguistically).

(26) Što si pametni-ji od svojih prijatelja, to su vam [muzički interesi] različitiji.
sto be smart-er than to-you friend.gen to.be.pres.3pl dat.pron music.nom interests more-diff
'The smarter you are as compared to your friend, the more different your musical interests will be.'
Finally, Japanese is an interesting case partly because it does not require or make use of a comparative morpheme in the CC. Though this is true of other languages, such as Chinese (McCawley, Lin, etc.), those others are languages in which a morpheme typically appears to mark the positive form of adjectives, leaving the comparative unmarked throughout the language (thus, the lack of a comparative morpheme essentially is the indication of the presence of the comparative). In contrast, Japanese has a comparative morpheme that is used in simple comparative sentences and also an equative morpheme (Pollard & Sag's 1994 'factor comparatives'), but neither of those appear in Correlational Comparatives in Japanese, such as the example in (28). Because Japanese would normally use a morpheme to express comparison in simple sentences (yori 'than', or moo/sarani 'more'), it is not the unmarked case, and thus its absence in the CC cannot be taken as evidence of a 'silent' comparative meaning as in Chinese. I suggest instead that verbal or adjectival repetition such as that of hasir in (28) can function like a comparative would in combination with an indicator of degree such as the word hodo, thus explaining why it is possible to lack a comparative morpheme (because a similar meaning is contributed in a different way based on the difference between Japanese and English). I also wonder whether this is related to the repetition of more in more and more in English CC-like sentences.

---

36 I choose the term 'repetition' rather than 'reduplication' in order to remain neutral about whether this is actually reduplication; the fact that both verbs are inflected, one with tense and one with a conditional marker, may suggest that this is not what is traditionally called 'reduplication'.

287
A prerequisite for a Japanese analysis is an understanding of the degree adverbial *hodo*. The linguistic literature on *hodo* is limited and not accessible to me (not being capable of reading Japanese fluently). *Hodo* appears in three different kinds of environments, all having to do with degrees, but with seemingly different senses, which is another arena in which further work is required.

One difficulty I foresee in the analysis of Japanese is the existence of CCs that do not use repetition in conjunction with *hodo*, namely those that are nominal rather than verbal, such as (29).

(29) *Gakuen-ga* ue-no *kodomo* hodo, *se-ga* takai. *school.year-NOM* upper-GEN child DEGREE height-NOM high ‘The older you are, the taller you are.’

The literal translation of (29) is more like 'the upper school year child (hodo), taller.' Here, 'upper' and 'school year' describe the child and the whole phrase is modified by *hodo*. As you can see, there is not an inflected verb or adjective to repeat in the first clause, but the sentence's interpretation is monotonically increasing anyway.

### 5.3 General Conclusion

The dissertation proposed in these pages contributes to the tradition of explicit formal theoretical work in natural language semantics by providing a concrete analysis that makes testable predictions. It simultaneously contributes to the empirical tradition in
linguistics though the delineation of benchmarks for the construction in question and the observation of new benchmarks that contribute to a deeper understanding of the CC. The following are some of the benefits of the analysis in this dissertation:

1. It is the only one that explains why a proportional rather than universal interpretation is the default.
2. For the above reason, it's the only one that has a way to explain why not all AQs can modify CCs.
3. It is the only approach that does not assimilate CCs and conditionality (though Brasoveanu does claim that not all CCs are conditionals, the only ones that are not conditional are regular correlatives without a monotonic correlational meaning).
4. Though Beck and Lin both use a Generalized Quantifier analysis of conditionals to model CCs, my claim that CCs contain GQs is new and different because I'm saying that the quantificational determiner is an overt (and necessary) part of the CC (namely, *the1*) rather than locating the GQ in an implicit adverb and drawing a parallel between a supposedly vacuous *if* and *the1*.
5. It accounts for a significantly larger number of the benchmarks than any previous analysis.

Further, this dissertation's analysis shows that what is unique about CCs is only the particular combination of language processes that exist as regularities elsewhere in that language, as revealed by these phenomena's parallels with the behavior of CCs. In other words, the syntactic (tecto and pheno) and semantic rules are the same for CCs and other sentences, the analyses of comparison, parastic scope, UDCs, generalized quantification, etc. are the same for CCs and other sentences, and the only thing that differentiates CCs from other sentences is (1) the lexical items like *the* that appear in the CC but not elsewhere and (2) the co-occurrence of so many kinds of phenomena such as comparison, long-distance dependency, quantification, anaphora, etc. in a single sentence.
Bibliography


Ladusaw, W. (1979). *Polarity sensitivity as inherent scope relations*. Unpublished PhD, University of Texas at Austin, Austin, TX.


Pollard, C. J. (2010). *Proof theoretic background for DyCG and DyCG rules*. Unpublished manuscript, Columbus, OH.


