3

Direct Compositionality on Demand*

CHRIS BARKER

3.1 Two Equally Valid Views of the Syntax–Semantics Interface

The problem is that many natural language expression types lead a double life, simultaneously here and there, masquerading as a local lump but somehow interacting directly with distant elements. The two main examples discussed here are bound anaphora, in which an anaphor depends for its value on some distant binder; and quantification, in which some local element takes semantic scope over a properly containing constituent. Such action-at-a-distance confronts theories of the syntax–semantics interface with a dilemma: should we interpret these elements locally, where they enter into the syntactic structure, or globally, where they take semantic effect?

Both approaches have staunch defenders. One well-established approach (e.g. Heim and Kratzer 1998) emphasizes the global perspective, relegating the interpretation of anaphors to variable assignment functions, and postponing

* This paper owes debts to three people: first and foremost, Pauline Jacobson, whose work inspired the workshop from which this paper developed, and who discussed many of the ideas below with me during my sabbatical visit to Brown in the fall of 2003. Second, David Dowty, whose remarks on the formal nature of direct compositionality expressed at the workshop and in conversation planted the seed that grew into this paper. Third, Gerhard Jäger, who first introduced me to the display property for type logics ("it is always possible to glue two adjacent constituents together before building the larger constituent. This requires a lot of cut applications…" (in an email from fall 2003)), which approximates my on-demand property. (See §3.8 for a discussion of the display property in comparison with the DCOD property.) Finally, I would like to point out that the crucial technical innovation that gives the logic presented below the on-demand property, namely, the rule of disclosure for quantification, qLR, was directly inspired by Jäger's (2005) similar rule for his binding analysis. In some sense, then, this paper merely emphasizes the importance of a symmetry already present in Jäger's LLC, and extends the same symmetry to quantification. Substantial improvements over the first draft are due to discussions with Chung-chieh Shan and Anna Szabolcsi.

the interpretation of quantifiers until their scope has been revealed via Quantifier Raising at a level of Logical Form.

Another tradition, going back at least to Montague, and recently championed by many working within Categorial Grammar and its related frameworks, emphasizes local interpretation. Jacobson, in a series of papers (e.g. Jacobson 1999), names this approach "Direct Compositionality": roughly (to be made somewhat more precise below), expressions must deliver their entire semantic payload at the moment they enter into a syntactic relationship.

I claim that both views are indispensable, and that any grammar which ignores either mode of meaning is incomplete. Can we hope for a system that adheres to direct compositionality, but without giving up the clarity and simplicity of a global view?

Of course, there is an uninteresting, trivial answer, which involves simply combining a directly compositional grammar with an empirically equivalent, redundant grammar that allows non-local action.

I will propose what I believe is a more interesting answer, in the form of DCOD (for "Direct Compositionality On Demand"), a grammar in which the long-distance and local analyses arise from one and the same set of rules, none of which are redundant. The on-demand property allows us to have our cake and eat it too: we can connect an anaphor directly with its antecedent, and we can connect a quantifier to its scope in a single intuitive leap; or else, if we prefer, we can articulate each derivation into incremental steps that reveal the semantic contribution of each syntactic constituent.

To give a slightly more detailed preview, for every derivation in DCOD in which an expression is bound at a distance or takes wide scope, there will be a syntactically and semantically equivalent derivation on which the semantic contribution of each constituent is purely local. Furthermore, the interconvertibility of the two styles of derivation does not simply follow from grafting a direct-compositional grammar onto an action-at-a-distance grammar; rather, the duality in the syntax–semantics interface follows from a natural symmetry in the grammar itself. The symmetry concerns Gentzen's rules of use and rules of proof. Roughly, in the grammar below, rules of use connect expressions directly over long distances, and embody the global view. Rules of proof help characterize the contribution of individual expressions within a complex constituent. Crucially, I introduce rules of disclosure, which establish an explicit connection between the long-distance semantic effect of an element with its local denotation. As a consequence, we can use the global-style rules of use without the slightest hesitation, confident that we can produce a parallel, strictly directly compositional elaboration upon demand.

The on-demand property, at least as I envision it below, is by no means a general characteristic of type-logical grammars. In particular, Jäger's (2005) LLC+q (LLC supplemented with rules for quantification) does not have the DCOD property. Nor do Display Calculi (Goré 1998; Bernardi 2002) necessarily have the property, despite the fact that Display Calculi guarantee a similar constraint, the "display property": that any syntactic constituent can be given a self-contained denotation. (More technically, that any structural element can be isolated on the left-hand side of a sequent.) The problem with these grammars is that the value determined by exercising the display guarantee often incorporates information about the derivation of elements external to the constituent in question. This is a fairly subtle but important point, and is developed in §3.8.

3.2 Three Characteristics of Direct Compositionality

The essence of direct compositionality is that every syntactic operation that combines two smaller expressions into a larger expression comes along with a semantic operation that combines the meanings of the smaller expressions to arrive at the meaning of the larger expression. This is just an informal characterization of Montague's Universal Grammar (as instantiated in, e.g., Montague 1974).

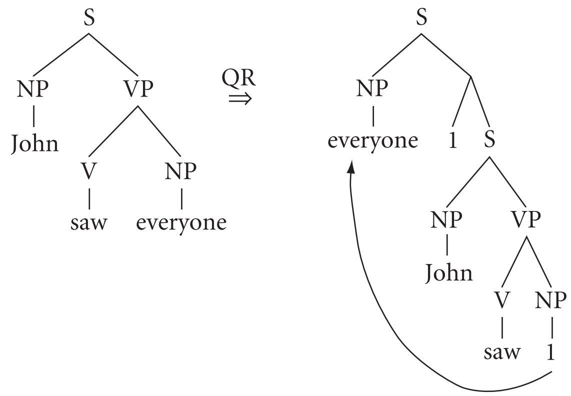
The conceptual simplicity and straightforwardness of direct compositionality has tremendous appeal. On the other hand, direct compositionality requires a certain amount of book-keeping: as we shall see, every semantically relevant aspect of a constituent must be available at the point at which the constituent enters into a syntactic structure, whether that aspect is relevant at that point or not. Thus combination rules must be adjusted to make use of information when it is required, and ignore it (but without discarding it!) when it is not.

3.2.1 *No Postponement*

On the direct compositional view, the order of syntactic combination is identical to the order of semantic combination.

This view of the syntax–semantics interface contrasts with the standard Quantifier-Raising (QR) approach to quantifier scope:

(1)



The reason that this approach fails to be directly compositional is that there is a point in the derivation at which *everyone* has already been combined syntactically with *saw* to form a verb phrase constituent (as evident in the left-most tree), but the verb phrase does not yet have a well-formed semantic denotation. In particular, there is no (relevant) way to directly combine a transitive verb such as *saw* of type $\langle \mathtt{e}, \langle \mathtt{e}, \mathtt{t} \rangle \rangle$ with a generalized quantifier such as *everyone* of type $\langle \langle \mathtt{e}, \mathtt{t} \rangle, \mathtt{t} \rangle$. Nor is there any obvious way of assigning a suitable denotation to the constituent *saw everyone*. It is only after the complete nuclear scope has been constructed (as in the right-most tree) that Quantifier Raising resolves the type mismatch and a complete meaning emerges.

The QR approach is rather like building a car on an assembly-line: it may be convenient to install the steering wheel before attaching the doors, even if the control cables and electronic connections that allow the steering wheel to guide the wheels do not get attached until later in the assembly process. Just so, in the QR model, the generalized quantifier *everyone* is inserted in direct object position, but its semantic control cables are left dangling until the rest of its clause has been assembled. I will call this sort of delayed evaluation POSTPONEMENT.

Forbidding postponement entails that no denotation has access to material contributed by an expression that is not part of the immediate syntactic expression. For instance, a pronoun must make its semantic contribution as soon as it enters into the syntactic construction, and cannot wait to find out what its antecedent is going to be.

The direct compositional ideal is a kind of zen semantics, living entirely in the moment of combination, unaware of what has happened or what is to come.

3.2.2 *Full Disclosure*

It follows from forbidding postponement that all syntactically and semantically relevant aspects of elements within a constituent must be accessible when the constituent combines with other elements. In other words, if a constituent contains a bindable pronoun or a quantifier, then that fact must be evident by inspecting the syntactic and semantic features of the constituent.

Jacobson (1999) works out in detail what full disclosure could look like for a directly compositional theory of binding; Shan and Barker (2006) show what full disclosure could look like for a theory of quantification.

3.2.3 *Self-Reliance*

The flip side of full disclosure is that the analysis of constituents should be self-contained. This means that the syntactic category and the semantic value should depend only on the lexical items dominated by the constituent and on the structure internal to the constituent. The analysis should certainly not depend in any way on any element external to the constituent.

I will suggest below that although the display property possessed by most type-logical grammars is capable of providing analyses of any constituent, those analyses often violate self-reliance. (See especially §3.8 below.) If the analysis of a subconstituent incorporates details that anticipate the specific structures in which it will be embedded, that is just a way of sneaking postponement through the back door.

3.3 DCOD, a Logic with Direct Compositionality on Demand

This section presents DCOD, the grammar analyzed in later sections. DCOD is a type-logical grammar. There are several quite different but more or less equivalent ways to present TLG. The two main styles are Natural Deduction vs. the Sequent Calculus presentation. I will use the sequent presentation here, since that will best facilitate the discussion of cut elimination further below. I have done my best to keep in mind readers who are not already familiar with sequent systems, but it may be helpful to consult more leisurely presentations of sequent logics such as Moortgat (1997), Restall (2000), or Jäger (2005).

3.3.1 *General Strategy*

The cut rule plays a special role in providing direct compositionality on demand. As discussed below, the cut rule expresses a kind of transitivity governing inference. Most discussions of the cut rule concentrate on proving that all cuts can be eliminated without reducing the generative power of the system, and thus that the cut rule is logically redundant. Cut elimination is important in order to prove such properties as logical consistency, guaranteed termination for the proof search algorithm, or that there will be at most a finite number of distinct interpretations for any given sentence. When considering these results, it is easy to get the impression that getting rid of cuts is always a very good thing.

But in fact it is *being able to* get rid of cuts that is the good thing. Cuts themselves are often quite useful: they correspond to proving a lemma, which you can then use and reuse in different contexts without having to reprove it each time. The connection with constituency is that it is possible to think of deriving, say, a noun phrase as a lemma. To say that a noun phrase is a constituent is to say that we could, if we chose, derive the noun phrase separately from the rest of the proof as a lemma, and then insert the lemma into the larger derivation in the position that the noun phrase occupies.

3.3.2 *DCOD*

Each rule in Example (2) relates either one or two antecedent sequents (appearing above the line) with exactly one consequent (below the line). Here, a sequent consists of a structure on the left-hand side of the sequent symbol ("$\Rightarrow$") and a single formula on the right-hand side. A structure is either a single formula or an ordered pair (Γ, Δ) in which Γ and Δ are both themselves structures. A formula is either a symbol such as np, n, or s, or else has the form $A\backslash B$, A/B, A^B, or $q(A, B, C)$, where A, B, and C are metavariables over formulas. For instance, the sequent $(np/n, n) \Rightarrow np$ says that a structure consisting of a determiner of category np/n followed by a noun of category n can form an NP of category np. From a linguistic point of view, formulas serve the role of syntactic categories, and structures indicate constituency. (You can think of a structure as a standard linguistic tree but without any syntactic category labels on the internal nodes.)

There are two kinds of information present in the rules given in Example (2): logical information, encoded by formulas expressing types; and semantic information, encoded by terms in the λ-calculus expressing how the meanings of the elements in the inference rule relate to one another. Thus $\mathtt{x}$: B stands for an expression of category B whose semantic value is named by x. In the derivations below, I will often omit the semantic part of the derivation when it enhances clarity.

Two rules are special, the axiom rule and the cut rule. The axiom rule is a simple tautology: given A, conclude A ("if it's raining, then it's raining").

(2)

$$\frac{}{\texttt{x}\colon A \Rightarrow \texttt{x}\colon A}\text{Axiom}$$

$$\frac{\Delta \Rightarrow \texttt{m}\colon A \quad \Gamma[\texttt{x}\colon A] \Rightarrow \texttt{p}\colon B}{\Gamma[\Delta] \Rightarrow \texttt{p}\{\texttt{m}/\texttt{x}\}\colon B}\text{Cut}$$

Rules of use (left rules):

$$\frac{\Delta \Rightarrow \texttt{x}\colon A \quad \Gamma[\texttt{y}\colon B] \Rightarrow \texttt{m}\colon C}{\Gamma[(\texttt{f}\colon (B/A), \Delta)] \Rightarrow \texttt{m}\{(\texttt{fx})/\texttt{y}\}\colon C}/\text{L}$$

$$\frac{\Delta \Rightarrow \texttt{x}\colon A \quad \Gamma[\texttt{y}\colon B] \Rightarrow \texttt{m}\colon C}{\Gamma[(\Delta, \texttt{f}\colon (A\backslash B))] \Rightarrow \texttt{m}\{(\texttt{fx})/\texttt{y}\}\colon C}\backslash\text{L}$$

$$\frac{\Delta \Rightarrow \texttt{m}\colon A \quad \Gamma[\texttt{x}\colon A][\texttt{y}\colon B] \Rightarrow \texttt{n}\colon C}{\Gamma[\Delta][\texttt{f}\colon B^A] \Rightarrow \texttt{n}\{\texttt{m}/\texttt{x}; (\texttt{fm})/\texttt{y}\}\colon C}\uparrow\text{L}$$

$$\frac{\Delta[\texttt{x}\colon B] \Rightarrow \texttt{m}\colon C \quad \Gamma[\texttt{y}\colon D] \Rightarrow \texttt{n}\colon E}{\Gamma[\Delta[\texttt{g}\colon q(B, C, D)]] \Rightarrow \texttt{n}\{(\texttt{g}(\lambda\texttt{x.m}))/\texttt{y}\}\colon E}q\text{L}$$

Rules of proof (right rules):

$$\frac{(\Gamma, \texttt{x}\colon A) \Rightarrow \texttt{m}\colon B}{\Gamma \Rightarrow \lambda\texttt{x.m}\colon (B/A)}/\text{R}$$

$$\frac{(\texttt{x}\colon A, \Gamma) \Rightarrow \texttt{m}\colon B}{\Gamma \Rightarrow \lambda\texttt{x.m}\colon (A\backslash B)}\backslash\text{R}$$

Rules of disclosure (left–right rules):

$$\frac{\Delta[\texttt{x}\colon A] \Rightarrow \texttt{m}\colon B}{\Delta[\texttt{f}\colon A^C] \Rightarrow \lambda\texttt{y.m}\{(\texttt{fy})/\texttt{x}\}\colon B^C}\uparrow\text{LR}$$

$$\frac{\Pi[\texttt{x}\colon A] \Rightarrow \texttt{m}\colon B}{\Pi[\texttt{g}\colon q(A, C, D)] \Rightarrow \lambda\texttt{f.g}(\lambda\texttt{x.}(\texttt{fm}))\colon q(B, C, D)}q\text{LR}$$

DCOD: a resource-sensitive logic with binding and quantification that guarantees Direct Compositionality On Demand. The only novel element compared with Jäger's (2005: 100) LLC is qLR.

The cut rule, which plays an important role in the discussion here and below, expresses a fundamental kind of logical transitivity: given $\Delta \Rightarrow A$, which says that Δ constitutes a proof of the formula A; and given $\Gamma[A] \Rightarrow B$, which says that Γ is a proof of B that depends on assuming A, it follows that $\Gamma[\Delta] \Rightarrow B$: substituting the reasoning that led to the conclusion A into the spot in the proof Γ where Γ depends on assuming A constitutes a valid complete proof of B. For instance, if you can prove that *the* plus *dog* forms a noun phrase (expressed in syntactic categories, $(np/n, n) \Rightarrow np$), and if you can prove that any noun phrase plus *barked* forms a complete sentence (i.e. $(np, np\backslash s) \Rightarrow s$), then the cut rule entitles you to conclude that ((*the dog*) *barked*) forms a complete sentence ($((np/n, n), np\backslash s) \Rightarrow s$).

In the cut rule, the notation $\Gamma[A]$ schematizes over structures Γ that contain at least one occurrence of the formula A somewhere within it. The notation $\Gamma[\Delta]$ in the consequent represents a structure similar to Γ except with Δ substituted in place of (the relevant occurrence of) A. (Examples immediately below will illustrate how this works.)

Apart from axiom and cut, most of the remaining rules introduce a single new logical connective (i.e. one not occurring in the antecedents) into the consequent, either on the left-hand side of the sequent symbol (the "left" rules) or else on the right-hand side (the "right" rules). As discussed below, the two crucial rules for the discussion here introduce a new connective on both sides of the sequent, and so are "left–right" rules. (Purely for the sake of expository simplicity, I have omitted logical rules for the • ("product") connective, which plays a prominent role in many type-logical discussions, but is not needed for any of the derivations below.)

A derivation is complete if all of its branches end (reading bottom up) using only instances of the Axiom as antecedents, and if each conclusion is derived from the antecedents above it via a legitimate instantiation of one of the logical rules. Here is a complete derivation proving that *the dog barked* is a sentence:

(3)
$$\dfrac{\dfrac{\dfrac{}{n \Rightarrow n}\text{Axiom} \quad \dfrac{}{np \Rightarrow np}\text{Axiom}}{(np/n, n) \Rightarrow \boxed{\text{np}}}/\text{L} \qquad \dfrac{\dfrac{}{np \Rightarrow np}\text{Axiom} \quad \dfrac{}{s \Rightarrow s}\text{Axiom}}{(\boxed{\text{np}}, np\backslash s) \Rightarrow s}\backslash\text{L}}{\left(\left(\begin{matrix}np/n \\ \texttt{the}\end{matrix}, \begin{matrix}n \\ \texttt{dog}\end{matrix}\right), \begin{matrix}np\backslash s \\ \texttt{barked}\end{matrix}\right) \Rightarrow \begin{matrix}s \\ \texttt{barked(the(dog))}\end{matrix}}\text{Cut}$$

I have placed boxes around the formulas that instantiate the As targeted by the cut rule. (From this point on, I will assume that applications of the axiom rule are obvious, and so do not need to be explicitly indicated.)

An expression is generated by a type-logical grammar just in case there are lexical items whose category labels match the formulas in the result (bottom-most) sequent and which appear in the same order as the formulas in the sequent. Thus if the word *the* has meaning `the` and category np/n, *dog* has meaning `dog` and category n, and *barked* has meaning `barked` and category $np\backslash s$, then the derivation here proves that *the dog barked* has category s, with semantic interpretation $\texttt{barked}(\texttt{the}(\texttt{dog}))$. Furthermore, in that derivation, *the dog* forms a constituent.

In this case, the structure of the derivation corresponds to the constituent structure of the conclusion sequent: the NP $(np/n, n)$ is a constituent in the conclusion sequent, and the main subderivation on the left (the first antecedent of the lowest inference) shows how to construct a subproof that $(np/n, n)$ forms an NP. Unfortunately, this graceful correspondence between the form of the derivation and constituency (as determined by the structure in the conclusion sequent) is not guaranteed. Indeed, arriving at such a guarantee is the main topic of this paper.

3.4 Rules of Use, Rules of Proof, and Rules of Disclosure

Following Lambek (1958) (in turn following Gentzen), the distinctive feature of TLG compared to plain categorial grammar is the ability to employ hypothetical reasoning. In terms of the sequent presentation used here, there are two types of logical rules, called left rules and right rules, or rules of use and rules of proof. Plain categorial grammar gets by with only (the equivalent of) the left rules, the rules of use. Because rules other than rules of use will be crucial for my main argument, this section briefly motivates the utility of rules of proof for linguistic analyses.

The usual motivating examples typically involve either function composition or relative clause formation. But these examples also require structural postulates that render function application associative, which would significantly complicate exposition at this point (structural postulates are discussed below in §3.9).

But even in the absence of associativity, rules of proof can perform useful linguistic work by deriving a certain class of lifted predicates. Dowty (2000) suggests that adjuncts may sometimes be re-analyzed as arguments. If *sing* has category v (where v abbreviates the category $(np\backslash s)/np$), then if *well* is a verbal adjunct of category $v\backslash v$, *sing well* is correctly predicted to be a complex v, as shown here:

(4)
$$\frac{v \Rightarrow v \qquad v \Rightarrow v}{\left(\begin{array}{c} v \\ \texttt{sing} \end{array}, \begin{array}{c} v\backslash v \\ \texttt{well} \end{array}\right) \Rightarrow \begin{array}{c} v \\ \texttt{well(sing)} \end{array}} \backslash\mathrm{L}$$

It is also possible for a higher-order verb to take a verbal modifier as an argument:

(5)
$$\dfrac{v\backslash v \Rightarrow v\backslash v \quad v \Rightarrow v}{\left(\begin{array}{c}v/(v\backslash v)\\ \mathtt{behave}\end{array}, \begin{array}{c}v\backslash v\\ \mathtt{well}\end{array}\right) \Rightarrow \begin{array}{c}v\\ \mathtt{behave(well)}\end{array}}\,/\mathrm{L}$$

A language learner may not be able to tell whether to assign *sing* to the category v or the category $v/(v\backslash v)$. This is harmless, however, since as the following derivation shows, it is a theorem of DCOD (and any Lambek grammar) that anything in the category v can behave as if it were of category $v/(v\backslash v)$, the same category as *behave*:

(6)
$$\dfrac{\dfrac{v \Rightarrow v \quad v \Rightarrow v}{(v, v\backslash v) \Rightarrow v}\,\backslash\mathrm{L}}{\begin{array}{c}v\\ \mathtt{sing}\end{array} \Rightarrow \begin{array}{c}v/(v\backslash v)\\ \lambda\mathtt{a.a(sing)}\end{array}}\,/\mathrm{R}$$

Thus whether you think of *sing* in *sing well* as the argument of *well* or else as a lifted predicate that takes *well* as an argument is completely optional as far as the logic is concerned. We can assume, then, that the simplest lexicon will be one in which *sing* has category v, but verbs that require a modifier must be of category $v/(v\backslash v)$.

One prediction that this analysis makes is that it should be possible to conjoin a simple verb like *sing* with a verb of higher category like *behave*. This is a good prediction, since *sing and behave well* certainly is a legitimate coordination:

(7)
$$\dfrac{v/(v\backslash v) \Rightarrow v/(v\backslash v) \qquad \dfrac{\dfrac{\dfrac{v \Rightarrow v \quad v \Rightarrow v}{(v, v\backslash v) \Rightarrow v}\,\backslash\mathrm{L}}{v \Rightarrow v/(v\backslash v)}\,/\mathrm{R} \qquad v/(v\backslash v) \Rightarrow v/(v\backslash v)}{(v, (v/(v\backslash v))\backslash(v/(v\backslash v))) \Rightarrow v/(v\backslash v)}\,\backslash\mathrm{L}}{\left(\begin{array}{c}v\\ \mathtt{sing}\end{array}, \left(\begin{array}{c}X\backslash(X/X)\\ \mathtt{and}\end{array}, \begin{array}{c}v/(v\backslash v)\\ \mathtt{behave}\end{array}\right)\right) \Rightarrow \begin{array}{c}v/(v\backslash v)\\ \lambda\mathtt{m.m(sing)} \wedge \mathtt{behave(m)}\end{array}}\,/\mathrm{L}$$

In this derivation, the phrase *and behave* is something that is looking for an expression of the same category as *behave* in order to form a coordinate structure. *Sing* is able to fill that role because of the /R rule, which is a crucial part of the deduction that *sing* is able to combine with an adverb. Note that in the semantic interpretation, the variable m (representing a manner) corresponding to the adverb *well* modifies *sing*, but serves as an argument to *behave*, as desired.

In any case, rules of proof can be motivated independently of the main issues of this paper.

The innovative rules in DCOD, however, are not straightforward rules of proof. True rules of proof such as \R or /R introduce a new connective only on the right side of the sequent. The rules of special interest here do introduce a new connective on the right side of the sequent, but they also introduce a new connective on the left side of the sequent at the same time. Following a suggestion of Chung-chieh Shan (p.c.), I will label such rules LR rules: simultaneously left rules and right rules. (Thus the rule Jäger (2005) refers to as ↑R I will call ↑LR.) These LR rules are not really rules of use, since they are not sufficient to license the use of, say, a pronoun or a quantifier; and they are not really rules of proof, since they do not discharge any hypotheses in the way that \R or /R do. (Section 3.7 will discuss other unusual properties of the LR rules with respect to cut elimination.) As we shall see, what these new rules do is transmit information about subconstituents to higher levels. Therefore I will call them **rules of disclosure**.

Thus DCOD leaves open the question of what a true rule of proof would be for binding or for quantification. This is an area of active research; for one detailed view of how to handle scope-taking in a type-logical grammar, see Barker and Shan (in press).

3.5 First Case Study: Binding

Here is the swooping (cut-free) derivation of *John*$_i$ *said he*$_i$ *left*.

(8)

$$\dfrac{np \Rightarrow np \qquad \dfrac{np \Rightarrow np \qquad \dfrac{s \Rightarrow s \qquad \dfrac{np \Rightarrow np \qquad s \Rightarrow s}{(np, np\backslash s) \Rightarrow s}\,\backslash\mathrm{L}}{(np, ((np\backslash s)/s, s)) \Rightarrow s}\,/\mathrm{L}}{(\boxed{\mathrm{np}}, ((np\backslash s)/s, (\boxed{\mathrm{np}}, np\backslash s))) \Rightarrow s}\,\backslash\mathrm{L}}{\left(\begin{matrix} np \\ \texttt{John} \end{matrix}, \left(\begin{matrix} (np\backslash s)/s \\ \texttt{said} \end{matrix}, \left(\begin{matrix} np^{np} \\ \texttt{he} \end{matrix}, \begin{matrix} np\backslash s \\ \texttt{left} \end{matrix}\right)\right)\right) \Rightarrow \begin{matrix} s \\ \texttt{said(left(he(j)))(j)} \end{matrix}}\,\uparrow\mathrm{L}$$

The form of the binding rule in Example (2) requires some comment. In the second antecedent, $\Gamma[A][B]$ matches a structure Γ that contains an occurrence α of the formula A and an occurrence β of formula B. (In addition, α must precede β, but linear precedence will not be an important factor below.) In the conclusion, $\Gamma[\Delta][A^B]$ indicates the structure constructed by starting with Γ, replacing α with Δ, and replacing β with the formula A^B. In the diagram immediately above, α is the left box, and β is the right box. In the conclusion, α is replaced with np (no net change), and β is replaced with np^{np}.

In this derivation, there is exactly one application of the binding rule $\uparrow$L, operating over a potentially unbounded distance. I have boxed the NP targets of the $\uparrow$L rule, that is the elements to be bound. Intuitively, the rule in effect coindexes the boxed elements, guaranteeing that their interpretations will be linked. Naturally, the pronoun denotes a function from NP meanings to NP meanings, in this case, the identity function (see, e.g., Jacobson 1999 for discussion on this point).

Crucially, this analysis does not provide self-contained interpretations for each constituent. That is, this derivation is not directly compositional. In particular, consider the constituent *he left*. If any substring of the larger sentence deserves to be a constituent, it is the embedded clause! The final sequent recognizes that *he left* is a syntactic constituent, since *he* and *left* are grouped together into a substructure (as indicated by the parentheses in the bottom-most sequent) yet there is no point at which the denotation of the constituent is treated as a semantic unit.

To see why, focus on the second (lower) instance of \L, repeated here with full Curry–Howard labeling:

(9)
$$\frac{\begin{matrix} np \\ \mathtt{x} \end{matrix} \Rightarrow \begin{matrix} np \\ \mathtt{x} \end{matrix} \qquad \left(\begin{matrix} np \\ \mathtt{y} \end{matrix}, \left(\begin{matrix} (np\backslash s)/s \\ \mathtt{said} \end{matrix}, \begin{matrix} s \\ \mathtt{p} \end{matrix} \right) \right) \Rightarrow \begin{matrix} s \\ \mathtt{said(p)(y)} \end{matrix}}{\left(\begin{matrix} np \\ \mathtt{y} \end{matrix}, \left(\begin{matrix} (np\backslash s)/s \\ \mathtt{said} \end{matrix}, \left(\begin{matrix} np \\ \mathtt{x} \end{matrix}, \begin{matrix} np\backslash s \\ \mathtt{left} \end{matrix} \right) \right) \right) \Rightarrow \begin{matrix} s \\ \mathtt{said(left(x))(y)} \end{matrix}} \backslash L$$

This is the step that articulates the embedded clause into a subject and a verb phrase, so this is the step that justifies the claim that *he* and *left* are structural siblings. The contribution to the semantic value made by this step is the expression $\mathtt{left(x)}$, where x is a variable introduced by the instance of the axiom rule that justifies the left-most antecedent. This variable does not receive a value until the binding rule ($\uparrow$L) applies in the next derivational step, at which point the value of the embedded subject is bound to the value of the matrix subject. But the matrix subject is external to the constituent in question, so this labeling constitutes semantic postponement, and violates semantic self-reliance.

In other words, the cut-free derivation above clearly associates the position of the pronoun with its antecedent, so it accounts beautifully for the long-distance aspect of binding. However, quite sadly, it does not provide a full account of the sense in which *he left* is a constituent in its own right.

The main result of this paper guarantees the existence of a distinct derivation in DCOD that arrives at the same constituent structure and that has an identical semantic value, but in which each constituent can be associated with a self-contained semantic value. It is this second, related, derivation that provides the missing piece of the puzzle, and that characterizes *he left* as a constituent.

Arriving at such an alternative derivation involves use of the rule of proof ($\uparrow$LR) and several applications of the cut rule. Here's one way to do it:

(10)

$$\dfrac{\dfrac{\dfrac{np \Rightarrow np \qquad s \Rightarrow s}{(np, np\backslash s) \Rightarrow s}\,\backslash\mathrm{L}}{\left(\begin{matrix} np^{np} \\ \texttt{he} \end{matrix}, \begin{matrix} np\backslash s \\ \texttt{left} \end{matrix}\right) \Rightarrow \begin{matrix} s^{np} \\ \lambda\texttt{x.left(he(x))} \end{matrix}}\,\uparrow\mathrm{LR} \qquad \dfrac{\dfrac{s \Rightarrow s \qquad np\backslash s \Rightarrow np\backslash s}{((np\backslash s)/s, s) \Rightarrow np\backslash s}\,/\mathrm{L}}{((np\backslash s)/s, s^{np}) \Rightarrow (np\backslash s)^{np}}\,\uparrow\mathrm{LR}}{\left(\begin{matrix} (np\backslash s)/s \\ \texttt{said} \end{matrix}, \left(\begin{matrix} np^{np} \\ \texttt{he} \end{matrix}, \begin{matrix} np\backslash s \\ \texttt{left} \end{matrix}\right)\right) \Rightarrow \begin{matrix} (np\backslash s)^{np} \\ \lambda\texttt{y.said(left(he(y)))} \end{matrix}}\,\mathrm{Cut}$$

This derivation for the verb phrase constituent *said he left* can participate in a complete derivation as follows:

$$\dfrac{(np\backslash s)/s, (np^{np}, np\backslash s) \Rightarrow (np\backslash s)^{np} \qquad \dfrac{np \Rightarrow np \qquad \dfrac{np \Rightarrow np \qquad s \Rightarrow s}{(np, np\backslash s) \Rightarrow s}\,\backslash\mathrm{L}}{(np, (np\backslash s)^{np}) \Rightarrow s}\,\uparrow\mathrm{L}}{\left(\begin{matrix} np \\ \texttt{John} \end{matrix}, \left(\begin{matrix} (np\backslash s)/s \\ \texttt{said} \end{matrix}, \left(\begin{matrix} np^{np} \\ \texttt{he} \end{matrix}, \begin{matrix} np\backslash s \\ \texttt{left} \end{matrix}\right)\right)\right) \Rightarrow \begin{matrix} s \\ \texttt{said(left(he(j)))(j)} \end{matrix}}\,\mathrm{Cut}$$

The final sequents of the two completed derivations are identical, so both derivations provide the same constituency and the same final semantic interpretation. However, the second derivation provides a self-contained denotation for each constituent, where "self-contained" means that each element in the denotation either turns out to be the denotation of a word in the final sequent (e.g., `left`) or else is a variable that is bound within the denotation (e.g., `x` in λ`x.left(x)` but not in `left(x)`).

This cut-full derivation beautifully captures the intuition that *he left* is a constituent. However, now the link between the pronoun and its binder has been sadly obscured.

In order to emphasize the role of the cut rule in encapsulating the meaning of a subconstituent, the diagram above provides Curry–Howard semantic labels only for the final sequent in a derivation and for the first antecedent of each application of the cut rule, that is the antecedent that corresponds to an encapsulated constituent. (It should be clear how to complete the labelings of each of the derivations based on the labeling annotations in Example (2).) Here is what a complete labeling for the constituent *he left* would look like:

(11)
$$\dfrac{\dfrac{\texttt{x}\colon np \Rightarrow \texttt{x}\colon np \qquad \texttt{p}\colon s \Rightarrow \texttt{p}\colon s}{(\texttt{x}\colon np,\ \texttt{left}\colon (np\backslash s)) \Rightarrow \texttt{left(x)}\colon s}\backslash\mathrm{L}}{\begin{pmatrix} np^{np} & np\backslash s \\ \texttt{he} & \texttt{left} \end{pmatrix} \Rightarrow \begin{matrix} s^{np} \\ \lambda\texttt{x.left(he(x))} \end{matrix}}\uparrow\mathrm{LR}$$

This complete labeling of the constituent shows how the ↑LR rule binds the argument of the pronoun, making the denotation self-contained. The troublesome element in the swooping derivation was the variable `x`, whose value was determined from outside the constituent. In the final labeling here, λ`x.left(x)`, `x` is bound by the lambda.[1] As Jäger points out, it is not a coincidence that the syntactic category and the denotation of *he left* coincides exactly with the analysis proposed in Jacobson's (1999) directly compositional theory; in other words, both analyses say the same thing about the directly compositional aspect of the binding relationship.

In the swooping derivation, there is no syntactic category labeling the structure corresponding to the constituent *he left*. In the directly compositional derivation, the syntactic category is revealed to be s^{np}, a sentence containing a bindable pronoun. Thus the DC derivation imposes full disclosure, annotating on the category the presence of bindable pronoun inside. Similarly, in the swooping derivation, there is no semantic denotation corresponding exactly to the constituent *he left*. There is a contribution to the Curry–Howard labeling, but one of the elements contributed (namely, the variable `x`) ultimately depends on material outside of the constituent for its final value, rather than on some function of the meanings of the words contained within the constituent. Thus in addition to a transparent constituent structure, the DC derivation also imposes semantic self-reliance.

On the other hand, only in the swooping derivation can we clearly see the net result of the series of incremental abstractions and substitutions that the DC derivation plods through. By linking the boxed *np*s directly in the swooping derivation, we get a satisfying account of what the pronoun takes as its antecedent. In addition, it is worth noting that the swooping derivation is considerably shorter and simpler.

Which derivation is superior? With DCOD, we do not need to decide: both are equally available. Furthermore, as we shall see below, it is possible

[1] The version of ↑LR given here is simplified from Jäger's (2005: 100) more general rule. Jäger's rule handles cases in which more than one pronominal element is bound by the same antecedent.

to prove that they are equivalent. Therefore they are equally valid analyses of the sentence in question, and neither one alone tells the full story.

3.5.1 *Deictic Pronouns*

One of the virtues of Jacobson's treatment is that there is no lexical distinction between bound pronouns and deictic pronouns: a deictic pronoun is simply a pronoun that never happens to get bound. Yet in the DCOD grammar, the only way of introducing a functional dependency (a category with shape B^A) is either by way of the $\uparrow$L rule, in which case there is an overt binder somewhere else in the derivation, or by way of the $\uparrow$LR rule, in which case the functional dependency is linked to the denotation of the result category, and in effect bound from above. Another way of saying this is that the denotations assigned by DCOD are all closed, that is, they are lambda terms containing no free variables. If DCOD pronouns are always bound by something, does this mean that we need to introduce a second strategy to handle deictic uses?

No. Since the $\uparrow$ rules in the DCOD grammar are (simplified) versions of Jäger's rules, they share the strengths and weaknesses of his analysis. Just like Jacobson's system, Jäger's grammar and the DCOD grammar automatically accommodate deictic uses. To see how this works, here is a derivation of the deictic reading of the same example that was given a bound reading above at the beginning of §3.5:

(12)

$$\dfrac{\dfrac{np \Rightarrow np \qquad \dfrac{s \Rightarrow s \qquad \dfrac{np \Rightarrow np \qquad s \Rightarrow s}{(np, np\backslash s) \Rightarrow s}\,\backslash L}{(np, ((np\backslash s)/s, s)) \Rightarrow s}\,/L}{(np, ((np\backslash s)/s, (np, np\backslash s))) \Rightarrow s}\,\backslash L}{\left(\begin{matrix}np\\ \texttt{John}\end{matrix}, \left(\begin{matrix}(np\backslash s)/s\\ \texttt{said}\end{matrix}, \left(\begin{matrix}np^{np}\\ \texttt{he}\end{matrix}, \begin{matrix}np\backslash s\\ \texttt{left}\end{matrix}\right)\right)\right) \Rightarrow \begin{matrix}s^{np}\\ \lambda\texttt{x.sd(lft(he(x)))(j)}\end{matrix}}\,\uparrow LR$$

The derivations are the same until the last step. Instead of applying the $\uparrow$L rule in order to bind the pronoun to the matrix subject, the $\uparrow$LR rule discloses the presence of the bindable pronoun at the top level of the derivation. (There is, of course, an equivalent directly compositional derivation as well.)

In fact, to the extent that Jäger's left and left–right rules are indeed two logical aspects of a single inference type, they arguably provide an analysis of bound vs. deictic pronouns that is more unified even than Jacobson's, since on her analysis, the analog of the binding rule ($\uparrow$L, her **z** type-shifter) and the analog of the disclosure rule ($\uparrow$LR, her Geach type-shifter, **g**) do not resemble each other in any obvious way.

3.6 Second Case Study: Quantification

The analysis of binding in DCOD is (intended to be) identical to Jäger's (2005) treatment. Therefore his LLC has direct compositionality on demand, at least with respect to binding. This section considers quantification, a different kind of action at a distance, and shows how the qLR rule in Example (2) guarantees the DCOD property for quantification as well.

Here is the swooping derivation of *John saw everyone*:

(13)
$$\dfrac{\dfrac{np \Rightarrow np \qquad \dfrac{np \Rightarrow np \qquad s \Rightarrow s}{(np, np\backslash s) \Rightarrow s}\backslash\mathrm{L}}{(np, ((np\backslash s)/np, np)) \Rightarrow s}/\mathrm{L} \qquad s \Rightarrow s}{\left(\begin{matrix} np \\ \texttt{John} \end{matrix}, \left(\begin{matrix} (np\backslash s)/np \\ \texttt{saw} \end{matrix}, \begin{matrix} q(np, s, s) \\ \texttt{everyone} \end{matrix}\right)\right) \Rightarrow \begin{matrix} s \\ \texttt{eo}(\lambda\texttt{x}.\texttt{saw}(\texttt{x})(\texttt{j})) \end{matrix}} q\mathrm{L}$$

The qL rule applies only after the entire scope of the quantificational element is in view, linking the quantifier with its nuclear scope in one leap. Nevertheless, the final structure clearly indicates that *saw everyone* is a syntactic constituent.

Now for the corresponding direct compositional derivation.

(14)
$$\dfrac{\dfrac{np \Rightarrow np \qquad np\backslash s \Rightarrow np\backslash s}{((np\backslash s)/np, np) \Rightarrow np\backslash s}/\mathrm{L}}{\left(\begin{matrix} (np\backslash s)/np \\ \texttt{saw} \end{matrix}, \begin{matrix} q(np, s, s) \\ \texttt{everyone} \end{matrix}\right) \Rightarrow \begin{matrix} q(np\backslash s, s, s) \\ \lambda\texttt{f}.\texttt{eo}(\lambda\texttt{x}.\texttt{f}(\texttt{saw}(\texttt{x}))) \end{matrix}} q\mathrm{LR}$$

This derivation for the verb phrase constituent *saw everyone* can participate in a complete derivation as follows:

$$\dfrac{(np\backslash s)/np, q(np, s, s) \Rightarrow q(np\backslash s, s, s) \qquad \dfrac{\dfrac{np \Rightarrow np \qquad s \Rightarrow s}{(np, np\backslash s) \Rightarrow s}\backslash\mathrm{L} \qquad s \Rightarrow s}{(np, q(np\backslash s, s, s)) \Rightarrow s} q\mathrm{L}}{\left(\begin{matrix} np \\ \texttt{John} \end{matrix}, \left(\begin{matrix} (np\backslash s)/np \\ \texttt{saw} \end{matrix}, \begin{matrix} q(np, s, s) \\ \texttt{everyone} \end{matrix}\right)\right) \Rightarrow \begin{matrix} s \\ \texttt{eo}(\lambda\texttt{x}.\texttt{saw}(\texttt{x})(\texttt{j})) \end{matrix}} \mathrm{Cut}$$

Once again, *saw everyone* is a constituent. It is not, however, a simple VP of category $np\backslash s$; rather, full disclosure requires that it be of category $q(np\backslash s, s, s)$, that is, a quantificational verb phrase: something that functions locally as a verb phrase, takes scope over a clause, and produces a clause as a result.

In other words, arriving at a directly compositional treatment has turned the verb phrase transparent, so that the category reveals the presence of a quantificational element somewhere within.

As in the previous case study, once again I will provide a full Curry–Howard labeling for the lemma analyzing the constituent under study, *saw everyone*:

(15)
$$\dfrac{\dfrac{\mathtt{x}\colon np \Rightarrow \mathtt{x}\colon np \qquad \mathtt{r}\colon (np\backslash s) \Rightarrow \mathtt{r}\colon (np\backslash s)}{(\,\mathtt{saw}\colon ((np\backslash s)/np),\ \mathtt{x}\colon np) \Rightarrow \mathtt{saw(x)}\colon (np\backslash s)}\,/\mathrm{L}}{\left(\begin{matrix}(np\backslash s)/np \\ \mathtt{saw}\end{matrix}, \begin{matrix}q(np,s,s) \\ \mathtt{everyone}\end{matrix}\right) \Rightarrow \begin{matrix}q(np\backslash s, s, s) \\ \lambda\mathtt{f.everyone}(\lambda\mathtt{x.f(saw(x))})\end{matrix}}\,q\mathrm{LR}$$

Here $\mathtt{f}$ is a function from verb phrase denotations of category $np\backslash s$ to clause denotations of category s. As a result, the semantic type of this constituent is $\langle\langle\langle \mathtt{e}, \mathtt{t}\rangle, \mathtt{t}\rangle, \mathtt{t}\rangle$: a function from a function from verb phrase meanings to truth values to truth values. This is exactly the type for a continuized verb phrase in the (directly compositional) continuation-based approach to quantification described in Barker (2002).

Once again, the swooping derivation is simpler, both conceptually and practically, an advantage that increases dramatically as the quantifier takes wider and wider scope. And once again the DC version imposes the discipline of full disclosure and self-reliance, giving a full accounting of the syntactic and semantic nature of each syntactic constituent.

3.6.1 *The Locus of Scope Ambiguity*

When more than one quantificational expression is present, the order in which the quantifiers are introduced into the derivation can give rise to scope ambiguity. Because we have the DCOD property, we know that every distinct construal will have an equivalent derivation in which each constituent is given a self-contained denotation. We might ask, therefore, which constituents must or can participate in scope ambiguities.

In order to explore this question, I will consider a sentence that contains three quantificational expressions, such as *Most people gave something to every child*. Assume that this sentence has a construal on which it entails that for every child, most people gave that child something, and that the scope relations are *every* > *most* > *some*. For expository simplicity, I will ignore the preposition *to*.

Let Q abbreviate the category $q(np, s, s)$. Then on the swooping derivation, the scope construal depends in a straightforward manner on the order of the instantiations of the qL rule:

(16)

$$\dfrac{\dfrac{\dfrac{\dfrac{np \Rightarrow np \qquad \dfrac{np \Rightarrow np \qquad \dfrac{np \Rightarrow np \quad s \Rightarrow s}{(np, np\backslash s) \Rightarrow s}\backslash\mathrm{L}}{(np, ((np\backslash s)/np, np)) \Rightarrow s}\backslash\mathrm{L}}{(np, ((((np\backslash s)/np)/np, np), np)) \Rightarrow s}\backslash\mathrm{L} \qquad s \Rightarrow s}{(np, ((((np\backslash s)/np)/np, Q_3), np)) \Rightarrow s}q\mathrm{L} \qquad s \Rightarrow s}{(Q_2, ((((np\backslash s)/np)/np, Q_3), np)) \Rightarrow s}q\mathrm{L} \qquad s \Rightarrow s}{(Q_2, ((((np\backslash s)/np)/np, Q_3), Q_1)) \Rightarrow s}q\mathrm{L}$$

Reading the derivation from the bottom upwards (as usual with sequent proofs), quantifiers that are eliminated earlier take wider scope. Here, $Q_1 > Q_2 > Q_3$, as desired. Since the qL rule can target any NP on the left-hand side of the sequent arrow, we have complete control over the scope relations among the quantifiers.

In the equivalent directly compositional derivation, we must cut out two subconstituents: the verb phrase, of course (category $np\backslash s$); and within the verb phrase, the constituent formed when the ditransitive verb combines with its first argument (category $(np\backslash s)/np$)). Because the verb phrase contains two quantifiers, full disclosure requires that the presence of both quantifiers must be registered on the category of the verb phrase. As a second abbreviation, let $Q(A)$ abbreviate the category $q(A, s, s)$. Then the category of the verb phrase after disclosure will be the category $Q(Q(np\backslash s)) = q(q(np\backslash s, s, s), s, s)$. Here, then, is a directly compositional analysis of the relevant construal of the verb phrase *gave something to everyone*:

(17)

$$\dfrac{\dfrac{\dfrac{np \Rightarrow np \quad (np\backslash s)/np \Rightarrow (np\backslash s)/np}{(((np\backslash s)/np)/np, np) \Rightarrow (np\backslash s)/np}\backslash\mathrm{L}}{(((np\backslash s)/np)/np, Q_3) \Rightarrow Q_3((np\backslash s)/np)}q\mathrm{LR} \qquad \dfrac{\dfrac{\dfrac{np \Rightarrow np \quad np\backslash s \Rightarrow np\backslash s}{((np\backslash s)/np, np) \Rightarrow np\backslash s}\backslash\mathrm{L}}{(Q_3((np\backslash s)/np), np) \Rightarrow Q_3(np\backslash s)}q\mathrm{LR}}{(Q_3((np\backslash s)/np), Q_1) \Rightarrow Q_1(Q_3(np\backslash s))}q\mathrm{LR}}{((((np\backslash s)/np)/np, Q_3), Q_1) \Rightarrow Q_1(Q_3(np\backslash s))}\mathrm{Cut}$$

Because we could have instantiated the two right-most applications qLR rules in the opposite order, there are two distinct possible analyses of the verb phrase. These alternatives correspond to the two relative scope relations between Q_1 and Q_3. In general, in a directly compositional derivation, the derivation of each constituent fully determines the relative scope of all (and only!) those quantifiers contained within that constituent.

At this point, we need only show how to compose the subject quantifier with the verb phrase:

(18)
$$\dfrac{\dfrac{\dfrac{\dfrac{np \Rightarrow np \quad s \Rightarrow s}{(np, np\backslash s) \Rightarrow s}\,\backslash\mathrm{L} \quad s \Rightarrow s}{(np,\, Q_3(np\backslash s)) \Rightarrow s}\,q\mathrm{L} \quad s \Rightarrow s}{(Q_2,\, Q_3(np\backslash s)) \Rightarrow s}\,q\mathrm{L} \quad s \Rightarrow s}{(Q_2,\, Q_1(Q_3(np\backslash s))) \Rightarrow s}\,q\mathrm{L}$$

To complete the derivation of the entire sentence, simply cut the derivation of the verb phrase given just above against this derivation of a complete clause.

The order of the qL rules here controls whether the subject outscopes both quantifiers in the verb phrase, or else takes intermediate scope (as shown here), or else takes narrow scope with respect to both verb phrase quantifiers. However, this part of the derivation has no power to affect the relative order of the quantifiers within the verb phrase, since that is fully determined by the derivation of the verb phrase, as described above.

Thus in a directly compositional derivation with full disclosure, the derivation of each constituent determines the relative scoping of all of the quantifiers it contains. In particular, even though the quantifiers involved take scope only over complete clauses (and not over verb phrases), the verb phrase nevertheless exhibits its own local scope ambiguity, and it is those local scope relations that determine the contribution of the verb phrase to the scope relations of the larger derivation in which it is embedded. In other words, semantic self-reliance requires that each constituent takes full responsibility for every aspect of the contribution of its contents within the larger derivation.

3.7 Direct Compositionality on Demand

Having provided two examples in detail, I now establish that it is always possible to provide both a swooping and a directly compositional analysis for any sentence generated by DCOD. Furthermore, the two analyses are guaranteed equivalent structurally and semantically, and interconvertible in both directions.

Converting from the DC analysis to the swooping analysis requires eliminating cuts, so I first show that the cut rule is admissible in DCOD (i.e. that cuts can always be eliminated). Then I show how to add cuts back in where desired in order to construct a fully DC analysis.

3.7.1 *Cut Elimination for DCOD*

Jäger (2005: 102) sketches cut elimination for LLC (see pages 43ff, 102ff, and 127), a logic similar to DCOD except in two respects: (i) DCOD leaves out the • connective (for purely expository reasons), which does not affect cut elimination; and (ii) DCOD adds qL and qLR. Therefore we need only show that adding qL and qLR does not interfere with cut elimination. The only potentially troublesome situations are ones in which some application of qLR is cut against an instance of either qL or qLR.

Consider first a derivation containing a cut of qLR against qL (for instance, as in the DC derivation of *John saw everyone* above in §3.6):

(19)
$$\dfrac{\dfrac{\Pi[A] \Rightarrow B}{\Pi[q(A,C,D)] \Rightarrow q(B,C,D)}\,q\mathrm{LR} \qquad \dfrac{\Delta[B] \Rightarrow C \qquad \Gamma[D] \Rightarrow E}{\Gamma[\Delta[q(B,C,D)]] \Rightarrow E}\,q\mathrm{L}}{\Gamma[\Delta[\Pi[q(A,C,D)]]] \Rightarrow E}\,\mathrm{Cut}$$

For each derivation of this form, there will necessarily be an equivalent derivation of the following form:

(20)
$$\dfrac{\dfrac{\Pi[A] \Rightarrow B \qquad \Delta[B] \Rightarrow C}{\Delta[\Pi[A]] \Rightarrow C}\,\mathrm{Cut} \qquad \Gamma[D] \Rightarrow E}{\Gamma[\Delta[\Pi[q(A,C,D)]]] \Rightarrow E}\,q\mathrm{L}$$

The replacement derivation still has a cut, but it is a cut of a lower degree (see, e.g., Jäger (2005) for a suitable definition of degree; the intuition is that the cut involves a lemma covering a smaller amount of material). As long as it is possible to replace any cut with one of strictly lower degree, cuts can be pushed upwards until they reach an axiom and can be entirely removed.

If qLR were a true rule of proof, there would be nothing more left to say concerning cut elimination. But because the qLR rule is two-sided—that is, it introduces the q connective on both the left and the right—then just as with Jäger's binding rules, we must also consider cutting an instance of qLR with another instance of qLR:

(21)
$$\dfrac{\dfrac{\Pi[A] \Rightarrow B}{\Pi[q(A,C,D)] \Rightarrow q(B,C,D)}\,q\mathrm{LR} \qquad \dfrac{\Gamma[B] \Rightarrow E}{\Gamma[q(B,C,D)] \Rightarrow q(E,C,D)}\,q\mathrm{LR}}{\Gamma[\Pi[q(A,C,D)]] \Rightarrow q(E,C,D)}\,\mathrm{Cut}$$

Once again, given the resources provided by the antecedents of the first derivation, we can reconfigure the proof with a cut of strictly lesser degree.

(22)
$$\dfrac{\dfrac{\Pi[A] \Rightarrow B \qquad \Gamma[B] \Rightarrow E}{\Gamma[\Pi[A]] \Rightarrow E}\,\text{Cut}}{\Gamma[\Pi[q(A, C, D)]] \Rightarrow q(E, C, D)}\,q\text{LR}$$

Thus the cut rule is admissible in DCOD, which is to say that any derivation in DCOD containing a cut can be replaced with an equivalent derivation that is cut-free.

3.7.2 *Adding Cuts Back in to Provide Direct Compositionality on Demand*

Now that we know that it is possible to eliminate cuts completely, we can consider adding back some cuts in order to provide direct compositionality.

Definition (*Directly composable*) Consider a specific derivation in DCOD with conclusion $\Gamma[(\Delta, \Pi)] \Rightarrow A$. That derivation is DIRECTLY COMPOSABLE just in case there exists a formula X and an equivalent derivation of the following form:

$$\dfrac{(\Delta, \Pi) \Rightarrow X \qquad \Gamma[X] \Rightarrow A}{\Gamma[(\Delta, \Pi)] \Rightarrow A}\,\text{Cut}$$

in which each of the antecedent derivations is also directly composable. (Two derivations count as equivalent if they have identical conclusion sequents with Curry–Howard labelings that are equivalent up to α and β equivalence.)

The first step in establishing that DCOD is directly compositional is the following simple but important observation:

Lemma (Categorization) Every structural constituent can be assigned a category.

Proof. Only two rules create complex constituents, that is structures of the form (Δ, Π), namely, $\backslash$L and /L. Consider $\backslash$L:

$$\dfrac{\Delta \Rightarrow A \quad \Gamma[B] \Rightarrow C}{\Gamma[(\Delta, A\backslash B)] \Rightarrow C}\,\backslash\text{L}$$

The rule itself provides us with a suitable choice for categorizing the constituent, namely, B, and we can replace the $\backslash$L step just given with the following reasoning:

$$\dfrac{\dfrac{\Delta \Rightarrow A \quad B \Rightarrow B}{(\Delta, A\backslash B) \Rightarrow B}\,\backslash\text{L} \qquad \Gamma[B] \Rightarrow C}{\Gamma[(\Delta, A\backslash B)] \Rightarrow C}\,\text{Cut}$$

In other words, using cut we can arrange for every application of $\backslash$L to have an axiom instance as its second antecedent.

The situation for /L is symmetric. □

The categorization lemma guarantees that each constituent can be associated with a syntactic category, but that is not enough to ensure full disclosure and semantic self-reliance. There are only two situations in DCOD in which self-reliance might be violated: when the binding rule $\uparrow$L links an anaphor inside the constituent in question with an antecedent outside the constituent; or when the quantification rule qL links a quantifier inside the constituent with a nuclear scope that properly contains the constituent in question. In each case, it will be necessary to adjust the syntactic category identified by the categorization lemma in order to align it with the principle of full disclosure.

3.7.2.1 *Full Disclosure for Binding* A violation of self-reliance due to binding has the following form:

(23)
$$\frac{\Sigma \Rightarrow A \quad \Pi[A][(\Delta[B], \Gamma)] \Rightarrow C}{\Pi[\Sigma][(\Delta[B^A], \Gamma)] \Rightarrow C} \uparrow\text{L}$$

This is just an application of $\uparrow$L where the anaphor (but not the antecedent) is inside the constituent (Δ, Γ).

First, assume that the derivations of the antecedents above the line are directly composable. Then there is some formula X such that we can cut out the constituent under consideration, (Δ, Γ), as a separate subproof.

(24)
$$\frac{\Sigma \Rightarrow A \quad \dfrac{(\Delta[B], \Gamma) \Rightarrow X \quad \Pi[A][X] \Rightarrow C}{\Pi[A][(\Delta[B], \Gamma)] \Rightarrow C}\,\text{Cut}}{\Pi[\Sigma][(\Delta[B^A], \Gamma)] \Rightarrow C} \uparrow\text{L}$$

Next, we use $\uparrow$LR to disclose the fact that the constituent contains an anaphor, replacing the category X with X^A. We can also bind the anaphor at the stage at which we form X^A, as long as we interchange the order of $\uparrow$L with cut:

(25)
$$\frac{\dfrac{(\Delta[B], \Gamma) \Rightarrow X}{(\Delta[B^A], \Gamma) \Rightarrow X^A}\,\uparrow\text{LR} \qquad \dfrac{\Sigma \Rightarrow A \quad \Pi[A][X] \Rightarrow C}{\Pi[\Sigma][X^A] \Rightarrow C}\,\uparrow\text{L}}{\Pi[\Sigma][(\Delta[B^A], \Gamma)] \Rightarrow C}\,\text{Cut}$$

Then the left-most antecedent of the cut inference constitutes a complete and self-contained proof that (Δ, Γ) is a constituent of category X^A, and the original derivation is directly composable.

The situation when the anaphor is inside Γ instead of Δ is closely analogous.

In constructing the DC derivation, an instance of $\uparrow$LR is cut against an instance of $\uparrow$L, exactly the sort of situation that the cut elimination theorem is at pains to eliminate. But this is exactly what we need in order to arrive at a self-contained analysis of the syntactic constituent under consideration.

3.7.2.2 *Full Disclosure for Quantification* A violation of self-reliance due to quantification has the following form:

(26)
$$\frac{\Pi[(\Delta[B],\Gamma)] \Rightarrow C \quad \Sigma[D] \Rightarrow E}{\Sigma[\Pi[(\Delta[q(B,C,D)],\Gamma]]] \Rightarrow E}\, q\mathrm{L}$$

This is just an application of qL where the quantifier and its scope are on different sides of the targeted constituent boundary. The argument proceeds exactly as for the binding case.

Once again, assume that the antecedents are directly composable, so that we can cut out the constituent under consideration, (Δ, Γ), as a separate subproof.

(27)
$$\frac{\dfrac{(\Delta[B],\Gamma) \Rightarrow X \quad \Pi[X] \Rightarrow C}{\Pi[(\Delta[B],\Gamma)] \Rightarrow C}\,\mathrm{Cut} \qquad \Sigma[D] \Rightarrow E}{\Sigma[\Pi[(\Delta[q(B,C,D)],\Gamma)]] \Rightarrow E}\, q\mathrm{L}$$

Next, we use qLR to disclose the fact that the constituent contains a quantifier, replacing the category X with $q(X, C, D)$. At the same stage, we can fix the scope of quantificational X with an application of qL, as long as we interchange the order of qL with cut:

(28)
$$\frac{\dfrac{(\Delta[B],\Gamma) \Rightarrow X}{(\Delta[q(B,C,D)],\Gamma) \Rightarrow q(X,C,D)}\, q\mathrm{LR} \quad \dfrac{\Pi[X] \Rightarrow C \quad \Sigma[D] \Rightarrow E}{\Sigma[\Pi[q(X,C,D)]] \Rightarrow E}\, q\mathrm{L}}{\Sigma[\Pi[(\Delta[q(B,C,D)],\Gamma)]] \Rightarrow E}\,\mathrm{Cut}$$

Thus the original derivation is directly composable.

Once again, we have an instance of qLR cut against an instance of qL, which we now recognize as the way full disclosure delivers semantic self-reliance.

Proposition (Direct Compositionality On Demand) For any valid sequent generated by DCOD, there is an equivalent swooping (cut-free) proof in which anaphors and their antecedents are coindexed by a single application of $\uparrow$L, and in which quantificational elements and their scope are related by a single application of qL. For each such derivation, there is another, equivalent, directly composable proof in which each constituent receives a self-contained analysis that obeys full disclosure and semantic self-reliance.

Proof. (sketch) For each constituent in the final conclusion structure, working from smallest to largest, and beginning with the axioms and working downwards, apply the categorization lemma and impose full disclosure. Since there are a strictly finite number of constituents, anaphors, and quantifiers, and since none of these operations introduces new constituents, the process is guaranteed to terminate. □

One of the pleasant properties of cut-elimination is that it leads to an algorithm for deciding whether a sequent is derivable. Since each application of a rule (other than cut) adds at least one logical connective, the conclusion will always be strictly more complex than the antecedents. One can simply try all applicable rules, stopping when a valid proof is found.

Since the on-demand theorem introduces cuts, it is worth considering whether there is an algorithm for finding the DC proof. Since there is an algorithm for finding a cut-free proof, and since there is an algorithm for constructing the equivalent DC proof, there is an algorithm for constructing the DC proof.

In fact, there is an alternative parsing strategy for logics that satisfy the DC on-demand property. Given a set of lexical categories, since the DC analysis of each constituent has the subformula property, the complete set of usable proofs for each constituent can be constructed in finite time. For a string of length n, the number of possible constituents is at most n^2, and for each constituent the time cost for trying all relevant cuts is proportional to n, giving a time cost of order n^3 in the length of the string.

3.8 Why the Display Property Alone is not Sufficient

There is a class of substructural logics called Display Logics that are relevant for type-logical grammar. Display Logics have what is called the Display Property. As Bernardi (2002: 33) puts it, “any particular constituent of a sequent can be turned into the whole of the right or the left side by moving other constituents to the other side [of the sequent symbol]”.

At first blush, the display property sounds like exactly what we want, since it guarantees that it is always possible to arrive at a syntactic category and a self-contained denotation for any syntactic constituent. Furthermore, there will be such an analysis for each distinct interpretation provided by the grammar.

Certainly no grammar can be directly compositional without having the display property. For instance, the standard QR story (e.g. as presented in Heim and Kratzer 1998) does not have the display property, and is not directly compositional, since there is no way to factor out a syntactic and semantic analysis of a verb phrase such as *saw everyone* (at least, not when the quantifier takes scope over an entire clause).

However, in this paper I am advocating the desirability of an even stronger property than the display property. I will first give two concrete examples of the sort of analyses the display property gives for constituency, then I will point out some shortcomings that motivate seeking a stronger property such as DCOD.

First, here is an analysis in which the display property provides a category and a denotation for the constituent *saw everyone* in *John saw everyone*:

(29)

$$\dfrac{\dfrac{\dfrac{\dfrac{\vdots}{[\text{as derived in §3.6}]}}{\boxed{\boxed{(np, ((np\backslash s)/np, q(np,s,s))) \Rightarrow s}}}}{((np\backslash s)/np, q(np,s,s)) \Rightarrow np\backslash s}\,\backslash\text{R} \qquad \dfrac{np \Rightarrow np \quad s \Rightarrow s}{(np, np\backslash s) \Rightarrow s}\,\backslash\text{L}}{\left(\begin{array}{c} np \\ \texttt{John}\end{array}, \left(\begin{array}{c}(np\backslash s)/np\\ \texttt{saw}\end{array}, \begin{array}{c} q(np,s,s)\\ \texttt{everyone}\end{array}\right)\right) \Rightarrow \begin{array}{c} s \\ \texttt{eo}(\lambda\texttt{x.saw(x)(j)})\end{array}}\,\text{Cut}$$

The peculiar thing about this derivation is that it begins with a complete derivation of the entire sentence: the boxed sequent is identical to the final conclusion sequent. Having the final conclusion before us in this way allows us to work backwards to figure out what the contribution of the verb phrase must have been, using a technique familiar from basic algebra:

$$\dfrac{\dfrac{p * (r * (y * x)) = s}{(r * (y * x)) = s/p}}{(y * x) = (s/p)/r}$$

Thus the logic of basic algebra has the display property.[2] In this instance of the display property approach, the constituent *saw everyone* has category $np\backslash s$, and although that hides the fact that the verb phrase is quantificational, it certainly works out to the correct final result. The problem is that this analysis for the constituent only works in situations in which the quantifier does not need to take wider scope. That is, the analysis of the constituent is sensitive to the material that surrounds it, violating the spirit of self-reliance.

In order to demonstrate this sensitivity to factors external to the constituent, consider an analogous treatment of *Someone claimed John saw*

[2] I borrow this analogy from Rajeev Goré.

everyone (with the interpretation on which *everyone* takes wide scope over *someone* in mind[3]). This time, the results are not as appealing:

(30)
$$
\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\vdots}{(np, ((np\backslash s)/s, (np, ((np\backslash s)/np, np)))) \Rightarrow s} \quad s \Rightarrow s}{(q(np, s, s), ((np\backslash s)/s, (np, ((np\backslash s)/np, np)))) \Rightarrow s}\,q\mathrm{L} \quad s \Rightarrow s}{(q(np, s, s), ((np\backslash s)/s, (np, ((np\backslash s)/np, q(np, s, s))))) \Rightarrow s}\,q\mathrm{L}}{((np\backslash s)/s, (np, ((np\backslash s)/np, q(np, s, s)))) \Rightarrow q(np, s, s)\backslash s}\,\backslash\mathrm{R}}{(np, ((np\backslash s)/np, q(np, s, s))) \Rightarrow ((np\backslash s)/s)\backslash((q(np, s, s)\backslash s))}\,\backslash\mathrm{R}}{\left(\begin{matrix}(np\backslash s)/np \\ \texttt{saw}\end{matrix}, \begin{matrix}q(np, s, s) \\ \texttt{everyone}\end{matrix}\right) \Rightarrow \begin{matrix} np\backslash(((np\backslash s)/s)\backslash((q(np, s, s)\backslash s))) \\ \lambda\mathtt{x}.\lambda\mathtt{R}.\lambda\mathscr{P}.\mathtt{eo}(\lambda\mathtt{y}.\mathscr{P}(\lambda\mathtt{z}.\mathtt{R}(\mathtt{saw}(\mathtt{y})(\mathtt{x}))(\mathtt{z})))\end{matrix}}\,\backslash\mathrm{R}
$$

Thus we have a category and a self-contained denotation for the constituent *saw everyone*. I have not filled in the details of how to cut this into a complete derivation of the original sentence, but that is easy (if tedious) to do.

Once again the starting point is a complete swooping derivation of the whole sentence. In general, the starting derivation must include the entire scope domain of every element within the target constituent. Then we exploit the rules of proof in order to pick off the elements in the structure until only the desired constituent remains.

And once again we have arrived at an analysis for the constituent that does not disclose the presence of a quantificational element within it (since the $q(np, s, s)$ present in the result category corresponds to the quantificational NP *someone* in matrix subject position, not to *everyone*, i.e. to a quantifier *external* to the constituent under study). This is certainly a possible approach; however, when comparing the two analyses of the same constituent (i.e. the two display-property factorizations of *saw everyone* in different sentences), some disturbing patterns emerge:

- *Complexity inversion.* Given a particular derivation, the analyses of smaller constituents tend to be more complex than the analyses of the larger constituents.
- *Unbounded complexity.* Given a particular constituent, but varying the larger expression in which it is embedded, there is no bound on the complexity of the syntactic and semantic analyses that the display property

[3] Many native speakers report that allowing a quantifier to take scope out of a tensed clause is difficult. Most people accept *Someone tried to see everyone*, which can be used to make the same point, although at the cost of spoiling the parallelism in the embedded clause.

associates with the constituent across the full range of its syntactic environments.

- *Spurious ambiguity.* Since any given constituent can have at most a finite number of distinct interpretations, it follows from unbounded complexity that there will be more distinct analyses than there are distinct interpretations.
- *Dependence on material outside of the constituent.* Since any given constituent will have at most finitely many parts, it follows from unbounded complexity that the analyses must incorporate some aspects of the expression that are external to the constituent in question.

In a nutshell, using the display property to assign analyses to embedded constituents in effect recapitulates the future derivation of the constituent. This means that it does not provide a completely satisfying characterization of the contribution of the constituent (at least, not of the contribution proper to the constituent itself). It is an indirect method at best, as if we tried to describe a hand exclusively by showing what it looks like inside a variety of gloves and mittens.

The analysis of quantification in Hendriks' (1993) Flexible types has similar properties, in that the category of a constituent containing a quantifier can be arbitrarily complex depending on how wide a scope the quantifier needs to take. See Barker (2005) for a discussion of Hendriks' system in the context of Jacobson's variable-free, directly compositional treatment of binding.

3.9 Structural Postulates

Unlike most type-logical grammars, the DCOD grammar as presented above does not contain any structural postulates. Most type-logical grammars contain structural postulates that at least make implication associative. There is strong linguistic motivation for making implication associative (at least under highly constrained circumstances), including analyses of so-called nonconstituent coordination (Steedman 1985; Dowty 1988, 1997) and various applications of function composition (e.g. Jacobson 1999), as well as many multimodal analyses (see Moortgat 1997 for a survey).

Structural postulates complicate the discussion of direct compositionality considerably. The reason is that direct compositionality is all about constituency, and the express purpose of structural postulates is to scramble constituent structure. For instance, here is a structural postulate that provides associativity:

$$\frac{(A, (B, C)) \Rightarrow D}{((A, B), C) \Rightarrow D}\ \text{ASSOC}$$

The double line indicates that the inference is valid in both directions: given the top sequent, infer the bottom one, or given the bottom sequent, infer the top one.

Without using this postulate, we can easily prove that [*John* [*saw* [*Mary*]]] is a sentence, that is $(np, ((np\backslash s)/np, np)) \Rightarrow s$. As a result of adding this postulate to the grammar, we can also prove that [[*John saw*] *Mary*], with the opposite constituency, is a sentence. The display property allows us to calculate an appropriate self-contained denotation for the pseudo-constituent *John saw*:

(31)

$$\dfrac{\dfrac{\dfrac{\dfrac{np \Rightarrow np \quad \dfrac{np \Rightarrow np \quad s \Rightarrow s}{(np, np\backslash s) \Rightarrow s}\backslash L}{(np, ((np\backslash s)/np, np)) \Rightarrow s}/L}{((np, (np\backslash s)/np), np) \Rightarrow s}\text{Assoc}}{\left(\begin{matrix} np \\ \texttt{John}\end{matrix}, \begin{matrix} (np\backslash s)/np \\ \texttt{saw}\end{matrix}\right) \Rightarrow \begin{matrix} s/np \\ \lambda\texttt{x.saw(x)(John)}\end{matrix}}/R \quad \dfrac{np \Rightarrow np \quad s \Rightarrow s}{(s/np, np) \Rightarrow s}/L}{\left(\left(\begin{matrix} np \\ \texttt{John}\end{matrix}, \begin{matrix} (np\backslash s)/np \\ \texttt{saw}\end{matrix}\right), \begin{matrix} np \\ \texttt{Mary}\end{matrix}\right) \Rightarrow \begin{matrix} s \\ \texttt{saw(Mary)(John)}\end{matrix}}\text{Cut}$$

As mentioned above, there is fairly compelling evidence motivating associativity as desirable from a linguistic point of view. For instance, adding associativity allows deriving Right Node Raising examples such as *John saw and Tom called Mary*, in which the alleged constituent *John saw* coordinates with *Tom called*.

At least for terminological purposes, it is convenient to discriminate between two notions of constituent: NATURAL constituency, as determined by the function/argument structure of the lexical predicates involved, vs. CALCULATED constituency, as derived from natural constituency via structural postulates. The display property will always provide an appropriate analysis for calculated constituents.

In contrast to my remarks in the previous section criticizing the result of using the display property to arrive at analyses of natural constituents, the display property technique seems to be exactly the right way to understand a calculated constituent such as *John saw*: it is the quotient of a complete sentence after factoring out the direct object.

Full disclosure of anaphors and quantifiers still applies to calculated constituents. For instance, in *John$_i$ saw, and his$_i$ mother called, Mary*, in a DC analysis the calculated constituent *his mother called* will have category $(s/np)^{np}$ with corresponding denotation.

3.10 Conclusions

Bernardi describes (a Natural Deduction version of) the qL rule given above, and remarks (2002: 103) that "in the multimodal setting ... the q connective of course cannot be a primitive connective". Instead, Bernardi suggests synthesizing q via a collection of multimodal logical and structural rules (see Moortgat (1997) for one concrete implementation of this strategy), and defining the swooping q as a "derived inference". This is perfectly coherent and feasible, of course; but it relegates the long-distance mode of analysis to a rule that is entirely redundant and eliminable ("admissible" in the logical jargon).

This paper explores the possibility of finding a grammar in which both views of constituency are simultaneously present, but each one of whose rules is indispensable. For instance, unlike Bernardi's derived inference rule for q, none of the rules in the DCOD logic given above is admissible. That is, eliminating any rule other than cut would reduce the number of valid sequents. In particular, it is only possible to prove the sequent

(32) $$\left(\begin{matrix} (np\backslash s)/np \\ \texttt{saw} \end{matrix} , \begin{matrix} q(np, s, s) \\ \texttt{everyone} \end{matrix} \right) \Rightarrow \begin{matrix} q(np\backslash s, s, s) \\ \lambda\texttt{f.everyone}(\lambda\texttt{x.f(saw(x))}) \end{matrix}$$

using qLR, so qLR is not admissible.

There are many pressures on the design of a grammar, and I do not expect that any system based on DCOD will serve all purposes. Rather, I offer DCOD here as an example showing that it is possible to reconcile the local and long-distance aspects of the syntax–semantics interface within a single unified grammar. With any luck, there will be other grammatical systems that can semantically link distant elements directly, yet still provide complete, self-contained constituent analyses with full disclosure: long-distance linking, but with direct compositionality on demand.

References

Barker, C. 2002. 'Continuations and the nature of quantification', *Natural Language Semantics*, 10: 211–42.

—— 2005. 'Remark on Jacobson 1999: Crossover as a local constraint', *Linguistics and Philosophy*, 28: 447–72.

——— and Shan, C.-C. In press. 'Types as graphs: Continuations in Type-Logical Grammar', *Journal of Logic, Language, and Information*.

Bernardi, R. 2002. Reasoning with polarity in categorial type logic. Ph. D. dissertation, Utrecht Institute of Linguistics, Utrecht University.

Dowty, D. 1988. 'Type raising, functional composition, and non-constituent coordination', in R. Oehrle, E. Bach, and D. Wheeler (eds), *Categorial Grammars and Natural Language structures*. Dordrecht: Reidel. 153–98.

——— 1997. 'Non-constituent coordination, wrapping, and multimodal categorial grammars', in M. L. Dalla Chiara, K. Doets, D. Mundici, and J. van Benthem (eds), *Structures and Norms in Science: Volume Two of the Tenth International Congress of Logic, Methodology and Philosophy of Science, Florence, August 1995*. Dordrecht: Kluwer. 347–68.

——— 2000. 'The dual analysis of adjuncts/complements in Categorial Grammar', ZAS Papers in Linguistics, 17.

Goré, R. 1998. 'Gaggles, Gentzen and Galois: How to display your favourite substructural logic', *Logic Journal of the IGPL*, 6: 669–94.

Heim, I. and Kratzer, A. 1998. *Semantics in Generative Grammar*. Oxford: Blackwell.

Hendriks, H. 1993. *Studied Flexibility*. Amsterdam: ILLC Dissertation Series.

Jacobson, P. 1999. 'Towards a variable-free semantics', *Linguistics and Philosophy*, 22: 117–84.

Jäger, G. 2005. *Anaphora and Type Logical Grammar*. Dordrecht: Springer. [Page numbers refer to the 2001 manuscript version.]

Lambek, J. 1958. 'The mathematics of sentence structure', *American Mathematical Monthly*, 65: 154–70.

Montague, R. 1974. 'The proper treatment of quantification in ordinary English', in R. Thomason (ed), *Formal Philosophy: Selected Papers of Richard Montague*, New Haven, CT: Yale University Press, 247–70.

Moortgat, M. 1997. 'Categorial type logics', in J. van Benthem and A. ter Meulen (eds), *Handbook of Logic and Language*. Amsterdam: Elsevier Science. 93–177.

Restall, G. 2000. *An Introduction to Substructural Logics.* London: Routledge.

Shan, C.-C. and Barker, C. 2006. 'Explaining crossover and superiority as left-to-right evaluation', *Linguistics and Philosophy*, 29: 91–134.

Steedman, M. 1985. 'Dependency and coordination in the grammar of Dutch and English', *Language*, 61: 523–68.