Anaphora and Quantification in Categorial Grammar

Gerhard Jäger

Zentrum für Allgemeine Sprachwissenschaft
Jägerstr. 10/11
D-10117 Berlin
Germany
jaeger@zas.gwz-berlin.de

Abstract. The paper proposes a type logical reformulation of Jacobson's ([9]) treatment of anaphoric dependencies in Categorial Grammar. To this end, the associative Lambek Calculus is extended with a new connective. In the first part, its proof theory is introduced and the logical properties of the resulting calculus are discussed. In the second part, this system is applied to several linguistic phenomena concerning the interaction of pronominal anaphora, VP ellipsis and quantifier scope. Finally, a possible extension to cross-sentential anaphora is considered.

1 Introduction

Anaphora phenomena are a challenge to Categorial Grammar (CG) for several reasons. To start with, the standard treatment of anaphoric pronouns as variables is not really viable in CG due to its essentially variable free design. This is self-evident in Combinatory Categorial Grammar (CCG) since here semantic operations are by definition restricted to combinators (in the sense of Combinatory Logic). Thus in CCG the grammar does not provide variable binding devices. Researchers working in the framework of Type Logical Grammar usually don't stress the variable free setup of the syntax-semantics interface, but things are not different here than in CCG. This becomes clear if one acknowledges the fact that the type logics used in CG can be seen as fragments of positive Intuitionistic Logic. Thus all type logical proof terms (= admissible semantic operations) can be expressed as combinatory terms (using only Curry's **S** and **K**).

If one accepts that anaphora does not involve variables, one is faced with another problem. Virtually by definition, anaphora involves a multiple use of semantic resources. To put it the other way round, anaphors are expressions that use a resource without consuming it. In resource conscious logics, re-use of resources is characteristic for systems that contain the structural rule of contraction, like Relevance Logic or Intuitionistic Logic. Such logics lack the finite reading property, i.e one and same valid sequent may have infinitely many proof terms. A simple but telling example is the sequent $p \to p \Rightarrow p \to p$. All terms $\lambda x.f^n x$ for finite n are valid relevant or intuitionistic proof terms. Natural language utterances are at most finitely ambiguous though. The finite reading property is thus essential for an adequate grammar logic. So a type logical treatment of anaphora should be strong enough to admit multiple use of resources, but weak enough to have the finite reading property.

In this paper, we first propose an extension of the Lambek Calculus **L** ([13]) that overcomes this problem. In the sequel, we illustrate its linguistic applications. After discussing the treatment of anaphoric pronouns, we study the interaction of this system with Moortgat's account of quantification ([15, 16]). Next we extend the analysis to VP ellipsis (VPE henceforth) and discuss the interplay of pronominal anaphora and quantification with VPE. Finally we sketch how cross-sentential anaphora can be handled in this setup.

2 The Logic L|

2.1 Background: Jacobson's Proposal

In a series of publications ([6–9]), Pauline Jacobson has shown how pronominal anaphora can be handled successfully in CCG while maintaining the variable freeness of this framework. The basic intuition of her proposal is the idea that anaphoric expressions denote functions from antecedent meanings to contextually determined meanings. Applied to anaphoric pronouns, this means that they denote the identity function on individuals. Due to the strict category-to-type correspondence in CG, this must be reflected in the syntactic category. To this end, she extends the inventory of type forming connectives of CCG with a third slash $|$. A sign of category $A|B$ is an item that needs an antecedent of category B to behave like a sign of category A. Accordingly, its denotation will be a function from B-denotations to A-denotation. In other words, $|$ creates a functional category like the other two slashes. Under this account, anaphoric pronouns have category $N|N$ and denote the identity function on individuals. The non-local nature of anaphora is taken care of by a generalized version of function composition which makes arbitrarily large portions of syntactic material transparent for anaphoric dependencies. The job of connecting an anaphor to its antecedent—multiplication of a resource in the constructive jargon—is taken over by a modified version of Curry's **S** that she calls **Z**.

The logic that is introduced in the next subsection can be seen as a translation of Jacobson's ideas into the type logical architecture. In particular, we will adopt Jacobson's slash and its intuitive interpretation, including the functional semantics of anaphors. Furthermore all relevant combinators of her system are theorems of our logic. Nonetheless—due to the fact that the Lambek Calculus is completely associative but CCG isn't—the empirical predictions of the two approaches do not completely coincide. In Sect. 5 we will try to demonstrate that the move from combinatory to type logical design is advantageous from a descriptive point of view.

2.2 Sequent Presentation

The set of formulas $\mathcal{F}$ of the logic **L**| is defined as the closure of some set $\mathcal{A}$ of atomic formulas under the binary operations $\backslash, \bullet, /$ and $|$.

The sequent presentation of **L**| extends Lambek's ([13]) syntactic calculus **L** with the logical rules for the third slash "|". Sequent rules are augmented with Curry-Howard terms.

$$\frac{}{x : A \Rightarrow x : A}\, id$$

$$\frac{X \Rightarrow M : A \qquad Y, x : A, Z \Rightarrow N : B}{Y, X, Z \Rightarrow N[x \leftarrow M] : B}\, Cut$$

$$\frac{X, x : A, y : B, Y \Rightarrow M : C}{X, z : A \bullet B, Y \Rightarrow M[x \leftarrow (z)_0, y \leftarrow (z)_1] : C}\, \bullet L$$

$$\frac{X \Rightarrow M : A \qquad Y \Rightarrow N : B}{X, Y \Rightarrow \langle M, N \rangle : A \bullet B}\, \bullet R$$

$$\frac{X \Rightarrow M : A \qquad Y, x : B, Z \Rightarrow N : C}{Y, y : B/A, X, Z \Rightarrow N[x \leftarrow (yM)] : C}\, /L$$

$$\frac{X, x : A \Rightarrow M : B}{X \Rightarrow \lambda x.M : B/A}\, /R \qquad \%\ X \text{ non-empty}$$

$$\frac{X \Rightarrow M : A \qquad Y, x : B, Z \Rightarrow N : C}{Y, X, y : A \setminus B, Z \Rightarrow N[x \leftarrow (yM)] : C}\, \setminus L$$

$$\frac{x : A, X \Rightarrow M : B}{X \Rightarrow \lambda x.M : A \setminus B}\, \setminus R \qquad \%\ X \text{ non-empty}$$

$$\frac{Y \Rightarrow M : B \qquad X, x : B, Z, y : A, W \Rightarrow N : C}{X, Y, Z, z : A|B, W \Rightarrow N[x \leftarrow M][y \leftarrow (zM)] : C}\, |L$$

$$\frac{x : B, y : p, X \Rightarrow \langle M, y, N \rangle : B \bullet p \bullet A \qquad x : B, X \Rightarrow \langle M, N \rangle : B \bullet A}{X \Rightarrow \lambda x.N : A|B}\, |R$$

% X non-empty
% p is atomic and does not occur in A, B, X
% $M =_{\alpha\beta\eta} x$

A few comments are in order. Even though the Jacobsonian slash has a functional semantics and thus resembles the standard categorial slashes, it has a different nature since it cannot be reconstructed as a residuation operation of some product operator. In the presence of cut, rule $|L$ is equivalent to the axioms

(1) a. $x : A, y : B|A \Rightarrow \langle x, (yx)\rangle : A \bullet B$
b. $x : A, y : B, z : C|A \Rightarrow \langle x, y, (zx)\rangle : A \bullet B \bullet C$

Together with the rule of proof $|R$ this expresses the intuition that a sign has category $A|B$ if and only if it behaves like a sign of category A in the presence of an antecedent of type B. Anaphor and antecedent may, but need not be adjacent. This motivates the two premises in $|R$. Since the atom p in rule $|R$ does not occur anywhere else, it behaves like a variable over the material between anaphor and antecedent. Furthermore it has to be ensured that neither the antecedent nor the material in between are affected by anaphora resolution. This motivates the constraint of the Curry-Howard terms in the premises of $|R$. So labeling is not just a book keeping device but a genuine restriction here.

$\mathbf{L}|$ has the desired proof theoretic properties. To start with, cut elimination is possible.

Theorem 1 (Cut Elimination)
Cut is admissible in $\mathbf{L}|$.

Proof. The proof relies on the following two lemmas:

Lemma 1 If Π is a correct proof and p an atomic formula that occurs in the conclusion of Π, then $\Pi[p \leftarrow B]$ is a correct proof as well, where $\Pi[p \leftarrow B]$ is the result of replacing all occurrences of p in Π by B.

Proof. By Induction over the complexity of Π. □

Lemma 2 With $X^\bullet$ we refer to the formula that results from replacing all commas in X by $\bullet$. Then it holds that

$$\vdash X \Rightarrow X^\bullet$$

Proof. Induction over the length of X. □

The cut elimination algorithm follows the one given in [13]. Principal cut for $|$ (see Fig. 1 for the case where the Z in $|L$ is non-empty and Fig. 2 for the case where it is empty) apparently poses a problem since it may replace one cut by three cuts of a higher degree. Nevertheless it can be shown that the algorithm always terminates, for the following reason: We call a proof *special* iff all atoms occurring in it are pairwise different unless a sequent rule requires them to be identical. Clearly every proof can be transformed into a special proof by renaming of atoms. If we restrict our attention to special proofs, the principal cut for $|$ reduces the total number of atoms occurring in the proof (since p completely disappears). This parameter is not increased by any other cut elimination step. This guarantees that cut elimination eventually terminates. The restriction to special proofs is no real restriction since we can always transform a given proof into a special proof via renaming, perform cut elimination and finally reverse the renaming. □

$$\dfrac{\dfrac{\dfrac{\Pi_1}{B, p, W \Rightarrow B \bullet p \bullet A} \qquad \dfrac{\Pi_2}{B, W \Rightarrow B \bullet W}}{W \Rightarrow A|B}\,|R \qquad \dfrac{\dfrac{\Pi_3}{Y \Rightarrow B} \qquad \dfrac{\Pi_4}{X, B, Z, A, U \Rightarrow C}}{X, Y, Z, A|B, U \Rightarrow C}\,|L}{X, Y, Z, W, U \Rightarrow C}\,Cut$$

$$\rightsquigarrow$$

$$\dfrac{\dfrac{}{Z \Rightarrow Z^{\bullet}}\,lm\ 2 \qquad \dfrac{\dfrac{\dfrac{\Pi_3}{Y \Rightarrow B} \qquad \dfrac{\Pi_1[p \leftarrow Z^{\bullet}]}{B, Z^{\bullet}, W \Rightarrow B \bullet Z^{\bullet} \bullet A}}{Y, Z^{\bullet}, W \Rightarrow B \bullet Z^{\bullet} \bullet A}\,Cut \qquad \dfrac{\dfrac{\dfrac{\Pi_4}{X, B, Z, A, U \Rightarrow C}}{\vdots}\,\bullet L}{X, B \bullet Z^{\bullet} \bullet A, U \Rightarrow C}\,\bullet L}{X, Y, Z^{\bullet}, W, U \Rightarrow C}\,Cut}{X, Y, Z, W, U \Rightarrow C}\,Cut$$

Fig. 1. Principal cut for $|$, Z non-empty

$$\dfrac{\dfrac{\dfrac{\Pi_1}{B, p, W \Rightarrow B \bullet p \bullet A} \qquad \dfrac{\Pi_2}{B, W \Rightarrow B \bullet W}}{W \Rightarrow A|B}\,|R \qquad \dfrac{\dfrac{\Pi_3}{Y \Rightarrow B} \qquad \dfrac{\Pi_4}{X, B, A, U \Rightarrow C}}{X, Y, A|B, U \Rightarrow C}\,|L}{X, Y, W, U \Rightarrow C}\,Cut$$

$$\rightsquigarrow$$

$$\dfrac{\dfrac{\dfrac{\Pi_3}{Y \Rightarrow B} \qquad \dfrac{\Pi_2}{B, W \Rightarrow B \bullet A}}{Y, W \Rightarrow B \bullet A}\,Cut \qquad \dfrac{\dfrac{\Pi_4}{X, B, A, U \Rightarrow C}}{X, B \bullet A, U \Rightarrow C}\,\bullet L}{X, Y, W, U \Rightarrow C}\,Cut$$

Fig. 2. Principal cut for $|$, Z empty

Although $|R$ lacks the subformula property, every sequent rule except cut increases complexity in the sense defined below:

Definition 1

1. $d(p) = 1$ % p atomic
2. $d(A \circ B) = d(A) + d(B) + 1$, where $\circ$ ranges over $\bullet$, $\backslash$ and $/$
3. $d(A|B) = d(A) + 2d(B) + 5$
4. $d(A_1, \ldots, A_n \Rightarrow B) = d(A_1) + \cdots d(A_n) + d(B)$

Theorem 2 The proof search space for $\mathbf{L}|$ is finite.

Proof. Cut free proof search reduces complexity, and at every point in proof search, only finitely many sequent rules are applicable. □

Corollary 1 (Decidability) **L**| is decidable.

Proof. Immediately from Theorem 2. □

Corollary 2 (Finite Reading Property) **L**| has the finite reading property.

Proof. Immediately from Theorem 2. □

2.3 Natural Deduction

The sequent system is indispensable since it guarantees decidability, but for practical purposes it is rather awkward. A presentation in natural deduction (ND) format is better suited to present concrete derivations. Besides, it has an appealing allusion to the tree format linguists are used to.

We start with a sequent style presentation of the natural deduction system (Fig. 3). Besides the identity rule and the cut rule (which are identical to the corresponding rules in the sequent system and therefore omitted), we have an introduction rule and an elimination rule for each connective.

$$\frac{}{x : A, y : B \Rightarrow \langle x, y\rangle : A \bullet B} \bullet I \qquad \frac{X \Rightarrow M : A \bullet B \qquad Y, x : A, y : B, Z \Rightarrow N : C}{Y, X, Z \Rightarrow N[x \leftarrow (M)_0][y \leftarrow (M)_1] : C} \bullet E$$

$$\frac{x : A, X \Rightarrow M : B}{X \Rightarrow \lambda x.M : A \setminus B} \setminus I \qquad \frac{X \Rightarrow M : A \qquad Y \Rightarrow N : A \setminus B}{X, Y \Rightarrow (NM) : B} \setminus E$$

$$\frac{X, x : A \Rightarrow M : B}{X \Rightarrow \lambda x.M : B/A} /I \qquad \frac{X \Rightarrow M : A/B \qquad Y \Rightarrow N : B}{X, Y \Rightarrow (MN) : A} /E$$

$$\frac{X \Rightarrow M : A \qquad Y \Rightarrow N : B \qquad Z \Rightarrow O : C|A}{X, Y, Z \Rightarrow \langle M, N, (OM)\rangle : A \bullet B \bullet C} |E$$

$$\frac{x : B, y : p, X \Rightarrow \langle N, y, M\rangle : B \bullet p \bullet A \qquad X : B, X \Rightarrow \langle N, M\rangle : B \bullet A}{X \Rightarrow \lambda x.M : A|B} |I$$

p not occurring in A, B, X
$M =_{\alpha\beta\eta} x$

Fig. 3. Natural Deduction **L**|

Natural deductions are more conveniently carried out in tree form. The building blocks are given in Fig. 4. Note that a complete deduction always ends in a single conclusion, despite the fact that • elimination and | elimination have multiple conclusions. For simplicity we combined the |-elimination rule with two subsequent applications of the •-elimination rule, thus removing the products in the conclusion of |-elimination. The parentheses in the premises of | introduction indicate that these premises must be derivable both with and without the material in parentheses. Note that the only structural constraint on anaphora

$$\frac{M:A \qquad N:B}{\langle M,N\rangle : A\bullet B}\bullet I \qquad\qquad \frac{M:A\bullet B}{(M)_0:A \qquad (M)_1:B}\bullet E$$

$$\cfrac{\cfrac{\begin{array}{cc}\overline{x:A}^{\,i} & \vdots\vdots \\ \vdots & \vdots\vdots\end{array}}{M:B}}{\lambda x.M : A\setminus B}\setminus I, i \qquad\qquad \frac{M:A \qquad N:A\setminus B}{(NM):B}\setminus E$$

$$\cfrac{\cfrac{\begin{array}{cc}\vdots\vdots & \overline{x:A}^{\,i} \\ \vdots\vdots & \vdots\end{array}}{M:B}}{\lambda x.M : B/A}/I, i \qquad\qquad \frac{M:A/B \qquad N:B}{(MN):A}/E$$

$$\cfrac{\cfrac{\begin{array}{ccc}\overline{x:B}^{\,i} & \overline{(y:p)}^{\,i} & \vdots \\ \vdots & \vdots & \vdots\end{array}}{\langle x,(y,)M\rangle : B\bullet(p\bullet)A}}{\lambda x.M : A|B}|I, i \qquad\qquad \frac{M:B \quad \cdots \quad N:A|B}{M:B \quad \cdots \quad (NM):A}|E$$

Fig. 4. Natural deduction in tree format

resolution is the requirement that the antecedent precede the anaphor. No command relations of whatever kind are involved.

For better readability and to stress the similarity to conventional coindexing of constituents, we simplify the notation for $|E$ somewhat (see Fig. 5). When

$$[M:B]_i \quad \cdots \quad \frac{[N:A|B]_i}{(NM):A}|E$$

Fig. 5. Simplified notation for $|E$

working with Natural Deduction in tree format, it has to be kept in mind that the domain of rule applications are complete proof trees, not arbitrary subtrees of a proof tree. This is particularly important when $|E$ is involved. Both parts of an anaphoric link belong to one and the same tree. Therefore it is illicit to let another rule operate on a subtree that includes one part of the anaphoric link and excludes the other. (An example of a violation of this constraint is given in Fig. 6. Here the premise $x : A$ is connected with the premise $z : C|A$ by an anaphoric link, i.e an application of $|E$, but the scope of $\setminus I$ includes the former and excludes the latter.) This blocks derivations where the proof term of the conclusion contains free variables that do not correspond to any premise.

$$\dfrac{\dfrac{\dfrac{\dfrac{}{[x:A]_i}\,1 \qquad \dfrac{}{y:B}}{\langle x,y\rangle : A\bullet B}\,\bullet I}{\lambda x.\langle x,y\rangle : A\setminus(A\bullet B)}\,\setminus I,1 \qquad \dfrac{\dfrac{}{[z:C|A]_i}}{(zx):C}\,|E}{\langle \lambda x.\langle x,y\rangle,(zx)\rangle : (A\setminus(A\bullet B))\bullet C}\,\bullet I$$

Fig. 6. An illicit Natural Deduction derivation

3 Pronouns and Quantification

Following Jacobson, we assume that pronouns like *he* have category $N|N$ and denote the identity function on individuals, i.e the associated semantic term is $\lambda x.x$. For a simple example like

(2) John said he walked

where the only potential antecedent of the pronoun is a proper noun, we have the two possible derivations shown in Fig. 7, corresponding to the coreferential and the free reading of the pronoun.

Things become somewhat more involved when we consider possible interaction of anaphora resolution with hypothetical reasoning. Nothing prevents us from using a hypothesis of the appropriate type as antecedent for anaphora resolution. For example, in the VP

(3) said he walked

the pronoun can be linked to the subject argument place of the VP, as Fig. 8 demonstrates. This VP can for instance be combined with a subject relative pronoun to yield the relative clause *who said he walked*. Another type of construction where this kind of derivation is crucial are sloppy readings of VPE that will be discussed below.

Binding to hypothetical antecedents is not restricted to slash introduction rules. Another obvious case in point is the interaction of anaphora with quantification. Here we adopt the type logical treatment of quantification that was proposed by Michael Moortgat (see for instance [15]). To repeat the basic ingredients very briefly, Moortgat proposes a new three place type constructor q. A sign a has category $q(A, B, C)$ iff replacing a sign of category A by a in the context of a super-constituent of type B, the result will have category C. This is reflected by the Natural Deduction rules in Fig. 9. The elimination rule involves hypothetical reasoning and can thus lead to binding of anaphors. Let us consider the example

$$\dfrac{\dfrac{\textit{John}}{[\mathrm{J} : N]_i}\,lex \qquad \dfrac{\dfrac{\textit{said}}{\mathrm{SAY} : N \backslash S/S}\,lex \qquad \dfrac{\dfrac{\dfrac{\textit{he}}{[\lambda x.x : N|N]_i}\,lex}{\mathrm{J} : N}\,|E \qquad \dfrac{\textit{walked}}{\mathrm{WALK} : N \backslash S}\,lex}{\mathrm{WALK}\ \mathrm{J} : S}\,\backslash E}{\mathrm{SAY}(\mathrm{WALK}\ \mathrm{J}) : N \backslash S}\,/E}{\mathrm{SAY}(\mathrm{WALK}\ \mathrm{J})\mathrm{J} : S}\,\backslash E$$

$$\dfrac{\dfrac{\dfrac{\dfrac{}{x : [N]_i}\,1 \qquad \dfrac{}{y : p}\,1}{\begin{array}{c}N \bullet p\\ \langle x, y\rangle\end{array}}\,\bullet I \qquad \dfrac{\dfrac{\textit{John}}{\mathrm{J} : N}\,lex \qquad \dfrac{\dfrac{\textit{said}}{\mathrm{SAY} : N \backslash S/S}\,lex \qquad \dfrac{\dfrac{\dfrac{\textit{he}}{\begin{array}{c}[N|N]_i\\ \lambda x.x\end{array}}\,lex}{x : N}\,|E \qquad \dfrac{\textit{walked}}{\mathrm{WK} : N \backslash S}\,lex}{\mathrm{WK}\ x : S}\,\backslash E}{\mathrm{SAY}(\mathrm{WK}\ x) : N \backslash S}\,/E}{\mathrm{SAY}(\mathrm{WK}\ x)\mathrm{J} : S}\,\backslash E}{\langle x, y, \mathrm{SAY}(\mathrm{WK}\ x)\mathrm{J}\rangle : N \bullet p \bullet S}\,\bullet I}{\lambda x.\mathrm{SAY}(\mathrm{WK}\ x)\mathrm{J} : S|N}\,|I, 1$$

Fig. 7. Derivations of *John said he walked*

$$\dfrac{\dfrac{\dfrac{}{[x : N]_i}\,1 \qquad \dfrac{\dfrac{\textit{said}}{\mathrm{SAY} : N \backslash S/S}\,lex \qquad \dfrac{\dfrac{\dfrac{\textit{he}}{[\lambda x.x : N|N]_i}\,lex}{x : N}\,|E \qquad \dfrac{\textit{walked}}{\mathrm{WALK} : N \backslash S}\,lex}{\mathrm{WALK}\ x : S}\,\backslash E}{\mathrm{SAY}(\mathrm{WALK}\ x) : N \backslash S}\,/E}{\mathrm{SAY}(\mathrm{WALK}\ x)x : S}\,\backslash E}{\lambda x.\mathrm{SAY}(\mathrm{WALK}\ x)x : S}\,\backslash I, 1$$

Fig. 8. Derivation of *said he walked*

$$\dfrac{\dfrac{\begin{array}{ccc}\vdots & \dfrac{x : q(A, B, C)}{y : A}\,i & \vdots\\ \vdots & \vdots & \vdots\end{array}}{\alpha : B}}{x(\lambda y.\alpha) : C}\,qE, i \qquad\qquad \dfrac{x : A}{\lambda y.yx : q(A, B, B)}\,qI$$

Fig. 9. Natural Deduction rules for $q(A, B, C)$

(4) Everybody said he walked

Quantifiers like *everybody* have category $q(N, S, S)$, so in the course of scoping the quantifier, a hypothesis of category N is temporarily introduced. This hypothesis can in turn serve as antecedent of *his*, as illustrated in Fig. 10.

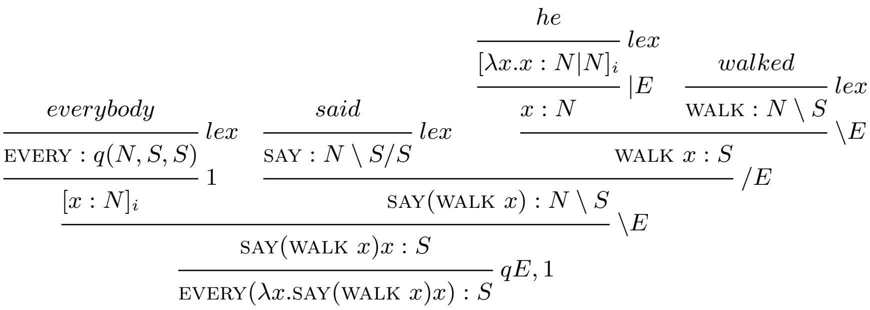

Fig. 10. Derivation of *Everybody said he walked*

If we reverse the order of the quantifier and the pronoun as in (5), the derivation of a bound reading will fail, even though the pronoun is in the scope of the quantifier.

(5) *He$_i$ said everybody$_i$ walked

This configuration—a Strong Crossover violation—is ruled out since the hypothesis that temporarily replaces the quantifier does not precede the pronoun. Thus |-elimination cannot be applied.

As any ND rule, q-elimination can only be applied to complete trees. If the hypothetical N that is used in qE serves as the antecedent of a pronoun, this pronoun must be in the scope of qE. Linguistically speaking, this means that a bound pronoun is always in the scope of its binder. This excludes for instance a wide scope reading of the indefinite object in (6) if the pronoun is bound by the subject.

(6) Every man saw a friend of his

The way the present system excludes such readings is similar to the one proposed in [19], even though the treatment of pronouns in general is different.

Finally it should be stressed that the only constraints on pronoun binding here are the requirements that 1. the quantifier precedes the pronoun, and 2. the pronoun is in the scope of the quantifier. So the derivation of a construction like (7), where the binding quantifier does not c-command the pronoun under the standard conception of constituency, does not differ substantially from the previous example (Fig. 11).

(7) Everybody's mother loves him

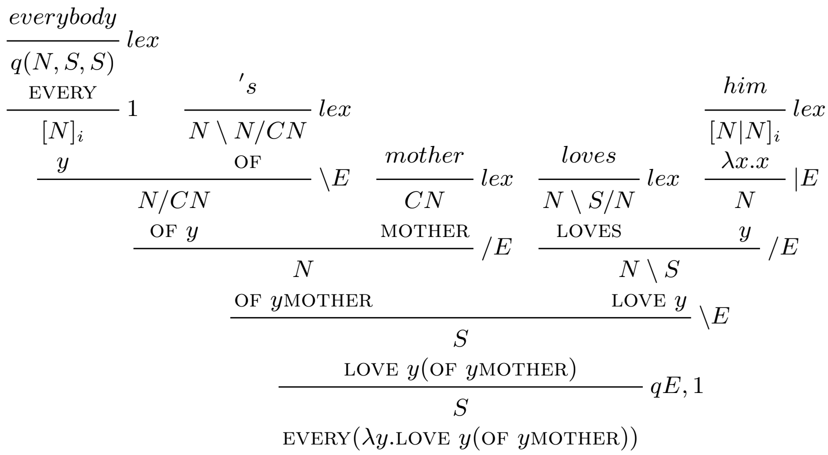

Fig. 11. Everybody's mother loves him

Again, if we change the order of pronoun and quantifier, the derivation will fail since the precedence requirement for $|E$ is not met.

(8) *His mother loves everybody

So the precedence requirement also accounts for Weak Crossover violations.

As mentioned above, binding to hypothetical antecedents is not restricted to quantification. Another obvious case in point is *wh*-movement. There are several proposals for a type logical treatment of this complex around in the literature (see for instance [14, 18]). Despite the differences in detail, they share the assumption that a hypothesis is put into the "base position" of *wh*-movement that is later discharged and bound by the operator. So we correctly predict the same patterns concerning the interaction of binding and scope and with respect to Crossover phenomena. Arguably, association with focus involves hypothetical reasoning as well (see [10] for an attempt to spell this idea out in a multi-modal framework). Accordingly, we find bound readings and Crossover effects if the antecedent of a pronoun is focused. (The latter observation was initially made in [1]. Example (10) is from [20].)

(9) Only JOHN hates his mother

(10) a. We only expect HIM to be betrayed by the woman he loves
b. We only expect him to be betrayed by the woman HE loves

The pronoun in sentence (9) can be bound, i.e. the sentence can mean *John is the only x such that x loves x's mother*. Example (10) illustrates that binding by focus displays Weak Crossover effects. Sentence (10a) has a reading saying that the referent of *him* is the only person x such that we expect x to be betrayed by the woman x loves. No such reading is available in (10b).

4 VP Ellipsis

This treatment of anaphora can straightforwardly be extended to VPE. Ignoring matters of tense and mood, we treat the stranded auxiliary in the second conjunct of constructions like (11) as a proform for VPs.

(11) John walked, and Bill did too

So *did* will be assigned the category $(N \setminus S)|(N \setminus S)$ and the meaning $\lambda P.P$, i.e the identity function on properties. The derivation for (11) is given in Fig. 12 (we also ignore the contribution of *too* since it is irrelevant for the semantics of VPE, though not for the pragmatics).

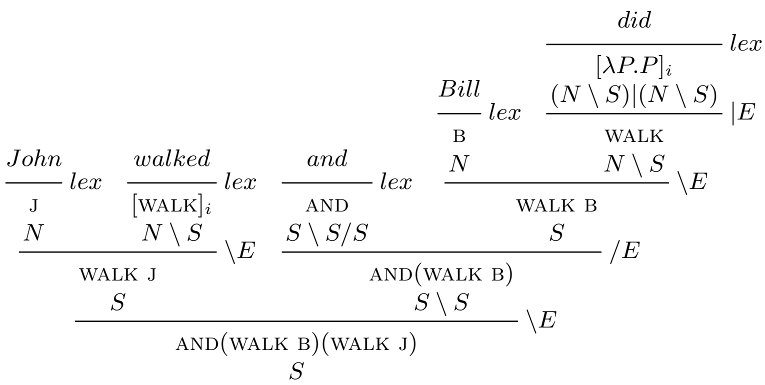

Fig. 12. John walked, and Bill did (too)

What makes VPE an interesting topic is of course its complex interaction with pronominal anaphora and quantification. Due to limitation of space, we cannot give an in-depth investigation of these issues here. Instead we will content ourselves with a discussion of some of the most frequently discussed examples from the literature.

The first non-trivial issue in this connection is the well-known strict/sloppy ambiguity in constructions like (12).

(12) John revised his paper, and Harry did too

The crucial step for the derivation of the sloppy reading is already given for an analogous example in Fig. 8: the pronoun is bound to the subject argument place of the source VP. From this we can continue the derivation completely in parallel to Fig. 12, and we end up with the meaning AND(SAY(WALK B)B) (SAY(WALK J)J). Crucially, here the pronoun was bound by a hypothetical antecedent. Of course it is also licit to bind to pronoun to the actual subject *John*

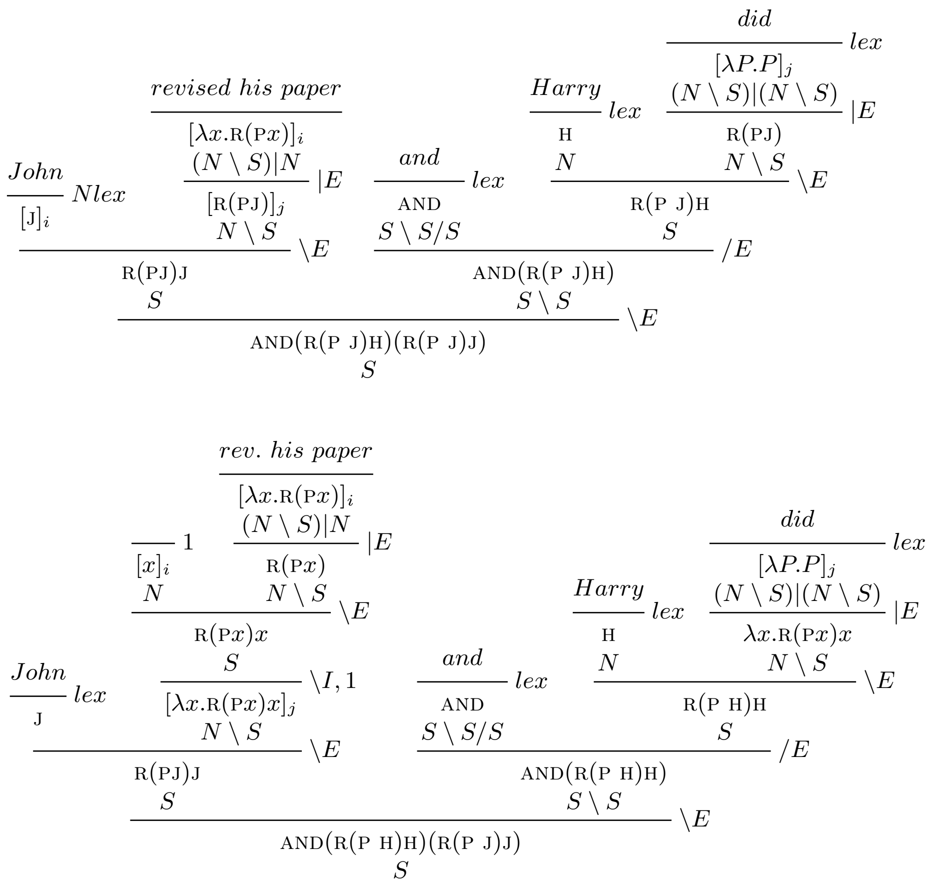

Fig. 13. Derivation of the strict and the sloppy reading of (12)

and then doing ellipsis resolution, which results in the strict reading. The derivation of both readings is given in Fig. 13.
Next we would like to draw attention to a kind of ambiguity that arises from the interplay of quantification and VPE. Consider the following example.

(13) a. John met everybody before Bill did
b. John met everybody before Bill met everybody
c. John met everybody before Bill met him

As Ivan Sag observed in [21], constructions like (13a) are ambiguous between a reading synonymous to (13b) and one synonymous to (13c). Under the present approach, reading (13b) arises if the quantifier is scoped before ellipsis resolution takes place. If scoping is postponed until after ellipsis resolution, the antecedent of the ellipsis still contains a hypothetical N, and accordingly the quantifier binds two occurrences of the corresponding variable. Fig. 14 gives the derivations for the source VPs of the two readings.

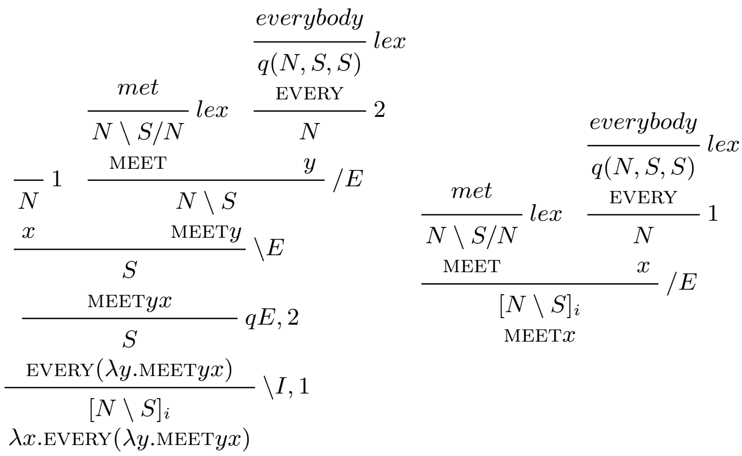

Fig. 14. Source VPs in (13a)

Again this phenomenon is not restricted to quantification. Whenever the derivation of the source VP involves hypothetical reasoning and it is possible to discharge the hypothesis after ellipsis resolution, multiple binding should be possible. This is in fact the case. For *wh*-movement, this was also observed in [21].

(14) a. (the man) that Mary met before Bill did
b. How many miles are you prepared to walk if the people want you to

The preferred reading of (14a) is *the man that Mary met before Bill met him.* Example (14b) is similar. Additionally it illustrates that this mechanism is not restricted to hypotheses of type N, and that constructions with “binding into ellipsis” need not correspond to a parallel construction without ellipsis and with a pronoun.

Last but not least, we encounter the same pattern in connection with focus. This was first noted [12], where the following example is given:

(15) I only went to TANGLEWOOD because you did

The sentence has a reading saying: The only place x such that I went to x because you went to x is Tanglewood.

Let me close this section with an example from [2] that demonstrates that the ambiguity of bound versus corefential interpretation of pronouns on the one hand and the strict/sloppy ambiguity on the other hand are independent phenomena:

(16) a. Every student revised his paper before the teacher did

b. Every student$_i$ revised his$_j$ paper before the teacher$_k$ revised his$_j$ paper
c. Every student$_i$ revised his$_i$ paper before the teacher$_j$ revised his$_j$ paper
d. Every student$_i$ revised his$_i$ paper before the teacher$_j$ revised his$_i$ paper

Sentence (16a) has three reading (paraphrased in (16b-d)). Next to the unproblematic cases where the pronoun is either free and strict (b) or bound and sloppy (c), there is an interpretation where the pronoun is bound but nevertheless strict (d). Gawron and Peters therefore assume a three-way ambiguity of pronoun uses—referential as in (b), role-linking as in (c), and co-parametric as in (d) (cf. [2]).

In the present systems, all three readings fall out immediately, even though the pronoun is unambiguous. If the pronoun is free, the derivation is analogous to Fig. 7. Readings (16c,d) are derived by first plugging in a hypothetical N into to the matrix subject position, giving the ellipsis a sloppy or strict construal (as in Fig. 13), and applying qE and thus replacing the hypothetical N by the quantifier.

5 Comparison to Jacobson's System

Despite the overall similarity between Jacobson's and the present treatment of anaphora, there are two important differences. As can be seen from rule $|R$, according to our proposal a pronoun is licensed primarily by a preceding antecedent. This antecedent may be hypothetical. In this case, the pronoun may be linked to an argument place of a super-ordinate functor. This option is employed in the derivation of bound and sloppy readings. Jacobson takes this method of binding to be basic. Due to the lack of unrestricted associativity in CCG, the restriction to super-ordinate functors is non-trivial here. This aspect of her system leads to two shortcomings that can be avoided in the type logical setting.[1]

First, not all bound pronouns can be treated in this fashion. In (17) (from [2]), there is no constituent that contains the pronoun and takes its antecedent as an argument (even under the flexible notion of constituency that CCG adopts).

(17) The soldiers turned some citizens in [each state]$_i$ over to its$_i$ governor

In the type logical formulation presented here, the discontinuity of the main verb may require an involved treatment, but since the antecedent precedes the pronoun, there is no problem with anaphora resolution.

As for the second point, reconsider the strict reading of (12). The only way to give the pronoun a bound reading under the combinatory approach is to bind it to the subject argument place of the verb *revised*. But this means that the property *to revise John's paper* doesn't occur as the meaning of any constituent. We only get the meaning *to revise one's paper* as the meaning of the source VP. So if we assume an identity-of-meaning approach to ellipsis, we only get the sloppy reading here if the pronoun is construed as bound. The best we can do to derive

[1] The same objections apply to the proposals in [5] and [22] as well.

the strict reading is to consider the pronoun as free and accidentally co-referring with *John*. But as (16) demonstrates, strict readings of bound pronouns are possible. The combinatory treatment of anaphora cannot handle constructions like this one.

6 Cross-sentential Anaphora

To extend the type logical approach to grammar to the discourse level, we have to introduce a new type, call it D, for discourses. Besides, type assignment has to guarantee at least that every sentence is a discourse, and that appending a sentence to a discourse yields a discourse again. So the following two sequents should be theorems of $\mathbf{L}|$:

(18) a. $S \Rightarrow D$
b. $D, S \Rightarrow D$

Clearly these cannot be derivable sequents if S and D are atomic. So we have to replace them by suitable complex types. Two options suggest themselves: both S and D should be identified either with $I \setminus I$ or I/I for some type I.[2] They both have a dynamic flavor: $I \setminus I$ is akin to the view of Discourse Representation Theory, File Change Semantics or Update Semantics (cf. [4, 11, 24]), where a sentence defines a function from information states to information states. I/I resembles DMG (Dynamic Montague Grammar, cf. [3]), where a sentence meaning is a function from possible continuations to "static" sentence meanings. I'll adopt the latter, DMG-style option here. The semantic type corresponding to I is t (or $\langle s, t\rangle$ if we incorporate intensionality). A sentence like *John walks* will receive the DMG-style meaning $\lambda p(\textsc{walk j} \wedge p)$. Now consider a sample discourse like

(19) John walked. He talked

If S is uniformly replaced by I/I in the type assignments, the relevant sequent becomes

(20) $N, N \setminus (I/I), N|N, N \setminus (I/I) \Rightarrow I/I$

Following the DMG philosophy again, we assume that verbs like *walk* and *talk* denote dynamic properties, i.e $\lambda xp.((\{\textsc{w/t}\}\textsc{alk } x) \wedge p)$. Fig. 15 shows that the sequent is valid and that the derived meaning is

(21) $\lambda p.((\textsc{walk j}) \wedge (\textsc{talk j}) \wedge p)$

The static truth conditions are obtained by applying this to the tautology.

This treatment can easily be extended to indefinites. Let us assume that an indefinite like *someone* has category $q(N, I, I)$ and meaning $\lambda P.\exists x P x$.

[2] To be precise, S should be identified with $\Box^{\downarrow}(I/I)$ or $\Box^{\downarrow}(I \setminus I)$ for some domain modality $\Box^{\downarrow}$ in the sense of [17] to take the special status of sentences into account. I'll ignore this issue here.

$$
\dfrac{
\dfrac{
\dfrac{\mathit{John}}{\substack{[N]_i\\ \mathrm{J}}}\,lex
\qquad
\dfrac{\mathit{walked}}{\substack{N\backslash I/I\\ \lambda xp.\mathrm{WALK}x\wedge p}}\,lex
}{\substack{I/I\\ \lambda p.\mathrm{WALK\ J}\wedge p}}\backslash E
\qquad
\dfrac{
\dfrac{
\dfrac{\dfrac{\mathit{He}}{\substack{[N|N]_i\\ \lambda x.x}}\,lex}{\substack{N\\ \mathrm{J}}}\,|E
\qquad
\dfrac{\mathit{talked}}{\substack{N\backslash I/I\\ \lambda xp.\mathrm{TALK}x\wedge p}}\,lex
}{\substack{I/I\\ \lambda P.\mathrm{TALK\ J}\wedge p}}\backslash E
\qquad
\dfrac{}{\substack{I\\ p}}\,1
}{\substack{I\\ \mathrm{TALK\ J}\wedge p}}/E
}{\substack{I\\ \mathrm{WALK\ J}\wedge\mathrm{TALK\ J}\wedge p}}/E
$$

$$
\dfrac{\substack{I\\ \mathrm{WALK\ J}\wedge\mathrm{TALK\ J}\wedge p}}{\substack{I/I\\ \lambda p.\mathrm{WALK\ J}\wedge\mathrm{TALK\ J}\wedge p}}/I,1
$$

Fig. 15. John walked. He talked

(22) Someone walked. He talked.

As Fig. 16 shows, a discourse like (22) receives type I/I and the meaning $\lambda p.\exists x((\mathrm{WALK}x)\wedge(\mathrm{TALK}x)\wedge p)$. The mechanism of cross-sentential/dynamic binding is thus essentially the same as for sentence internal binding. Other quantifiers like *every man* can be prevented from taking discourse scope by the same multimodal mechanisms that block them from outscoping *and* in coordinate structures (see for instance [18]).

7 Incremental Interpretation

This treatment of cross-sentential anaphora is compatible with the intuitive requirement that discourse interpretation works incrementally. Note that in all compositional theories of dynamic semantics, the meaning of a sentence includes information about how many old discourse markers are picked up, and how many novel discourse markers are introduced. Under the type logical perspective, this information has to be encoded in the type of a sentence. So we admit a limited polymorphism for the category of sentences. To take two simple examples, *A man walks* will have (among others) the type $I/(I|N)$ since it licenses a subsequent pronoun. *He walks* will have type $(I/I)|N$, indicating that it contains an old discourse referent that needs an antecedent. Generally, we assume that a sentence containing n (locally unbound) pronouns and introducing m discourse entities will have category $(I/(I(|N)^m))(|N)^n$. Semantically, such a sentence will denote a function from n individuals and an m-place relation to a proposition. Sentence concatenation corresponds to a family of generalized versions of function composition, where each argument place of the second sentence may either be filled by one of the discourse markers introduced by the first sentence, or

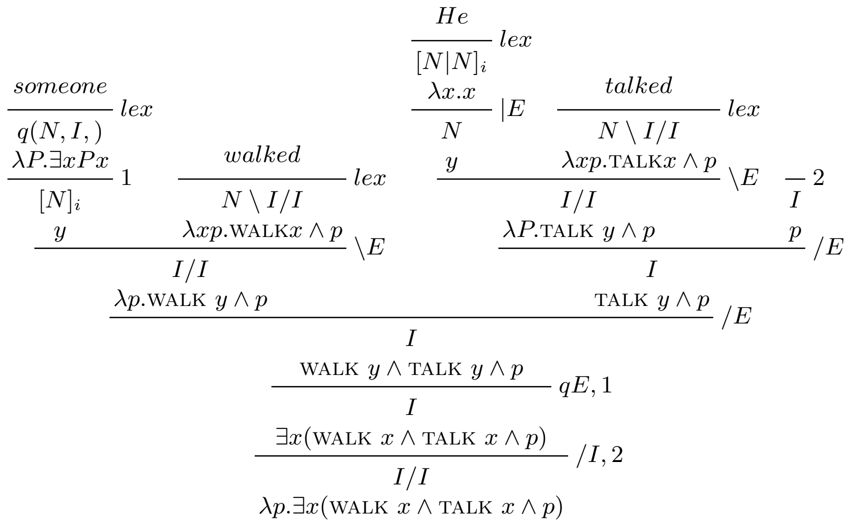

Fig. 16. Someone walked. He talked

projected to the discourse as a whole. [23] briefly considers a proposal similar to this, but rejects it due to the apparent proliferations of sentence combining operations. This is no obstacle here though, since all these operations are derivable in **L**|.

8 Future Work

Two directions for further research suggest themselves. In its present shape, the system cannot cope with cataphora. This phenomenon is very limited and many cases should arguably be treated as accidental coreference instead as a grammatical dependency (cf. [25] for an insightful discussion), but there are undeniable cases of grammatically determined backward binding.

Furthermore, it is yet unclear how other types of ellipsis should be incorporated. VP ellipsis is simple insofar as it leaves a proform, the stranded auxiliary. Insofar as ellipsis resolution is rooted in the lexicon here. This is not the case with gapping, stripping etc. So apparently an adequate extension of **L**| has to include devices to create anaphoric types in syntax. In other words, **L**| is still to weak for a treatment of ellipsis in general. Any strengthening of the logic is in risk of losing the finite reading property though.

9 Conclusion

This paper proposed the logic **L**| as a type logical reconstruction of Pauline Jacobson's treatment of anaphora in CG. It was shown that **L**| is weak enough to have the finite reading property, but strong enough to handle the multiplication of resources that we find in anaphoric dependencies. Paired with Moortgat's type logical approach to quantification, we are able to cope with a substantial amount of phenomena concerning pronominal anaphora, VP ellipsis and quantification. Finally it was sketched how cross sentential anaphora can be handled under this approach.

References

1. Noam Chomsky. Conditions on rules in grammar. *Linguistic Analysis*, 2:303–351, 1976.
2. Jean Mark Gawron and Stanley Peters. *Anaphora and Quantification in Situation Semantics*. CSLI, Stanford, 1990.
3. Jeroen Groenendijk and Martin Stokhof. Dynamic Montague Grammar. In Jeroen Groenendijk, Martin Stokhof, and David Ian Beaver, editors, *Quantification and Anaphora I*, DYANA deliverable R2.2a. Amsterdam, 1991.
4. Irene Heim. *The Semantics of Definite and Indefinite Noun Phrases*. PhD thesis, University of Massachusetts, Amherst, 1982.
5. Mark Hepple. *The Grammar and Processing of Order and Dependency: A Categorial Approach*. PhD thesis, University of Edinburgh, 1990.
6. Pauline Jacobson. Bach-Peters sentences in a variable-free semantics. In Paul Dekker and Martin Stokhof, editors, *Proceedings of the Eighth Amsterdam Colloquium*. University of Amsterdam, 1992.
7. Pauline Jacobson. Binding connectivity in copular sentences. In M. Harvey and L. Santelmann, editors, *Proceedings of SALT IV*, pages 161–178. Cornell University, 1994.
8. Pauline Jacobson. i-within-i effects in a variable-free semantics and a categorial syntax. In Paul Dekker and Martin Stokhof, editors, *Proceedings of the Ninth Amsterdam Colloquium*. University of Amsterdam, 1994.
9. Pauline Jacobson. Towards a variable-free semantics. *Linguistics and Philosophy*, 22(2):117–184, 1999.
10. Gerhard Jäger. Focus without variables: A multi-modal analysis. In Kimary N. Shahin, Susan Blake, and Eun-Sook Kim, editors, *Proceedings of WCCFL 17*. CSLI Publications, 1999.
11. Hans Kamp and Uwe Reyle. *From Discourse to Logic. Introduction to Modeltheoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory*. Kluwer, Dordrecht, 1993.
12. Angelika Kratzer. The representation of focus. In Arnim von Stechow and Dieter Wunderlich, editors, *Handbook Semantics*. de Gruyter, Berlin, New York, 1991.
13. Joachim Lambek. The mathematics of sentence structure. *American Mathematical Monthly*, 65:154–170, 1958.
14. Michael Moortgat. *Categorial Investigations. Logical and Linguistic Aspects of the Lambek Calculus*. Foris, Dordrecht, 1988.
15. Michael Moortgat. Generalized quantification and discontinuous type constructors. In W. Sijtsma and A. von Horck, editors, *Discontinuous Constituency*. De Gruyter, Berlin, 1996.
16. Michael Moortgat. In situ binding: A modal analysis. In Paul Dekker and Martin Stokhof, editors, *Proceedings or the 10th Amsterdam Colloquium*, pages 539–549. ILLC, University of Amsterdam, 1996.

17. Michael Moortgat. Multimodal linguistic inference. *Journal of Logic, Language and Information*, 5(3/4):349–385, 1996.
18. Glynn Morrill. *Type Logical Grammar.* Kluwer, 1994.
19. Fernando C. N. Pereira. Categorial semantics and scoping. *Computational Linguistics*, 16(1):1–10, 1990.
20. Mats Rooth. *Association with Focus.* PhD thesis, University of Massachusetts, Amherst, 1985.
21. Ivan A. Sag. *Deletion and Logical Form.* PhD thesis, MIT, 1976.
22. Anna Szabolcsi. *Bound Variables in Syntax. (Are there any?).* Institute of Linguistics, Hungarian Academy of Science, Budapest, 1988.
23. Anna Szabolcsi. Reconstruction, anaphora, and pronouns as identity maps. In Paul Dekker, Martin Stokhof, and Yde Venema, editors, *Proceedings of the 11th Amsterdam Colloquium.* University of Amsterdam, 1997.
24. Frank Veltman. Defaults in update semantics. *Journal of Philosophical Logic*, 25:221–261, 1996.
25. Edwin Williams. Blocking and anaphora. *Linguistic Inquiry*, 28(4):577–628, 1997.