University of Pennsylvania Working Papers in Linguistics

Volume 5
Issue 2 *Current Work in Linguistics*

Article 2

1-1-1998

A Multi-Modal Analysis of Anaphora and Ellipsis

Gerhard Jaeger

A Multi-Modal Analysis of Anaphora and Ellipsis

A Multi-Modal Analysis of Anaphora and Ellipsis*

Gerhard Jäger

1. Introduction

The aim of the present paper is to outline a unified account of anaphora and ellipsis phenomena within the framework of Type Logical Categorial Grammar.[1] There is at least one conceptual and one empirical reason to pursue such a goal. Firstly, both phenomena are characterized by the fact that they re-use semantic resources that are also used elsewhere. This issue is discussed in detail in section 2. Secondly, they show a striking similarity in displaying the characteristic ambiguity between strict and sloppy readings. This supports the assumption that in fact the same mechanisms are at work in both cases.

(1) a. John washed his car, and Bill did, too.
  b. John washed his car, and Bill waxed it.

In (1a), the second conjunct can mean that Bill washed Bill's car or that he washed John's car. Similarly, (1b) is ambiguous between a reading where Bill waxed John's and one where he waxed his own car. In the latter reading, *it* is usually called a *paycheck pronoun* or a *lazy pronoun*.

There is also a fundamental difference between ellipsis and anaphora, however. While ellipses require a strong syntactic and semantic *parallelism* between their own linguistic environment and the environment of their antecedents, nothing comparable can be observed in the case of (nominal) anaphors. This is immediately obvious in the case of strict readings, but sloppy, i.e. lazy readings show a considerable amount of tolerance here, too.

(2) John already spent his paycheck, but in Bill's case, it hasn't been handed out yet.

Arguably, a sort of semantic/pragmatic parallelism can be observed here, but certainly there is no syntactic parallelism, *his paycheck* being an object and *it* a subject.

*I am indebted to Natasha Kurtonina and Alexander Williams for valuable discussions and comments. Besides I profited from the suggestions of two anonymous referees. The research that led to this paper was funded partially by the Deutsche Forschungsgemeinschaft and partially by the National Science Foundation. I'd like to express my gratitude to both institutions.

[1] As introductions to this theory of grammar, the interested reader is referred to Carpenter 1997, Moortgat 1997, Morrill 1994.

This observation is somewhat surprising in view of the fact that it is frequently assumed that ellipsis interpretation is based on the recognition of syntactic parallelism between source clause and target clause (as for instance in Dalrymple et al. 1991 or in Hobbs and Kehler 1997). If this is true, anaphora and ellipsis are fundamentally different phenomena. To put it the other way around, a unified account of anaphora and ellipsis cannot make reference to parallelism. Hence the parallelism constraints that undeniably show up cannot originate in the ellipsis interpretation module itself but have to be located elsewhere in grammar. This is not too bad after all, since contexts that license VP ellipsis—like coordinations, comparative construction, question-answer sequences etc.—display parallelism effects even when there is no ellipsis. Such a line of argumentation has the advantage that the ellipsis module doesn't have to account for contrasts like the following:

(3) Who washed his car?

a. John did, although Bill already had.

b. John did, and Bill did, too.

While (3b) only has a uniformly sloppy reading, the preferred reading of (3a) is the one where both John and Bill washed John's car. The availability of the latter reading rests on the fact that John's washing his car is unlikely in case Bill already washed John's car, but not in case of Bill having washed his own car. Since not contrast but similarity is required by the conjunction *and ... too* in (3b), the corresponding reading is blocked. It strikes me as undesirable to give the syntax-semantics interface (which is arguably the locus of ellipsis interpretation) access to this kind of common sense knowledge. Thus the ellipsis interpretation module should give access to both readings in both cases.

2. Semantic Resources and Compositionality

If one assumes (a) a version of the Principle of Compositionality and (b) that meanings of natural language expressions can adequately be represented by means of expressions of the typed λ-calculus, one immediately arrives at the following claim (which is hardly more than a truism):

For each sign S consisting of n lexemes, in each of its readings there is an expression M of the typed λ-calculus with $x_1, \ldots, x_n$ occuring each exactly once such that

$$M[x_1 \leftarrow N_1, \ldots, x_n \leftarrow N_n] = S'$$

where S' represents the meaning of S and N_i the meaning of the i'th lexeme.

The term M can be said to represent the semantic structure of the sign. It is an obvious question to ask whether there are restrictions on the form of these structures in natural language semantics. It is uncontroversial to assume that every λ-operator should bind at least one variable occurrence. This disallows such unnatural meaning recipes like $((\lambda y.x_1)x_2)$, which would predict that the meaning of a sign can be completely independent of one of its lexical components.

A less obvious restriction that is frequently considered requires that each λ-operator in M binds at most one variable occurrence. This corresponds to the appealing intuition that each lexical resource is used exactly once. There are *prima facie* counterexamples to this view, but most of them can nevertheless be handled, as will be illustrated below. To do so, it is crucial to assume that the single-bind condition does not apply to lexical meanings. In the examples that will be discussed, (b) gives the meanings of the lexical items involved, (c) the desired sentence meaning after normalization, and (d) gives the term M in the sense of the definition above.

Reflexives

(4) a. John shaves himself.
  b. $N_1 = j, N_2 =$ *shave'*$, N_3 =?$
  c. $S' =$ *shave'*$j\ j$
  d. $M = (\lambda y.x_2yy)x_1$

At a first glance, the meaning of the subject is used twice here, while the meaning of the reflexive—whatever it may be—doesn't make any contribution at all. This puzzling situation can be overcome by assigning the meaning $\lambda T\lambda y.Tyy$ to the reflexive. Now the structure of the example gives rise to the meaning recipe $M = x_3x_2x_1$, which is perfect.[2]

Coordination Ellipsis

(5) a. John walks and talks.
  b. $N_1 = j, N_2 =$ *walk'*$, N_3 =$ *and'*$, N_4 =$ *talk'*
  c. $S' =$ *and'*$($*walk'* $j)($*talk'* $j)$
  d. $M = (\lambda y.x_3(x_2y)(x_4y))x_1$

Here again the meaning of the subject occurs twice. We can handle this by giving *and* the meaning

$$\lambda x\lambda y\lambda z.\mathit{and'}(xz)(yz)$$

[2]This analysis of reflexivation was proposed at various places, witness Keenan and Faltz 1985, Szabolcsi 1989.

This is basically already proposed in Montague's PTQ system and generalized to other types in Partee and Rooth 1983.

Other kinds of anaphors and ellipses

(6) a. John claims that he will win.
    b. *claim'*$(win'\ j)j$

Here the representation of the matrix subject occurs twice while the embedded subject completely disappears. Things are similar in the case of VP ellipsis:

(7) a. John walks, and Bill does, too.
    b. *and'*$(walk'b)(walk'j)$

Apparently the whole VP of the first conjunct gets recycled here. There are several ways to deal with these constructions. The burden of multiplying meaning could be transferred to the lexical semantics of the pronoun *he* in (6), and similarly to the auxiliary *does* in (7). In the case of bound anaphors, this has been proposed by Szabolcsi 1989 and Dalrymple et al. 1997. However, these systems only capture pronouns that are syntactically bound. Since ellipsis phenomena are largely identical within one sentence and across sentence boundaries, syntactic binding is unlikely to extend to ellipsis.

A currently quite popular approach assumes that the output of meaning composition is an underspecified representation where each lexical resource is used exactly once. The final meaning is achieved by resolving the underspecification, thereby possibly identifying several subexpressions. A paradigmatic example of this idea is Dalrymple et al. 1991, where the compositional meaning of (7) is supposed to be $and(P\ b)(walk'j)$, with P representing the meaning of *does (,too)*. This parameter is, in a final step, nailed down to the meanings it is supposed to have by means of a system of term equations.

Although such an approach has many attractive features, it strikes me as desirable to incorporate the semantics of anaphora and ellipsis into the compositional machinery. The only way of doing so seems to lie in a relaxation of the prohibition against multiple binding in syntax. To estimate the consequences and intricacies of such a move, we have to have a closer look on the relation between meaning recipes and syntactic structure.

3. The Syntax-Semantics Interface in Categorial Grammar

Compositionality of Interpretation requires that each syntactic operation is accompanied by a corresponding operation on meanings. Categorial Grammar strengthens this idea by assuming that not only syntactic and semantic objects, but also syntactic and semantic operations each form an algebra, and that there is also a homomorphism from syntactic to semantic operations. In the type logical version of Categorial Grammar, the syntactic operations are taken to be theorems (valid sequents) of a logical calculus generated from a single axiom scheme by application of a small set of inference rules. Correspondingly, semantic operations are generated from a single combinatorial scheme by closure under certain operations.

Syntactic categories, i.e. formulae of the syntax logic in question, are recursively built from a finite set of atomic categories $AtForm$ by means of the connectives "/" (rightward looking implication), "\" (leftward looking implication) and "•" (product). A *sequent* is a derivation $\Gamma \Rightarrow A$, where Γ is a binary tree of formulae (written as a bracketed string), and A is a formula. To transform such a logic into a full-blown grammar, two further ingredients have to be added, namely a set of designated categories (usually simply $\{S\}$), and an assignment of at least one category to each lexical item. A sequence of lexical items is recognized as a sentence by this grammar iff a sequent of corresponding categories can be bracketed in such a way that a designated category is derivable. The simplest logic fitting into this framework is the non-associative Lambek Calculus **NL** (Lambek 1961) which only has the axiom of identity and inference rules introducing a logical connective either at the left-hand or at the right-hand side of a sequent. Lambek 1961 proved that we can add the Cut rule without increasing the set of derivable sequents.

On the semantic side, there is a set of types which is the closure of a finite set of atomic types under the operations "$\rightarrow$" (function space) and "$\circ$" (Cartesian product). The homomorphism leading from categories to types is a straightforward generalization from the one in Montague's PTQ system (Montague 1974), requiring that "\" and "/" are sent to "$\rightarrow$" and"•" to "$\circ$". The only basic semantic operations are the identity maps on the domain of each type. The operations on semantic operations are most transparently defined as manipulations of polynomials in the simply typed λ-calculus (with product and projections). There is a one-one correspondence between inference rules and semantic meta-operations (which is of course just an instance of the Curry-Howard correspondence). Hence syntax and semantics can be presented simultaneously by augmenting the premises of the sequents in the Gentzen-style presentation with variables and the conclusions with polynomials over these variables. The axioms and rules of **NL** are presented below, where $\Gamma[A]$ stands for a binary tree with A as one of its leafs, and $\Gamma[B]$ for the result of replacing A with B in Γ.

(8)
$$\frac{}{x : A \Rightarrow x : A}[\text{id}]$$

$$\frac{\Delta \Rightarrow t_1 : A \quad \Gamma[x : A] \Rightarrow t_2 : C}{\Gamma[\Delta] \Rightarrow t_{2\,[x \leftarrow t_1]} : C}[\text{Cut}]$$

$$\frac{\Delta \Rightarrow t_1 : B \qquad \Gamma[x : A] \Rightarrow t_2 : C}{\Gamma[(y : A/B, \Delta)] \Rightarrow t_{2\,[x \leftarrow (y t_1)]} : C}[/L]$$

$$\frac{(\Gamma, x : B) \Rightarrow t : A}{\Gamma \Rightarrow \lambda x.t : A/B}[/R]$$

$$\frac{\Delta \Rightarrow t_1 : B \qquad \Gamma[x : A] \Rightarrow t_2 : C}{\Gamma[(\Delta, y : B\backslash A)] \Rightarrow t_{2\,[x \leftarrow (y t_1)]} : C}[\backslash L]$$

$$\frac{(x : B, \Gamma) \Rightarrow t : A}{\Gamma \Rightarrow \lambda x.t : B\backslash A}[\backslash R]$$

$$\frac{\Gamma[(x : A, y : B)] \Rightarrow t : C}{\Gamma[z : A \bullet B] \Rightarrow t_{[x \leftarrow \pi_1(z), y \leftarrow \pi_2(z)]} : C}[\bullet L]$$

$$\frac{\Gamma \Rightarrow t_1 : A \qquad \Delta \Rightarrow t_2 : B}{(\Gamma, \Delta) \Rightarrow \langle t_1, t_2 \rangle : A \bullet B}[\bullet R]$$

Confining ourselves to product-free types, it is easy to see by an induction on the complexity of proofs that polynomials derived as meaning recipes are terms of a limited fragment of the typed λ-calculus obeying the following constraints (cf. van Benthem 1987)

1. Each sub-term contains a free variable,
2. no sub-term contains more than one occurrence of the same variable, and
3. each λ binds a free variable.

A more flexible system is achieved by allowing arbitrary rebracketing of the antecedent of a sequent. This is captured by the structural rule of associativity, which turns **NL** into the associative Lambek Calculus **L**:[3]

(9)
$$\frac{\Gamma[(\Delta, (\Pi, \Sigma))] \Rightarrow t : A}{\Gamma[((\Delta, \Pi), \Sigma)] \Rightarrow t : A}[A]$$

However, the meaning recipes derived by **L** still confirm the mentioned constraints.

[3]The double line indicates that the rule can be applied in both directions.

4. Contraction and Permutation

Consider a simple elliptic sentence like

(10) John walks, and Bill, too

Since the parallel elements in the first and the second conjunct are not adjacent to *and*, assigning a polymorphic type like $X \setminus X/X$ to the conjunction would not enable us to derive the sentence in **L**. Besides, a viable solution should be able to cope with examples like (11) too, where no particular lexical item can be made responsible for the phenomenon:

(11) John walks. Bill too.

Therefore it seems desirable to assign (10) a meaning recipe like

(12) $x_3(x_2x_1)(x_2x_4)$

which uses one variable twice. To derive (12) as semantic structure of (10), **L** has to be extended by the structural rules of *contraction* and *permutation* to **LPC**:

(13) $$\frac{\Gamma[(x:A, y:A)] \Rightarrow t:B}{\Gamma[x:A] \Rightarrow t_{[y\leftarrow x]}:B}[C] \qquad \frac{\Gamma[(\Delta,\Pi)] \Rightarrow t:A}{\Gamma[(\Pi,\Delta)] \Rightarrow t:A}[P]$$

The essential steps of the derivation are (omitting brackets in the premises since these are redundant in **L**):

$$\cfrac{\cfrac{\cfrac{x:\mathbf{n}, y:\mathbf{n}\setminus\mathbf{s}, z:(\mathbf{s}\setminus\mathbf{s})/\mathbf{s}, u:\mathbf{n}, v:\mathbf{n}\setminus\mathbf{s} \Rightarrow z(vu)(yx):\mathbf{s}}{x:\mathbf{n}, y:\mathbf{n}\setminus\mathbf{s}, z:(\mathbf{s}\setminus\mathbf{s})/\mathbf{s}, v:\mathbf{n}\setminus\mathbf{s}, u:\mathbf{n} \Rightarrow z(vu)(yx):\mathbf{s}}[P]}{x:\mathbf{n}, y:\mathbf{n}\setminus\mathbf{s}, v:\mathbf{n}\setminus\mathbf{s}, z:(\mathbf{s}\setminus\mathbf{s})/\mathbf{s}, u:\mathbf{n} \Rightarrow z(vu)(yx):\mathbf{s}}[P]}{x:\mathbf{n}, y:\mathbf{n}\setminus\mathbf{s}, z:(\mathbf{s}\setminus\mathbf{s})/\mathbf{s}, u:\mathbf{n} \Rightarrow z(yu)(yx):\mathbf{s}}[C]$$

However, the unrestricted usage of contraction would lead to a heavy overgeneration. For instance, *John shows Mary* would be predicted to be a grammatical sentence with an interpretation like *John shows Mary herself*. More generally, in van Benthem 1991 it is shown that **LPC** based grammars only recognize regular languages, which makes them useless for linguistic purposes. Therefore we have to impose constraints on the applicability of these rules to avoid such a collapse.[4]

[4]Dalrymple et al. 1997 briefly consider the option to use the exponentials of Linear Logic for this purpose. While this is viable in their framework where the resource logic is used only for meaning assembly, it wouldn't be restrictive enough in a categorial setting.

5. A Multi-Modal System

Research in recent years has shown that none of the pure categorial logics (like **NL**, **L**, **LP** or **LPC**) is well-suited for a comprehensive description of natural language, each of them by itself being either too restrictive or too permissive. That's why combinations of several systems have attracted much attention. In the simplest case, such a multi-modal logic has more than one n-place product connective together with the corresponding residuation connectives. Each family is characterized by the usual logical rules and a set of characteristic structural rules. In more elaborate systems, these different modes of composition are allowed to communicate via certain *interaction postulates*. This allows for instance to distinguish between head adjunction and phrasal composition (cf. Moortgat and Oehrle 1996) or modeling discontinuity (Morrill 1995). This technique can be exploited to control the availability of contraction and permutation in the context of anaphora and ellipsis interpretation.

Pretheoretically, the proposal can be circumscribed as follows: Every node in a syntactic tree can be augmented by an arbitrary number of indices. Every index has itself a syntactic category and is marked with a polarity (+ or –). Anaphors carry a negative index of the appropriate category (N in the case of nominal anaphors, $N \setminus S$ in the case of VP anaphors) by means of a lexical specification. Every node in the tree can freely be augmented with a positive index of the same category. Positive indices can be moved to every node to their right. If a particular node simultaneously carries a positive index A and a negative index B such that $A \Rightarrow B$ is a theorem of our grammar logic, than both indices can be deleted.

To flesh this out formally, we propose to use a second mode of combination "$\sim$" (with corresponding residuation operations $\hookleftarrow$ and $\hookrightarrow$) besides concatenation. We augment **L** with the logical rules in (14) and the structural rules in (15), where $(\cdots)$ and $\{\cdots\}$ denote the bracketings corresponding to $\bullet$ and $\sim$ respectively. The new mode of combination is intended to capture the combination of a regular constituent with a (positive) index, ie. both $A \sim B$ and $\{A, B\}$ stand for B_{+A}. Leftward residuation corresponds to negative indexing, i.e. $A \hookrightarrow B \approx B_{-A}$, and [$\hookrightarrow$L] amounts to deletion of matching indices.

(14)
$$\frac{\Delta \Rightarrow t_1 : B \qquad \Gamma[x : A] \Rightarrow t_2 : C}{\Gamma[\{y : A \hookleftarrow B, \Delta\}] \Rightarrow t_{2\,[x \leftarrow (y t_1)]} : C}[\hookleftarrow \mathrm{L}]$$

$$\frac{\{\Gamma, x : B\} \Rightarrow t : A}{\Gamma \Rightarrow \lambda x.t : A \hookleftarrow B}[\hookleftarrow \mathrm{R}]$$

$$\frac{\Delta \Rightarrow t_1 : B \qquad \Gamma[x : A] \Rightarrow t_2 : C}{\Gamma[\{\Delta, y : B \hookrightarrow A\}] \Rightarrow t_{2\,[x \leftarrow (y t_1)]} : C}[\hookrightarrow L]$$

$$\frac{\{x : B, \Gamma\} \Rightarrow t : A}{\Gamma \Rightarrow \lambda x.t : B \hookrightarrow A}[\hookrightarrow R]$$

$$\frac{\Gamma[\{x : A, y : B\}] \Rightarrow t : C}{\Gamma[z : A \sim B] \Rightarrow t_{[x \leftarrow \pi_1(z), y \leftarrow \pi_2(z)]} : C}[\sim L]$$

$$\frac{\Gamma \Rightarrow t_1 : A \qquad \Delta \Rightarrow t_2 : B}{\{\Gamma, \Delta\} \Rightarrow \langle t_1, t_2 \rangle : A \sim B}[\sim R]$$

In (15) the structural rules for the hybrid system **LA** (**L**ambek Calculus with **A**naphora) are given. Allowing contraction for $\sim$ amounts to free indexing of any constituent (with a positive index). Index movement (IM) is formalized by an appropriate interaction postulate between $\bullet$ and $\sim$. The structural rule (P) ensures that the collection of indices attached to a node are unordered, i.e. they form a multiset.

(15) $$\frac{\Gamma[\{x : A, y : A\}] \Rightarrow t : B}{\Gamma[y : A] \Rightarrow t_{[x \leftarrow y]} : B}[C]$$

$$\frac{\Gamma[(\Delta, \{\Pi, \Theta\})] \Rightarrow t : A}{\Gamma[(\{\Pi, \Delta\}, \Theta)] \Rightarrow t : A}[IM]$$

$$\frac{\Gamma[\{\Pi, \{\Sigma, \Delta\}\}] \Rightarrow t : A}{\Gamma[\{\Sigma, \{\Pi, \Delta\}\}] \Rightarrow t : A}[P]$$

6. VP Ellipsis

To illustrate the system with a simple example, take the sentence

(16) John walks, and Bill does, too.

Informally, *does* gets a negative VP index from the lexicon. Semantically, it is interpreted as the identity function on VP meanings. In a first step, the VP *walks* gets a positive VP index with the same meaning as its host. In a second step, this index is moved to the second conjunct, where it is finally matched against the negative index and both are erased, resulting in application of the meaning of *does* to the meaning of the index. Formally, we assume the following lexical assignment:

(17) • John–j : $\mathbf{n}$

- Bill–b : $\mathbf{n}$
- walks–*walk'* : $\mathbf{n} \setminus \mathbf{s}$
- and–*and'* : $(\mathbf{s} \setminus \mathbf{s})/\mathbf{s}$
- does–$\lambda x.x$: $(\mathbf{n} \setminus \mathbf{s}) \hookrightarrow (\mathbf{n} \setminus \mathbf{s})$

In detail, the derivation looks as follows (the order is reversed since the derivation ends with the sequent to be proved, while we started with it in the informal description). We start with a sequent that is derivable in $\mathbf{L}$ (where parentheses are tacitly assumed to be left-associative):

$$x : \mathbf{n}, y : \mathbf{n} \setminus \mathbf{s}, z : (\mathbf{s} \setminus \mathbf{s})/\mathbf{s}, w : \mathbf{n}, r : (\mathbf{n} \setminus \mathbf{s}) \\ \Rightarrow z(rw)(yx) : \mathbf{s}$$

L$\hookrightarrow$ splits the second VP into a negatively indexed VP and a positive VP-index:

$$x : \mathbf{n}, y : \mathbf{n} \setminus \mathbf{s}, z : (\mathbf{s} \setminus \mathbf{s})/\mathbf{s}, w : \mathbf{n}, \{v : \mathbf{n} \setminus \mathbf{s}, u : (\mathbf{n} \setminus \mathbf{s}) \hookrightarrow (\mathbf{n} \setminus \mathbf{s})\} \\ \Rightarrow z(uvw)(yx) : \mathbf{s}$$

Associativity of concatenation allows rebracketing:

$$x : \mathbf{n}, y : \mathbf{n} \setminus \mathbf{s}, z : (\mathbf{s} \setminus \mathbf{s})/\mathbf{s}, (w : \mathbf{n}, \{v : \mathbf{n} \setminus \mathbf{s}, u : (\mathbf{n} \setminus \mathbf{s}) \hookrightarrow (\mathbf{n} \setminus \mathbf{s})\}) \\ \Rightarrow z(uvw)(yx) : \mathbf{s}$$

IM moves the positive VP index to the left

$$x : \mathbf{n}, y : \mathbf{n} \setminus \mathbf{s}, z : (\mathbf{s} \setminus \mathbf{s})/\mathbf{s}, (\{v : \mathbf{n} \setminus \mathbf{s}, w : \mathbf{n}\}, u : (\mathbf{n} \setminus \mathbf{s}) \hookrightarrow (\mathbf{n} \setminus \mathbf{s})) \\ \Rightarrow z(uvw)(yx) : \mathbf{s}$$

Another application of associativity gives us

$$x : \mathbf{n}, y : \mathbf{n} \setminus \mathbf{s}, z : (\mathbf{s} \setminus \mathbf{s})/\mathbf{s}, \{v : \mathbf{n} \setminus \mathbf{s}, w : \mathbf{n}\}, u : (\mathbf{n} \setminus \mathbf{s}) \hookrightarrow (\mathbf{n} \setminus \mathbf{s}) \\ \Rightarrow z(uvw)(yx) : \mathbf{s}$$

In sum, the last three steps amount to moving the index one item to the left. If we repeat this two more times, we get

$$x : \mathbf{n}, \{v : \mathbf{n} \setminus \mathbf{s}, y : \mathbf{n} \setminus \mathbf{s}\}, z : (\mathbf{s} \setminus \mathbf{s})/\mathbf{s}, w : \mathbf{n}, u : (\mathbf{n} \setminus \mathbf{s}) \hookrightarrow (\mathbf{n} \setminus \mathbf{s}) \\ \Rightarrow z(uvw)(yx) : \mathbf{s}$$

Now we have a VP carrying a positive VP index, and we can apply contraction, thereby identifying the meaning of host VP and index.

(18) $x : \mathbf{n}, y : \mathbf{n} \setminus \mathbf{s}, z : (\mathbf{s} \setminus \mathbf{s})/\mathbf{s}, w : \mathbf{n}, u : (\mathbf{n} \setminus \mathbf{s}) \hookrightarrow (\mathbf{n} \setminus \mathbf{s})$
$\Rightarrow z(uyw)(yx) : \mathbf{s}$

After inserting the lexical meanings, we obtain the reading

$$\textit{and'}((\lambda x.x)\textit{walk'}\ b)(\textit{walk'}\ j)$$

which reduces to

$$\textit{and'}(\textit{walk'}\ b)(\textit{walk'}\ j)$$

The mechanism works similar in the case of nominal anaphors. Since a possessive pronoun like *his* behaves like a definite determiner except that it requires a nominal antecedent, the type assignment $\mathbf{n} \hookrightarrow (\mathbf{n}/\mathbf{cn})$ seems appropriate. So we assume the following lexical entries:

(19)
- washed–$wash : (\mathbf{n} \setminus \mathbf{s})/\mathbf{n}$
- his–$\lambda xy.\textit{of'}\ y\ x : \mathbf{n} \hookrightarrow (\mathbf{n}/\mathbf{cn})$
- car–*car'* : $\mathbf{cn}$

We derive the reading (20c) for (20a), corresponding to the provable sequent (20b). The derivation is given in figure 1.

(20) a. John washed his car.
b. $x : \mathbf{n}, y : (\mathbf{n} \setminus \mathbf{s})/\mathbf{n}, z : \mathbf{n} \hookrightarrow (\mathbf{n}/\mathbf{cn}), w : \mathbf{cn} \ \Rightarrow\ y(zxw)x : \mathbf{s}$
c. $\textit{wash'}((\lambda xy.\textit{of'}\ y\ x)j\ \textit{car'})j\ (= \textit{wash'}(\textit{of'}\ \textit{car'}\ j)j)$

Before we proceed to the interaction of VP ellipsis and anaphora, observe that (20) shows a spurious ambiguity. After (20b) is derived, we can either stop or apply the rule "\L", which gives us the sequent

(21) $y : (\mathbf{n} \setminus \mathbf{s})/\mathbf{s}, z : \mathbf{n} \hookrightarrow (\mathbf{n}/\mathbf{cn}), w : \mathbf{cn} \ \Rightarrow\ \lambda x.y(zxw)x : \mathbf{n} \setminus \mathbf{s}$

This means that it is possible to resolve the anaphor *his* against the subject argument place of *washed*, assigning the meaning $\lambda x.\textit{wash'}(\textit{of'}\ \textit{car'}\ x)x$ to the VP *washed his car*. In terms of indices, this means that not only overt constituents but also open argument places license the introduction of positive indices.[5] In (20) this ambiguity is spurious since after combining this VP with the subject *John*, we end up with the meaning (20c) again.

(22) John washed his car, and Bill did, too.

[5]This can be seen as a reconstruction of Reinhart's 1983 distinction between coreferential and bound pronouns.

In (22), on the other hand, this ambiguity, though spurious in the first conjunct, makes a difference for the interpretation of the second one. If we plug in (21) into the conclusion of (16) via the Cut rule, we immediately derive the sloppy reading of (22). This amounts to first resolving *his* against the subject argument place of *washed* and afterward resolving *did* against the VP derived in this way. If, on the other hand, *his* is resolved against *John* prior to resolution of *did*, the strict reading results.

Dalrymple et al. 1991 present an example of a cascaded ellipsis that allows us to distinguish different ellipsis theories on a very fine-grained level.[6]

(23) John revised his paper before the teacher did, and Bill did, too.

Only those readings are considered where the first occurrence of *his* refers back to *John*, and the second *did* to the whole first conjunct. For simplicity, we treat *the teacher* as a proper noun with meaning t here. *Before* is analyzed on a par with *and* in the previous example. We use abbreviations like *JJBB* for the reading where John revised **John's** paper before the teacher revised **John's** paper and Bill revised **Bill's** paper before the teacher revised **Bill's** paper.

By combining the VP derivation in (21) with the VP ellipis structure in (18) by means of Cut and subsequently applying "R\" we obtain the semantic term

$$\lambda x.\mathit{before'}(\mathit{revise'}(\mathit{of'paper'}t)t)(\mathit{revise'}(\mathit{of'paper'}x)x)$$

for the VP *revised his paper before the teacher did*. Combining this again with (18) by means of Cut gives us the uniformly sloppy reading *JTBT*.

Starting proof search with with resolving *his* against John and resolving the two VP ellipses afterwards results in the uniformly strict reading *JJJJ*.

The derivation of *JJBB* is a bit more involved. First observe that the following sequent is valid, corresponding to the strict reading of the first conjunct:

$$x : \mathbf{n}, r : (\mathbf{n} \setminus \mathbf{s})/\mathbf{n}, h : \mathbf{n} \hookrightarrow (\mathbf{n}/\mathbf{cn})p : \mathbf{cn}, b : (\mathbf{s} \setminus \mathbf{s})/\mathbf{s}, t : \mathbf{n},$$

$$d : (\mathbf{n} \setminus \mathbf{s}) \hookrightarrow (\mathbf{n} \setminus \mathbf{s}) \Rightarrow b(d(r(h\ x\ p)t))(r(h\ x\ p)x) : \mathbf{s}$$

Applying "\R" to this sequent and inserting lexical entries gives us the reading

$$\lambda x.\mathit{before'}(\mathit{revise'}(\mathit{of'paper'}x)t)(\mathit{revise'}(\mathit{of'paper'}x)x) : \mathbf{n} \setminus \mathbf{s}$$

for *revised his paper before the teacher did*. This in turn serves as antecedent for the second *did*, resulting finally in the *JJBB* reading.

[6]The remainder of this section serves to compare the predictions of the present theory with the HOU approach and is not essential for the rest of the paper.

Recall that for index matching not identity is required but derivability of the negative from the positive index (this follows immediately from "$\hookrightarrow$L"). This is exploited in *JTJT*. In a first step, *revised his paper* is recognized as a VP with an unresolved anaphor, i.e. category $\mathbf{n} \hookrightarrow (\mathbf{n} \setminus \mathbf{s})$ with the reading

$$\lambda xy.\textit{revise'}(\textit{of'paper'}x)y$$

This is copied to a positive index, which is in turn moved to the first *did*. Since

$$v : \mathbf{n} \hookrightarrow (\mathbf{n} \setminus \mathbf{s}) \Rightarrow \lambda x.v\ x\ x : \mathbf{n} \setminus \mathbf{s}$$

is a theorem of **LA**, the negative VP index of *did* can be discharged by using the result of applying this operation to the positive index, resulting in a sloppy reading for *the teacher did*. The anaphor *his* is still free to be resolved against *John*, giving us the *JT* reading for the first conjunct. Resolving the second *did* finally results in *JTJT*.

Chosing a strict reading for the first ellipsis (i.e. a reading *JJ??*) severely restricts the possible interpretations of the second one. Resolution of the first *did* can only take place after *his* is resolved, be it against *John* or against a hypothetical variable to be abstracted away later. The former results in *JJJJ*, the latter in *JJBB*. No further options are available. Hence neither *JJBJ* nor *JJJB* can be derived in the present theory.

This result is not too bad since the four predicted readings unequivocally exist and *JJJB* is definitely impossible. Native speaker intuitions differ with respect to *JJBJ*. It is unavailable with the example (23), but Dalrymple et al. 1991 claim that it improves in the structurally equivalent example (24):

(24) Dewey announced his victory after the newspaper did, but so did Truman.

7. Associativity?

In view of the considerations in section 1, the theory seems to be too restrictive in its present form. In particular, it excludes the mixed sloppy/strict reading in (25) (from Dahl 1974):

(25) John realizes that he is a fool, but Bill does not, even though his wife does.

An obvious way to relax the constraints of the theory is to allow a lexical assignment like

(26) does–$\lambda x.x : (\mathbf{n} \hookrightarrow (\mathbf{n} \setminus \mathbf{s})) \hookrightarrow (\mathbf{n} \hookrightarrow (\mathbf{n} \setminus \mathbf{s}))$

for the first *does*. This would enable us to resolve it against *realizes that he is a fool* before *he* is resolved. In this way, the silent *he* can be resolved independently from the overt one, yielding (among others) the desired reading.

While it seems to be *ad hoc* to assume such a lexical ambiguity for *does*, this type assignment can be derived if we add a version of the Geach Rule to our calculus:

$$x : A \hookrightarrow B \;\Rightarrow\; \lambda yz.x(yz) : (C \hookrightarrow A) \hookrightarrow (C \hookrightarrow B)$$

Inserting the identity function (as the lexical meaning of *does*) for x gives us the semantic term $\lambda yz.yz$ for the derived category, which is equivalent to $\lambda y.y$. In terms of sequent rules, this amounts to extending **LA** to a new system, call it **LAA**, which includes the structural rule of Associativity for both modes of combination:

(27) $$\frac{\Gamma[\{\Delta, \{\Pi, \Sigma\}\}] \Rightarrow t : A}{\Gamma[\{\{\Delta, \Pi\}, \Sigma\}] \Rightarrow t : A}[\mathrm{A}\!\sim]$$

The decision between **LA** and **LAA** as appropriate calculus for anaphora and ellipsis is an empirical issue that has to be decided for each class of phenomena separately.

As far as English VP ellipsis is concerned, **LAA** predicts a very high degree of freedom. Besides the six readings for (23), it also admits readings like *JTTT* etc. Two comments are in order here. First, something similar to *JTTT* seems to be marginally possible indeed (judgments range from "impossible" to "perfect"):

(28) [Every bum on the streets of New York]$_j$ is more concerned about his$_j$ safety than this crowd loving president Clinton$_i$ is.

a. Fortunately for him$_i$, his$_i$ bodyguard is too.

b. Fortunately for him$_i$, his$_i$ bodyguard is more concerned about his$_i$ safety than he$_i$ is concerned about his$_i$ safety.

Second, restrictions on anaphora resolution in constructions without ellipsis do not substantially differ from those with ellipsis. (29) shows exactly the same range of readings like (23).

(29) John revised his paper before the teacher revised his paper, and Bill revised his paper before the teacher revised his paper, too.

If *too* is understood as establishing a parallelism between *John* and *Bill*, we have just the same four or five readings we have in (23). This fact is well-known (see for instance Gardent 1997). One way to account for this is the assumption that the deaccenting of the VPs in (29) that correspond to the elided material in (23) is the primary cause for this similarity. Ellipsis and deaccenting could be analyzed as largely two instances of the same phenomenon. Nevertheless another perspective is possible as well. The restrictions on anaphoric relationships that show up could be analyzed as consequences of the semantics/pragmatics of *too*, which simultaneously requires deaccenting of the second conjunct. This would make the differences between *and ... too*, *but*, *even though* etc. less mysterious. If such a line of research proves to be successful, this would allow a highly unrestrictive theory of ellipsis interpretation like the one implied by $\mathbf{LAA}$.

An $\mathbf{LAA}$ based account seems definitely be preferable in the case of nominal anaphors, since this automatically captures *paycheck* pronouns.

(30) a. Bill spent his money, and John saved it.

b.
- spent–*spend'* : $(\mathbf{n} \setminus \mathbf{s})/\mathbf{n}$
- saved–*save'* : $(\mathbf{n} \setminus \mathbf{s})/\mathbf{n}$
- money–*money'* : $\mathbf{cn}$

c. $and'(save'(of'money'j)j)(spend'(of'money'b)b)$

The crucial part of the derivation is given in figure 2. Most importantly, *it* can get the derived category $(\mathbf{n} \hookrightarrow \mathbf{n}) \hookrightarrow (\mathbf{n} \hookrightarrow \mathbf{n})$, again with the interpretation as identity functions (over Skolem functions). Hence *his money* with the pronoun still unresolved (which denotes the Skolem function from individuals to their cars) can serve as antecedent for *it*.

In the case of stripping, $\mathbf{LA}$ seems to be the appropriate logic, although judgments are somewhat fuzzy here. In (31a) all contextual factors favor a mixed sloppy/strict reading (as indicated in (31b)), which is nevertheless only very marginally possible.[7]

(31) a. Every candidate believes that he can win, even Smith, but not his wife.

b. Every candidate believes that he can win, even Smith believes that he can win, but his wife does not believe that Smith can win.

---

[7]Native speakers of German reject the corresponding example altogether.

8. Comparison with Jacobson's Theory of Anaphora

There is a striking correspondence between the present proposal and Jacobson's theory of anaphora (cf. Jacobson 1992b, Jacobson 1994, Jacobson 1996). Technically, the difference between this theory and the **LA**-based one is just the difference between the combinatory and type-logical variant of Categorial Grammar. Recall that in Combinatory Categorial Grammar, we have just the product free types of **L**. The axiom scheme of **L** is an axiom scheme in CCG as well. The main point of departure lies in the inference rules that can be used. Every version of CCG uses "/L", "\L" and "Cut", while "/R" and "\R" are not available. Besides, this deductive system can be extended by other axioms or axiom schemes, some of which are derivable in **L**, some aren't. Jacobson extends this basic system with a new type forming connective that corresponds to $\hookrightarrow$ in **LA**. Its behavior is governed by the following rules:

(32) a. $x : A/B \Rightarrow \lambda yz.x(yz) : (C \hookrightarrow A)/(C \hookrightarrow B)$
b. $x : A \setminus B \Rightarrow \lambda yz.x(yz) : (C \hookrightarrow A) \setminus (C \hookrightarrow B)$
c. $f : A \setminus (B \setminus C) \Rightarrow \lambda gx.f(gx)x : (B \hookrightarrow A) \setminus (B \setminus C)$
d. $f : (C/B)/A \Rightarrow \lambda gx.f(gx)x : (C/B)/(B \hookrightarrow A)$
e. $f : A \setminus (C/B) \Rightarrow \lambda gx.f(gx)x : (B \hookrightarrow A) \setminus (C/B)$
f. $f : (B \setminus C)/A \Rightarrow \lambda gx.f(gx)x : (B \setminus C)/(B \hookrightarrow A)$

With the exception of (32d) and (32e), these rules are theorems of **LA**. Extending **LA** with a mirror image of "IM" would even capture all of them. Since neither (32d) nor (32e) are used in Jacobson's analyses, this means that her results on Bach-Peters sentences, i-within-i effects, functional questions and right node raising carry over to the present approach without problems. This does not hold, however, for her approach to ACD (Jacobson 1992a) and to weak crossover phenomena.

On the other hand, our approach copes with VP ellipsis in a way that is not viable in Jacobson's system. In its published form, it does not recognize (1a) as a grammatical sentence if we use our lexical assignment. This can be fixed by minor amendments (for instance by assigning *and* the category $((\mathbf{n} \setminus \mathbf{s}) \setminus (\mathbf{n} \setminus \mathbf{s}))/\mathbf{s}$ and the meaning $\lambda pVx.\mathit{and'}p(Vx)$), but this would generate only the sloppy reading. At the present point it is still an open question which revisions would enable Jacobson's system to generate all readings of ellipses constructions. It is not unlikely that we in fact need the full power of conditionalization to achieve this goal (recall that the presence vs. absence of conditionalization is the crucial difference between type logical grammar and CCG). So future work has to show whether and how the advantages of the combinatory and the type-logical approach can be combined.

9. Conclusion and Further Research

In this paper, I have outlined a theory of anaphora and ellipsis which shows two desirable properties from a conceptual point of view:

- The semantics is fully compositional. As a consequence, there is no need for a level of Logical Form where ellipsis resolution takes place. Since ellipsis phenomena are usually considered to be a strong indication for the presence of LF, this might have consequences for grammar architecture as a whole. Neither does the theory presented here crucially depend on the typed λ-calculus as a semantic representation language. That it has been used throughout the paper is merely a matter of convenience; everything could be reformulated in terms of set theory or Combinatory Logic without loss of generality.

- The theory is variable free. This removes a great deal of arbitrariness from semantic derivations. In traditional theories, anaphors and ellipses are translated as variables (i.e. they denote functions from assignment functions to objects of the appropriate type). Since there are infinitely many variables, one and the same pronoun is predicted to be infinitely ambiguous. Though this is compatible with the letter of the Principle of Compositionality, it is clearly against its spirit, since identical expressions with identical syntactic structure should have identical denotations. Here, resolution ambiguities are treated as structural ambiguities, corresponding to essentially different proofs of the same sequent.

Let me conclude with a list of open questions that have to be addressed by further research.

The lexical type assignment hasn't been discussed yet. While it seems reasonable to treat nominal anaphors as identity functions over individuals with the syntactic category $\mathbf{n} \hookrightarrow \mathbf{n}$ by lexical stipulation, the similar assumption for English auxiliaries is less obvious, and the stripping cases cannot be handled in this way at all. To deal with examples as in (33), we have to assign the type $(\mathbf{n} \setminus \mathbf{s}) \hookrightarrow \mathbf{s}$ to *Bill*.

(33) John walks, and Bill too.

Doing this in the lexicon would be completely *ad hoc*. This assignment could be derived from the basic type $\mathbf{n}$ if the following rule were a theorem:

$$\dfrac{\dfrac{\dfrac{\mathbf{n} \Rightarrow \mathbf{n} \quad \mathbf{n}, \mathbf{n}\backslash\mathbf{s}/\mathbf{n}, \mathbf{n}/\mathbf{cn}, \mathbf{cn} \Rightarrow \mathbf{s}}{\mathbf{n}, \mathbf{n}\backslash\mathbf{s}/\mathbf{n}, \{\mathbf{n}, \mathbf{n} \hookrightarrow (\mathbf{n}/\mathbf{cn}), \mathbf{cn}\} \Rightarrow \mathbf{s}}{\scriptstyle[L\hookrightarrow]}}{\{\mathbf{n}, \mathbf{n}\}, \mathbf{n}\backslash\mathbf{s}/\mathbf{n}, \mathbf{n} \hookrightarrow (\mathbf{n}/\mathbf{cn}), \mathbf{cn} \Rightarrow \mathbf{s}}{\scriptstyle[IM+Ass]}}{\mathbf{n}, \mathbf{n}\backslash\mathbf{s}/\mathbf{n}, \mathbf{n} \hookrightarrow (\mathbf{n}/\mathbf{cn}), \mathbf{cn} \Rightarrow \mathbf{s}}{\scriptstyle[C]}$$

Figure 1: Partial Derivation of *John washed his car*

$$\dfrac{\dfrac{\dfrac{\dfrac{\{\{\mathbf{n}, \mathbf{n} \hookrightarrow \mathbf{n}\}, \mathbf{n} \hookrightarrow \mathbf{n}\} \Rightarrow \mathbf{n}}{\{\mathbf{n}, \{\mathbf{n} \hookrightarrow \mathbf{n}, \mathbf{n} \hookrightarrow \mathbf{n}\}\} \Rightarrow \mathbf{n}}{\scriptstyle[A\sim]}}{\{\mathbf{n} \hookrightarrow \mathbf{n}, \mathbf{n} \hookrightarrow \mathbf{n}\} \Rightarrow \mathbf{n} \hookrightarrow \mathbf{n}}{\scriptstyle[\hookrightarrow R]}}{\mathbf{n} \hookrightarrow \mathbf{n} \Rightarrow (\mathbf{n} \hookrightarrow \mathbf{n}) \hookrightarrow (\mathbf{n} \hookrightarrow \mathbf{n})}{\scriptstyle[\hookrightarrow R]} \quad \dfrac{\dfrac{\mathbf{n} \hookrightarrow \mathbf{n} \Rightarrow \mathbf{n} \hookrightarrow \mathbf{n} \quad \dfrac{\dfrac{\mathbf{n} \Rightarrow \mathbf{n} \quad \mathbf{n}, \mathbf{n}\backslash\mathbf{s}/\mathbf{n}, \mathbf{n} \Rightarrow \mathbf{s}}{\{\mathbf{n}, \mathbf{n}\}, \mathbf{n}\backslash\mathbf{s}/\mathbf{n}, \mathbf{n} \hookrightarrow \mathbf{n} \Rightarrow \mathbf{s}}{\scriptstyle[\hookrightarrow L+IM+IP]}}{\mathbf{n}, \mathbf{n}\backslash\mathbf{s}/\mathbf{n}, \mathbf{n} \hookrightarrow \mathbf{n} \Rightarrow \mathbf{s}}{\scriptstyle[C]}}{\{\mathbf{n} \hookrightarrow \mathbf{n}, (\mathbf{n}, \mathbf{n}\backslash\mathbf{s}/\mathbf{n}, (\mathbf{n} \hookrightarrow \mathbf{n}) \hookrightarrow (\mathbf{n} \hookrightarrow \mathbf{n}))\} \Rightarrow \mathbf{s}}{\scriptstyle[\hookrightarrow L+IM+IP]}}{\mathbf{n}, \mathbf{n}\backslash\mathbf{s}/\mathbf{n}, (\mathbf{n} \hookrightarrow \mathbf{n}) \hookrightarrow (\mathbf{n} \hookrightarrow \mathbf{n}) \Rightarrow (\mathbf{n} \hookrightarrow \mathbf{n}) \hookrightarrow \mathbf{s}}{\scriptstyle[\hookrightarrow R]}}{\mathbf{n}, \mathbf{n}\backslash\mathbf{s}/\mathbf{n}, \mathbf{n} \hookrightarrow \mathbf{n} \Rightarrow (\mathbf{n} \hookrightarrow \mathbf{n}) \hookrightarrow \mathbf{s}}{\scriptstyle[Cut]}$$

Figure 2: Partial Derivation of *Bill spent his money, and John saved it* in the sloppy reading.

(34) $B/A \Rightarrow A \hookrightarrow B$

However, adding this to **LA** as it is would lead to heavy over-generation, allowing for unrestricted deletion of substrings if preceded by an identical substring. It is inevitable to restrict (34) appropriately, thereby taking the interaction of ellipsis with intonation and focus into consideration.

Nothing has been said so far about the model theory of **LA**. It is no more than a technical exercise to identify a class of multi-modal ordered groupoids such that **LA** is sound and complete, but this would make the prosodic structures very abstract and make them resemble GB's S-structures more than surface structures. This is against the surface-oriented creed of Categorial Grammar. Therefore it is desirable to have a model theory where prosodic algebras are just sets of strings with concatenation as the only operation, and to assign the burden of the second mode of combination to the semantic algebra instead.

Finally, it should be checked to what degree the insights of Dynamic Semantics can be incorporated into the present approach. Such an attempt is promising both from a technical and an empirical point of view. Since both Dynamic Logic and substructural logics describe cognitive actions rather than states, natural connections are likely to exist. Nevertheless, the area is largely unexplored (but see van Benthem 1991,Oehrle 1997). Empirically, such a "dynamic turn" seems inevitable anyway, in order to handle discourse ellipsis as for instance in question-answer sequences.

References

Carpenter, Bob. 1997. *Type-Logical Semantics*. MIT Press.

Dahl, Östen. 1974. How to open a sentence. In *Logical Grammar Report*. University of Göteborg.

Dalrymple, Mary, John Lamping, Fernando C. N. Pereira, and Vijay Saraswat. 1997. Quantifiers, Anaphora and Intensionality. *Journal of Logic, Language and Information* 6(3):219–273.

Dalrymple, Mary, Stuart M. Shieber, and Fernando C.N. Pereira. 1991. Ellipsis and Higher-Order Unification. *Linguistics and Philosophy* 14(4):399–452.

Gardent, Claire. 1997. Parallelism, HOU and Deaccenting. Claus Report 85. Universität des Saarlandes, Saarbrücken.

Hobbs, Jerry R., and Andrew Kehler. 1997. A Theory of Parallelism and the Case of VP Ellipsis. In *Proceedings of the 35th Annual Meeting of the Association for Computational Linguistics*, 394–401. July.

Jacobson, Pauline. 1992a. Antecedent Contained Deletion in a Variable-Free Semantics. In *Proceedings of SALT 2*, ed. Chris Barker and David Dowty. 193–213. Working Papers in Linguistics, No. 40. Ohio State University, Columbus.

Jacobson, Pauline. 1992b. Bach-Peters sentences in a variable-free semantics. In *Proceedings of the Eighth Amsterdam Colloquium*, ed. Paul Dekker and Martin Stokhof. University of Amsterdam.

Jacobson, Pauline. 1994. i-within-i effects in a variable-free semantics and a categorial syntax. In *Proceedings of the Ninth Amsterdam Colloquium*, ed. Paul Dekker and Martin Stokhof. University of Amsterdam.

Jacobson, Pauline. 1996. The Syntax/Semantics Interface in Categorial Grammar. In *The Handbook of Contemporary Semantic Theory*, ed. Shalom Lappin. 89–116. Blackwell Publishers.

Keenan, Edward L., and Leonard M. Faltz. 1985. *Boolean Semantics for Natural Language*. Reidel, Dordrecht.

Lambek, Joachim. 1961. On the Calculus of Syntactic Types. In *Structure of Language and Its Mathematical Aspects*, ed. Roman Jakobson. Providence, RI.

Montague, Richard. 1974. *Formal Philosophy*. Yale University Press, New Haven.

Moortgat, Michael. 1997. Categorial Type Logics. In *Handbook of Logic and Language*, ed. Johan van Benthem and Alice ter Meulen. Chap. 2, 93–178. Elesvier, MIT Press.

Moortgat, Michael, and Richard T. Oehrle. 1996. Structural Abstraction. In *Proofs and Linguistic Categories: Applications of Logic to the Analysis and Implementation of Natural Language*. 49–64. Società Italiena di Logica e Filosofia della Scienza and Centro Interdipartmentale di Ricerca in Epistemologia e Storia delle Scienze "F. Enriques", Bologna. Proceedings of the 1996 Roma Workshop.

Morrill, Glynn. 1994. *Type Logical Grammar*. Kluwer.

Morrill, Glynn. 1995. Discontinuity in Categorial Grammar. *Linguistics and Philosophy* 18:175–219.

Oehrle, Richard T. 1997. Binding as Deduction. to appear in Geert-Jan M. Kruijff, Glynn V. Morrill and Richard T. Oehrle (eds.), Formal Grammar 1997.

Partee, Barbara, and Mats Rooth. 1983. Generalized Conjunction and Type Ambiguity. In *Meaning, Use, and Interpretation of Language*, ed. Rainer Bäuerle, Christoph Schwarze, and Arnim von Stechow. de Gruyter, Berlin, New York.

Reinhart, Tanya. 1983. *Anaphora and Semantic Interpretation*. Croom Helm.

Szabolcsi, Anna. 1989. Bound Variables in Syntax (are there any?). In *Semantics and Contextual Expressions*, ed. Renate Bartsch, Johan van Benthem, and Peter van Emde Boas. 295–318. Foris.

van Benthem, Johan. 1987. Categorial Grammar and lambda calculus. In *Mathematical Logic and Its Applications*, ed. Dimitur Genchev Skordev. 39–60. Plenum, New York.

van Benthem, Johan. 1991. *Language in Action*. Elesvier, Amsterdam.

Institute for Research in Cognitive Science
University of Pennsylvania
3401 Walnut Street, Suite 400A
Philadelphia, PA 19104
gjaeger@linc.cis.upenn.edu