



Centrum voor Wiskunde en Informatica

REPORTRAPPORT

Hypothetical-Reasoning and Radical Non-constituent Coordination
in Categorical Logic

N. Francez

Information Systems (INS)

INS-R9707 September 30, 1997

Report INS-R9707
ISSN 1386-3681

CWI
P.O. Box 94079
1090 GB Amsterdam
The Netherlands

CWI is the National Research Institute for Mathematics and Computer Science. CWI is part of the Stichting Mathematisch Centrum (SMC), the Dutch foundation for promotion of mathematics and computer science and their applications.

SMC is sponsored by the Netherlands Organization for Scientific Research (NWO). CWI is a member of ERCIM, the European Research Consortium for Informatics and Mathematics.

Copyright © Stichting Mathematisch Centrum
P.O. Box 94079, 1090 GB Amsterdam (NL)
Kruislaan 413, 1098 SJ Amsterdam (NL)
Telephone +31 20 592 9333
Telefax +31 20 592 4199

Hypothetical-Reasoning and Radical Non-constituent Coordination in Categorical Logic

Nissim Francez

CWI, P.O. Box 94079, Amsterdam, The Netherlands
and OTS, Utrecht university

on leave from the computer science dept., Technion-IL, Haifa, Israel

ABSTRACT

The paper investigates the connection between non-constituent coordination, as implemented in categorical grammar by means of a polymorphic type-assignment to lexical conjunctions, and hypothetical reasoning in Categorical Logics. A way of extending the logic is suggested, so that coordination can be applied to types depending on undischarged assumptions. By a certain “resource manipulation” of assumptions (of hypothetical reasoning), a late-discharge is facilitated, leading to what is referred to as the *radical non-constituent coordination*, whereby only basic types (and not functional types of any kind) are coordinated.

1991 Computing Reviews Classification System: F.4.1, F.4.2, J.5

Keywords and Phrases: Non-constituent coordination, Categorical Grammar, Hypothetical-reasoning, Lambek calculus

Note: work carried out under project INS.3, PDP.

1. INTRODUCTION

One of the acclaimed advantages of Categorical Grammar (CG) as a theory of the syntax of natural languages is its ability to smoothly provide a syntactic analysis for sentences with what is commonly known as *non-constituent coordination* (*ncc*); via the Curry-Howard correspondence, the semantics should fall out naturally too. Thus, coordinations like

- (1) Mary kissed John and hugged Bill
- (2) Mary kissed and Sue hugged John

are treated on a par. This is done¹ via the type assignment of $\forall X : ((X \rightarrow X) \leftarrow X)$ to all lexical conjunctions. We abbreviate below this type to κ . This type is used in conjunction with the following *ncc*-rule (which covers constituent-coordination as a special case): (*ncc*) $\frac{X, \kappa, X}{X}$ (which abbreviates two actual arrow-elimination steps – see below). Note that we use the more self-explanatory arrow-notation for categories, instead of the original slash-notation. To allow the whole range of *ncc* derivations (see [2] for the extent of *ncc*), the above rule is accompanied by composition-rules and type-raising rules, leading to what is known as *Combinatory Categorical Grammar* (*CCG*) (see [12]). As a consequence of the universal quantification on types in *ncc* (even when linguistically motivated restrictions are imposed on the possible values of X in κ), a proliferation of conjoinable types emerges. An alternative route to the additional rules is obtained by passing to more expressive formalisms, known as *Categorical Logics*, such as the various variants of the Lambek calculus \mathcal{L} , which have rules for *hypothetical reasoning*. In such systems, the additional rules become derived proof-rules. For a recent survey, see [7].

The question arises, what is the relationship between derivations using *ncc* and hypothetical reasoning. There seems to be no explicit account of such a relationship in the literature. In [7], There is an implicit consideration of this combination. The only example (on p. 114, repeated here in Section 2) has the following characteristic: when the *ncc*-rule is applied, both coordinated types had their assumptions all discharged. Thus, arrow-introduction rules (i.e., hypothetical reasoning), is applied before coordination via *ncc*. Similarly, all the examples in [8] (Ch. Three, sec. 3) have the same characteristic, although

¹We consider here the categorematic formulation as in [7]. There is a syncategorematic formulation too, e.g. in [12], which for **and** uses a rule $X \text{ and } X \Rightarrow X$, for any category X . In both formulations, sometimes some restrictions on the possible values of X are imposed.

the issue is not explicitly raised. A similar observation obtains regarding [1], where there is also only implicit consideration of this issue at hand, all the examples have the same characteristic mentioned above. Thus, even though the interaction of the ncc-rule with hypothetical reasoning is not explicitly stated to satisfy the above characteristic, it seems that this is the intended combination of an ncc-rule with \mathcal{L} . It is consistent with viewing the arguments coordinated via the ncc-rule as non-constituents.

Note that “blind application” of the ncc-rule in \mathcal{L} to types depending on undischarged assumptions is indeed not clear. Resorting only to peripheral abstraction, after conjoining two types (each of which peripherally depending, say, on one assumption), one of the two assumptions ceases to be peripheral and becomes unabstractable for the rest of the derivation; Resorting to some non-peripheral abstraction scheme may over-generate. An example is given in Section 2, after the notation is reviewed. Intuitively, the reason is that the two gaps that serve as assumptions should be filled by *the same filler* for the application of the ncc-rule to yield the right result. This is explained in more detail below.

Note that when the semantics is considered too, the notation of the ncc-rule is somewhat misleading; though all occurrences of X in the ncc-rule have the same category type as value (the one to which X is instantiated in κ), these occurrences have different interpretations as the semantic counterparts of their value. Thus, in the traditional “colon notation”, one would write the ncc-rule more fully as $(ncc) \frac{X:\alpha, \kappa:\lambda f.\lambda g.f \sqcap g, X:\beta}{X:\alpha \sqcap \beta}$ under the usual assumption that *all* semantic domains are closed w.r.t. ‘ \sqcap ’. Indeed, in [2], ‘ \sqcap ’ is not further analysed. However, in [12] a definition of conjoining functions is given, following the traditional approach found, for example in [9] and [4], according to which

$$(\lambda x_1.f(x_1) \sqcap \lambda x_2.g(x_2)) =_{df.} \lambda x.(f(x) \sqcap g(x))$$

(and similarly for Curried functions of greater arity). When used for non-constituent coordination, this definition has the property that while syntactically functor-categories are conjoined, semantically only result-category interpretations are conjoined. This mismatch suggest some syntactic operation, under which the semantic definition will fall-out by (a suitable extension of) the Curry-Howard correspondence. We mention here that the semantics of [12], as well as our revision presented below, deal only with narrow scope of quantifiers (under the usual GQ interpretation) w.r.t. coordination. To have a full control over scope, an orthogonal devise of scope modalities is used as an enhancement of \mathcal{L} , see e.g. [1] (and in particular its revised and extended version). We return to the relationship with semantics after the discussion of the revised ncc-rule.

In the present paper, we have two main purposes. On the more technical level, the paper provides an explicit treatment of the combination of the ncc-rule with \mathcal{L} . In particular, it provides a way of applying the ncc-rule *prior* to the arrow-introduction rules, under a certain side condition manipulating assumptions (thereby avoiding overgeneration), and only afterwards discharging the (manipulated) assumption(s). As for the application of \mathcal{L} as a formalism for grammar formulation for NL, the proposal here provides another view of non-constituent coordination. On the categorial level, the coordinated categories are always constituents. On the phrasal level, the coordinated phrases are what we call *pseudo-constituents*: constituents with gaps. This way, non-constituent coordination is perceived as a filler-gap structure, similar to extraction in relative clauses. The key observation is, that several gaps may be associated with the *same* filler. Thus, sentences such as (2) (repeated here as (3)) and (4) get the same treatment.

(3) Mary kissed and Sue hugged John

(4) The man whom Mary kissed and Sue hugged smiled.

This is an elaboration of the observation in [6], whereby ncc-sentences are viewed as having shared sub-structures. According to [6], sentence (3) would be represented as

Mary kissed
and John
(5) Sue hugged

while our proposal includes the gaps in the shared material:

Mary kissed _
and John
(6) Sue hugged _

$$\begin{array}{l}
(\text{ax}) \ A \triangleright A \\
(\rightarrow E) \ \frac{U_1 \triangleright B, \ U_2 \triangleright (B \rightarrow A)}{(U_1 U_2) \triangleright A}, \quad (\leftarrow E) \ \frac{U_2 \triangleright (A \leftarrow B), \ U_1 \triangleright B}{(U_2 U_1) \triangleright A} \\
(\rightarrow I) \ \frac{(BU) \triangleright A}{U \triangleright (B \rightarrow A)}, \quad (\leftarrow I) \ \frac{(UB) \triangleright A}{U \triangleright (A \leftarrow B)}
\end{array}$$

Figure 1: The \mathcal{L} -calculus

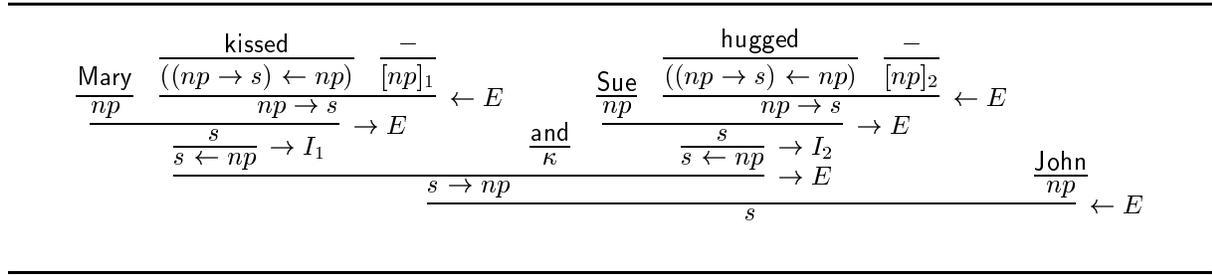
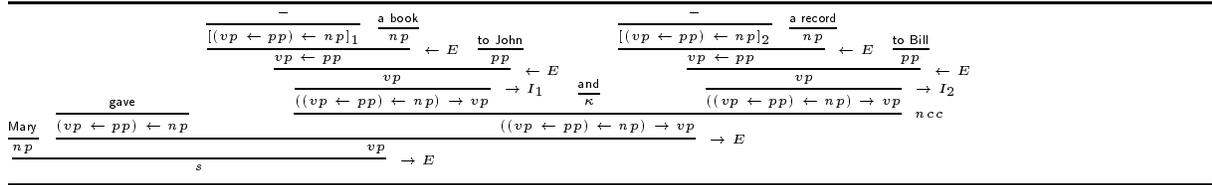
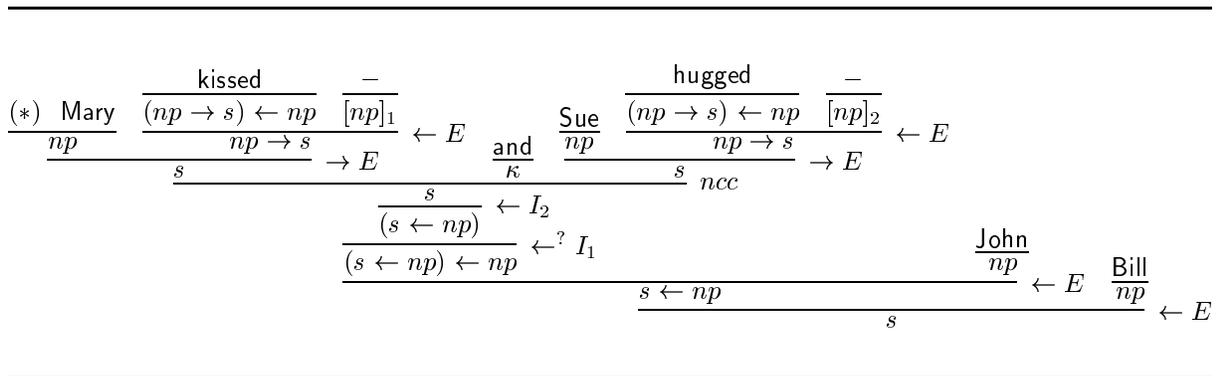
$$\begin{array}{l}
(\Rightarrow \tau) \ \frac{U \triangleright A}{U \triangleright (B \rightarrow A) \leftarrow B}, \quad (\Leftarrow \tau) \ \frac{U \triangleright A}{U \triangleright (B \leftarrow A) \rightarrow B} \\
(\Leftarrow c) \ \frac{U_1 \triangleright A \leftarrow B, \ U_2 \triangleright B \leftarrow C}{(U_1 U_2) \triangleright A \leftarrow C}, \quad (\Rightarrow c) \ \frac{U_1 \triangleright A \rightarrow B, \ U_2 \triangleright B \rightarrow C}{(U_1 U_2) \triangleright A \rightarrow C}
\end{array}$$

Figure 2: Additional CCG rules

The two shared substructures (with their gaps) will be referred to as pseudo-constituents, and the intended effect of the application of a revised ncc-rule is to *identify the two gaps into one gap, by identifying the assumptions (for hypothetical reasoning) introduced by these gaps*. Then, upon the application of an arrow-introduction rule, the conjoined pseudo-constituents have one antecedent, while discharging the assumption resulting from the identification; this represents a *simultaneous* discharge of the two original assumptions. Finally, by using a filler for the gap via an arrow-elimination rule, both original gaps can be perceived as having been filled by the same filler. In the final section, we present what we call *radical non-constituent coordination thesis*, that in a way eliminates the whole ncc as being different in essence from constituent-coordination. It enjoys the additional advantage of restricting coordination to basic categories only, thereby avoiding the above-mentioned proliferation of types. It also blocks some of the well-known overgenerated nccs by the traditional approach.

2. PRELIMINARIES

In this section we briefly review the basic notions and the notation used. The (associative) Lambek calculus \mathcal{L} [5] is a calculus for deriving valid *declarative units* of the form $U \triangleright A$, where U is a sequence of categories, and A is a category. The meaning of such a unit is that the sequence U reduces to A . The validity of such units is defined in terms of denotations over a string model (see [7]). A natural deduction version of the calculus for \mathcal{L} is presented in Figure 1. Here U represents the resources, sequences of categories, and A, B range over categories. Traditional categorial grammar is equivalent to the Ajdukiewicz-fragment \mathcal{A} , which does not have the hypothetical reasoning rules of arrow-introduction ($\leftarrow I$) and ($\rightarrow I$). The combinatory-categorial grammar (CCG) is obtained by augmenting \mathcal{A} with additional specific rules. Figure 2 presents two of these rules, the type-raising rules and the composition-rules. A detailed exposition, as well as the third rule, substitution (not needed for the current concerns), can be found in [12]. Note that the above rules are validity preserving and can be derived in \mathcal{L} , though not in \mathcal{A} .

Figure 6: A left-ncc derivation in \mathcal{L} Figure 7: A right-ncc derivation in \mathcal{L} Figure 8: Incorrect left-ncc \mathcal{L} -derivation with ncc-rule?

$$\begin{aligned}
(ncc) \quad & \frac{U_1 \triangleright X, U_2 \triangleright \kappa, U_3 \triangleright X}{(U_1 U_2 U_3) \triangleright X} \quad X \text{ assumption} - \text{independent} \\
(lncc) \quad & \frac{U_1[A]_1 \triangleright X, U_2 \triangleright \kappa, U_3[A]_2 \triangleright X}{(U_1 U_2 U_3[A]_{1 \equiv 2}) \triangleright X} \\
(rncc) \quad & \frac{[A]_1 U_1 \triangleright X, U_2 \triangleright \kappa, [A]_2 U_3 \triangleright X}{(U_1 U_2 U_3[A]_{1 \equiv 2}) \triangleright X}
\end{aligned}$$

Figure 9: The revised ncc-rules for \mathcal{L}

on assumptions of the same type and of equi-peripherality.

Let us restrict for a while the discussion to pseudo-constituents with a *single* gap, which is peripheral. Under the simplified single-gap assumptions, we alignment is possible in any of the following cases:

1. Both instances of X do not depend on any assumptions, as is the case in constituent-coordination; or
2. Both instances of X depend on left-peripheral assumptions (of the same type), as is the case for right-ncc; or
3. Both instances of X depend on right-peripheral assumptions (of the same type), as is the case for left-ncc.

The revised ncc-rule operates only on alignable types, by identifying the (equiperipheral) assumptions on which they depend, and replacing them by a *new* assumption of the same type, which is peripheral to the resulting type in the same way the original assumptions were peripheral to the combined types. An important effect of this definition is, that any continuation of a derivation that was possible for the original pseudo-constituents, is now possible for the resulting pseudo-constituent, where a discharge of an original assumption is replaced by a discharge of the combined assumption when an arrow-introduction rule is applied. In the general case (presented below), alignment of assumption has the same effect but for more general dependency on assumption than merely peripherality. The simplified revised ncc-rule (under the single-gap assumption) constitutes now of three separate rules, presented in Figure 9. The notational convention adopted in this rule is, that assumptions indexed by a single index are present in the original resources, while those indexed by an expression containing ‘ \equiv ’ are generated in the proof itself. Note that rule is a natural extension of the idea behind the original ncc-rule, that coordinated type should be identical. There is no semantic reason for disallowing coordination, say, of $np \rightarrow s$ and $np \leftarrow s$, both having predicates of the form $\lambda x[p(x)]$ as their meanings. However, that would allow derivation of

(8) (*) A man who - slept and Mary kissed -

which is syntactically bad. Note also, that the suggested rule is in accordance with the general point of view of Lambek-calculi as resource-management logics. Here, assumptions are recognized as special resources, and are given their own resource-manipulation rules. This resource-manipulation might be seen as a controlled-modulation (as alluded to in [7]) regarding a limited amount of the contraction structural-rule, absent in its full generality from \mathcal{L} , in the context of ncc.

Before turning to the semantics, let us first inspect some examples of the application of the revised ncc-rule. First, note that derivation of constituent-coordination, like the one in Figure 3, remain valid, as constituent-types are always alignable, not depending on assumptions. Figure 10 presents the derivation for (3). In this derivation, the lncc-rule discharges the two right-peripheral $np_i, i = 1, 2$ are discharged, and the resulting s depends on the newly-generated assumption (right-peripheral to the whole ncc) $[np]_{1 \equiv 2}$, the latter is then discharged by the arrow-introduction rule. Note that while the conjoined phrases are indeed pseudo-constituents (having an np-gap each in the object position), the categories conjoined (i.e., the value X is instantiated to in κ) are s (though depending still on assumptions). The revised derivation for a right-ncc is shown in Figure 11. To see further the effect of the revised ncc-rule, consider a derivation of a sentence containing “split constituent” ncc (in the nomenclature of [2]).

(9) Mary drove [to Chicago] yesterday and Detroit today.

The square brackets delimit the constituent, while the underline indicates the coordinated phrases. The derivation is presented in Figure 12. We use the derived rule $\Rightarrow_{\mathbf{C}}$ for shortening the derivation. As a more complicated example, in which the revised ncc-rule is applied twice, consider the derivation in Figure 13 of a two-sided ncc in

$$(ncch) \frac{[A^1]_{i_1} \dots [A^n]_{i_n} U_1 [B^1]_{k_1} \dots [B^n]_{k_m} \triangleright X, \quad U_2 \triangleright \kappa, \quad [A^1]_{j_1} \dots [A^n]_{j_n} U_3 [B^1]_{l_1} \dots [B^n]_{l_m} \triangleright X}{([A^1]_{i_1 \equiv j_1} \dots [A^n]_{i_n \equiv j_n} U_1 U_2 U_3 [B^1]_{k_1 \equiv l_1} \dots [B^n]_{k_m \equiv l_m}) \triangleright X}$$

Figure 15: The revised ncc-rule for general alignable types

meanings of the undischarged assumptions (say of type np) the respective coordinated types depend upon. Then, the type of the result is $\alpha(x_1) \sqcap \beta(x_2)[x_1 := x, x_2 := x]$, amounting to $\alpha(x) \sqcap \beta(x)$, where x is *fresh* variable, not used in the derivation so far. It is the meaning of the generated assumption $[np]_{1 \equiv 2}$. Here $\alpha[x := y]$ means the substitution of the variable y for all free occurrences of x in α . This semantic rule captures the intention of identifying the two assumptions semantically. Note that after applying later an arrow-introduction rule, say $\leftarrow I_{1 \equiv 2}$, the obtained meaning of the result becomes $\lambda x.(\alpha(x) \sqcap \beta(x))$, the exact meaning stipulated by Steedman as holding by definition, as discussed above. The substitution of a fresh variable for all occurrences two different free variables (in the two coordinated meanings) can be seen as the extension of the Curry-Howard correspondence to the syntactic operation manifesting itself in the revised ncc-rule. It might be desirable to separate the assumption-manipulation from the coordination-rule, and make it an autonomous operation in the calculus, potentially having additional applications. We do not have any other application at this stage, and thus keep the assumptions-manipulation as local to the revised ncc-rule. We note while passing that a treatment of ncc using open formulae is hinted at in [3]. There, the main argument is about the role of variables in the semantics, and the main issue is pronoun binding. In a brief consideration of an ncc example, Jacobson hints towards a conjunction of formulae with free-variables, but nothing relates it to hypothetical reasoning. The semantics of conjunction is stipulated similar to Steedman's. Furthermore, as gaps in (what we call) pseudo-constituents are referred to as 'traces' (movement?), and not viewed as assumptions for hypothetical reasoning. The two gaps in the syntax of the ncc are not distinguished to start with, so the issue of their identification does not arise.

We now turn to the formulation of the revised ncc-rule for the general case of alignable categories, with any number of equal-type, equal-peripherality assumptions. Recall that \mathcal{L} has only peripheral hypothetical-reasoning. Adding non-peripheral one would inquire a corresponding change in the revised ncc-rule. Henceforth, we refer to the revised ncc-rule as *ncch* (ncc with hypothetical types). The rule is schematic over the numbers of assumptions. It is presented in Figure 15. All superscripts are for distinguishing types; all subscripts, assumed pairwise different, are for numbering assumptions.

Note that the special-case rule presented before is obtained, respectively, for the three cases $n = m = 0$, $n = 0 \wedge m = 1$ and $n = 1 \wedge m = 0$. Note also how the rule identifies corresponding pairs of equi-typed, equi-peripheral assumptions, to generate for each pair a combined assumption with the same type and peripherality. We defer the presentation of examples of applications of the ncch-rule to Section 4, where a more radical view of the ncc phenomenon is presented, based on ncch.

An alternative implementation of the idea of identifying assumptions, leading to a late discharge, can be obtained by using a multi-modal logic with a "bracketing modality" (see [7], Ch. 4, and the references there), that would introduce another composition operation, that distributes over products within a bracketed context. This would also relax locally the absence of a contracting structural rule. We do not pursue here further this possibility.

4. RADICAL NON-CONSTITUENT COORDINATION

Let us start by considering an alternative derivation for (7), as presented in Figure 16. First, note that this derivation uses two left-peripheral assumptions, hence ncch is indeed needed. The fact we would like to draw attention to is, that the coordinated type is s . This is not accidental! By resorting to ncch, one can always introduce sufficiently many assumptions, so as the pseudo-constituents coordinated are open sentences (with a semantics containing free variables for all assumptions). There is one exception, though, for np-coordination, having a non-distributive predication, like

(13) Mary and John met.

For a recent account of boolean semantics for np-coordination (both collective and distributive) see [13]. We may, therefore, conclude, that nothing is lost if the logic *forbids* ncc-coordination of functional

5. ACKNOWLEDGMENTS

This research was carried out under a grant B 62-443 from Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO). I wish to thank NWO for their support. I thank Jan van Eijck, Michel Moortgat and Yoad Winter for some useful comments on a preliminary draft of the paper.

REFERENCES

1. Bob Carpenter. A deductive account of scope. In *Proc. of the 13th west coast conference on formal linguistics*, San-Diego, CA, 1994. CSLI publications. Extended version submitted to *Linguistics and Philosophy*.
2. David Dowty. Non-constituent coordination, wrapping, and multimodal categorial grammars. In *Proc. of the international congress on logic, methodology, and philosophy*, Florence, Italy, August, 1996. Kluwer.
3. Pauline Jacobson. The locality of interpretation: the case of binding and coordination. In *Proceedings of the 6th conference on semantics in Linguistic Theory (SALT)*. Cornell working papers in Linguistic, 1996.
4. Ed Keenan and Leonard Faltz. *Boolean semantics for natural language*. Reidel, Dordrecht, 1985.
5. Joachim Lambek. The mathematics of sentence structure. *American mathematical monthly*, 65:154 – 170, 1958.
6. David Milward. Non-constituent coordination: Theory and practice. In *Proc. of the 15th international conference on Computational linguistics (COLING94)*, pages 935–941, Kyoto, Japan, 1994.
7. Michael Moortgat. Categorial type logics. In Johan van Benthem and Alice ter Muelen, editors, *Handbook for Logic and Language*. Elsevier, 1997.
8. Glyn V. Morrill. *Type Logical Grammar: Categorial Logic of Signs*. Kluwer Academic Publishers, 1994.
9. Barbara Partee and Mats Rooth. Generalized conjunction and type ambiguity. In Rainer Bäuerle, Christoph Schwartze, and Arnim von Stechow, editors, *Meaning, use and interpretation of language*. De Gruyter, Berlin, 1983.
10. Dag Prawitz. *Natural deduction*. Alqvist and Wiksell, Uppsala, 1965.
11. Mark Steedman. Categorial grammar (tutorial overview). *Lingua*, 90:221–258, 1993.
12. Mark Steedman. *Surface Structure and Interpretation*. MIT Press, Linguistic Inquiry monographs 30, 1996.
13. Yoad Winter. A unified semantic treatment of singular np coordination. *Linguistics and Philosophy*, 19:337–391, 1996. See also forthcoming (1998) Ph.D thesis.