Resource sharing in Type Logical Grammar

Gerhard Jäger

1 Introduction

Categorial Grammars are generally resource conscious formal systems. They describe how linguistic signs can be combined and transformed, but linguistic material is usually neither multiplied nor destroyed under categorial analyses. In other words, Categorial Grammars avoid counterparts of copy transformations and deletion rules.

Given this, anaphora phenomena like pronoun binding and ellipsis are a challenge for this family of frameworks, since anaphoric expressions by definition use semantic resources without consuming them. Thus it is not surprising that especially bound pronouns have received considerable attention in the literature.

If one aims at incorporating anaphora binding into the core of the Categorial machinery—this means that binding is not achieved by employing variable binding devices, one has to chose between two options. Either one leaves the general resource conscious setup of Categorial Grammar intact and locates the resource multiplying behavior of anaphors in their lexical entry. Representatives of this variety are for instance [Szabolcsi, 1989] in a Combinatory setting, and [Moortgat, 1996] under a Type Logical perspective.

Alternatively, one can extend the admissible grammatical operations to cover the kind of resource multiplication that occur in connection with anaphora in syntax. Pioneering work along this line has been done by Jacobson ([Jacobson, 1999], Jacobson, this volume) within the Combinatory framework, partially building on Type Logical research by [Hepple, 1990].

The present article pursues the second line of research within Type Logical Grammar. We will propose an extension of the Lambek calculus with a new type logical connective that is specifically designed to deal with anaphoric resource multiplication. It is modeled after Jacobson's analogous Combinatory device. Section 2 introduces the proof theory of this system and presents some of its meta-logical properties. Section 3 sketches its linguistic application to basic cases of pronominal anaphora and Verb Phrase Ellipsis (VPE) in English.[1] Sections 4 and 5 discuss more complex cases of VPE in English and motivate a refined lexical analysis of the auxiliaries in VPE constructions. Section 6 summarizes the findings.

[1] These sections partially overlap with [Jäger, 2000], where this formal system was originally introduced.

2 The system in theory

We choose the associative Lambek calculus **L** (cf. [Lambek, 1958]) as background mainly for reasons of simplicity.[2] Following ideas from [Morrill et al., 1990], we use a Natural Deduction presentation of **L** and of the extension that we are going to propose. This has mainly pragmatic motivation; experience shows that sequent systems and axiomatic systems are useful to prove statements about a given calculus, while Natural Deduction is the most convenient way to work within the system. The Introduction and Elimination rules of the Lambek connectives are given in figure 1. Following standard praxis, we use sequences of formulas as premises of sequents, i.e. the associativity of the calculus is compiled into the presentation format.

Now let us turn attention to the intended extension of **L** with logical means to handle anaphora phenomena. Following Jacobson's suggestion, we assume that the anaphoric character of a given linguistic expression is mirrored in its type. Our notational convention is the following: An anaphoric expression that requires an antecedent of type B and—provided such an antecedent is around—behaves like an expression of type A will receive type $A|B$. Since an anaphoric pronoun behaves like an N provided it finds an antecedent of type N, it will receive type $N|N$. What remains to be specified is the structural configuration that has to hold between an anaphor and its antecedent. Here we only require that the antecedent precedes the anaphor. Empirical support for this claim will be given below.

Given this, we can give a semi-formal definition of anaphora in our Categorial setting:

> An expression a has type $A|B$ if and only if
> a behaves like an expression of type A whenever it is preceded by an expression of type B.

We also follow Jacobson in her treatment of the semantics of anaphora resolution. The meaning of an anaphor is a function from the meaning of its antecedent to its concrete meaning in the given context. This implies that anaphora resolution comes down to function application semantically; the meaning of the anaphor is applied to the meaning of its antecedent. Since the contextual meaning of an anaphoric pronoun is identical to the meaning of its antecedent, the lexical meaning of a pronoun is thus the identity function $\lambda x.x$.

The Natural deduction rules for the anaphora slash $|$ at the end of figure 1 mimic the quasi-definition given above. The elimination rules $|E_{1/2}$ correspond to the "if" part: if we can prove that certain resources Z have type $C|A$, and they are preceded by resources X from which we can derive the antecedent type A, we may safely consider Z to be of type C. Since anaphora resolution has to take place within one proof, all resources involved have to be glued together by means of the product operator. It does not play any role for this operation whether additional material Y intervenes between antecedent and anaphor as in $|E_1$ or they are adjacent as in $|E_2$.

The introduction rule $I|$ corresponds to the "only-if" part of the definition. To show that certain resources X behave like an A whenever they are preceded by a B, we have to assume a hypothetical antecedent B to the left of X. Then we have to prove that X

[2] See [Jäger, 1999] for a more general setting.

can be transformed into an A without further assumptions. This proof must not depend on the question whether or not the hypothetical antecedent is adjacent to X—therefore we distinguish two premises in the rule, one with and one without a variable p over intervening material—and in case some material intervenes, we must not make any assumption on its type or semantics. (This is ensured by the side conditions that p and y do not occur anywhere else in the proof.) If this succeeds, we may infer that X has type $A|B$, and we derive its semantics by λ-abstracting over the meaning of the hypothesized antecedent.

In [Jäger, 2000] the proof theoretic properties of the resulting calculus $\mathbf{L}_|$ are investigated. We give the main results here without proof:

Theorem 1 (Decidability): $\mathbf{L}_|$ is decidable.

Theorem 2 (Finite Reading Property): $\mathbf{L}_|$ has the finite reading property.

These results are noteworthy since anaphora involves sharing, i.e. multiplication of resources, and logics that admit this usually lack the finite reading property.

The proofs of these two theorems rely heavily on the fact that the Gentzen style sequent calculus for $\mathbf{L}_|$ is well-behaved, i.e. it admits Cut elimination. The Natural Deduction calculus is well-behaved in an analogous way: all proofs are strongly normalizable.

Theorem 3 (Strong Normalization): There are neither infinite sequences of normalization steps of Natural Deduction proofs nor infinite sequences of Cut elimination steps of sequent proofs in $\mathbf{L}_|$.

The proof is given in the appendix.

Natural deductions are more conveniently carried out in tree form. The building blocks are given in figure 2. Note that a complete deduction always ends in a single conclusion, despite the fact that $\bullet$-elimination and $|$-elimination have multiple conclusions. For simplicity we combined the $|$-elimination rule with two subsequent applications of the $\bullet$-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.

For better readability and to stress the similarity to conventional coindexing of constituents, we simplify the notation for $|E$ somewhat (see figure 3).

When 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 figure 4. 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 $\backslash I$ 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.

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

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

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

$$\frac{x : A, X \Rightarrow M : B}{X \Rightarrow \lambda x.M : A \setminus B}\, \setminus I$$

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

$$\frac{X, x : A \Rightarrow M : B}{X \Rightarrow \lambda x.M : B/A}\, /I$$

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

$$\frac{X \Rightarrow M : A \qquad Y \Rightarrow N : B \qquad Z \Rightarrow O : C|A}{X, Y, Z \Rightarrow \langle M, N, (OM) \rangle : A \bullet B \bullet C}\, |E_1$$

$$\frac{X \Rightarrow M : A \qquad Z \Rightarrow O : C|A}{X, Z \Rightarrow \langle M, (OM) \rangle : A \bullet C}\, |E_2$$

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

p not occurring in A, B, X
y not occurring in M

Fig. 1: Natural Deduction $\mathbf{L}_|$

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

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

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

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

Fig. 2: Natural deduction in tree format

3 The system at work

3.1 Pronouns

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

(1) John said he walked

where the only potential antecedent of the pronoun is a proper noun, we have the two possible derivations shown in figure 5, 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

(2) said he walked

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

Fig. 3: Simplified notation for $|E$

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

Fig. 4: An illicit Natural Deduction derivation

the pronoun can be linked to the subject argument place of the VP, as figure 6 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 [Moortgat, 1996]). 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 figure 7. The elimination rule involves hypothetical reasoning and can thus lead to binding of anaphors. Let us consider the example

(3) 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 figure 8.

If we reverse the order of the quantifier and the pronoun as in (4), the derivation of a bound reading will fail, even though the pronoun is in the scope of the quantifier.

(4) *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

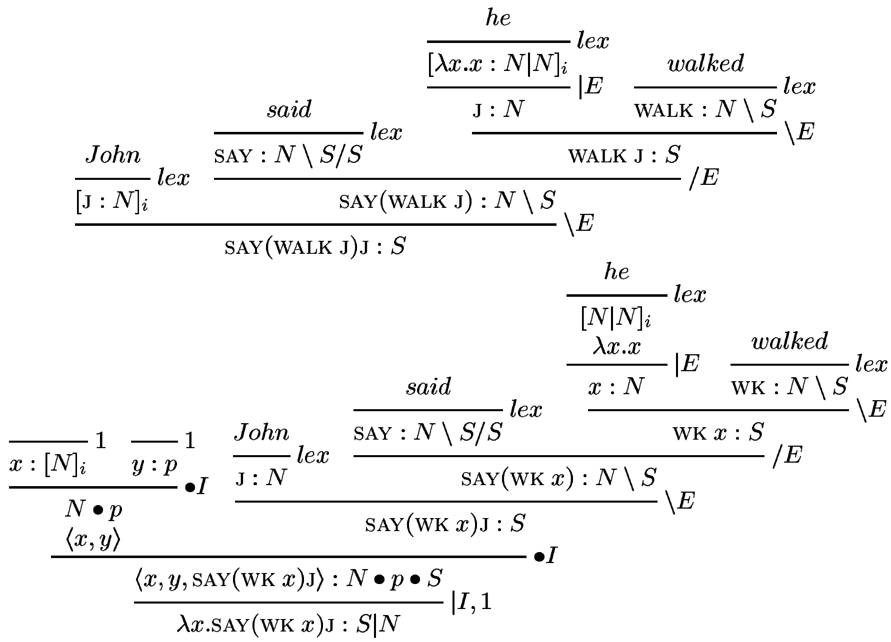

Fig. 5: Derivations of *John said he walked*

scope of its binder. This excludes for instance a wide scope reading of the indefinite object in (5) if the pronoun is bound by the subject.

(5) Every man saw a friend of his

The way the present system excludes such readings is similar to the one proposed in [Pereira, 1990], 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 (6), where the binding quantifier does not c-command the pronoun under the standard conception of constituency, does not differ substantially from the previous example (figure 9).

(6) 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.

(7) *His mother loves everybody

So the precedence requirement also accounts for Weak Crossover violations.

$$\dfrac{\dfrac{\dfrac{}{[x:N]_i}\,1 \quad \dfrac{\dfrac{said}{\textsc{say}:N\setminus S/S}\,lex \quad \dfrac{\dfrac{\dfrac{he}{[\lambda x.x:N|N]_i}\,lex}{x:N}\,|E \quad \dfrac{walked}{\textsc{walk}:N\setminus S}\,lex}{\textsc{walk}\ x:S}\,\setminus E}{\textsc{say}(\textsc{walk}\ x):N\setminus S}\,/E}{\textsc{say}(\textsc{walk}\ x)x:S}\,\setminus E}{\lambda x.\textsc{say}(\textsc{walk}\ x)x:S}\,\setminus I,1$$

Fig. 6: Derivation of *said he walked*

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

Fig. 7: Natural Deduction rules for $q(A,B,C)$

3.2 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 (8) as a proform for VPs.

(8) 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 (8) is given in figure 10 (we also ignore the contribution of *too* since it is irrelevant for the semantics of VPE, though not for the pragmatics).
What makes VPE an interesting topic is of course its complex interaction with pronominal anaphora and quantification. The first non-trivial issue in this connection is the well-known strict/sloppy ambiguity in constructions like (9).

(9) 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 figure 6: the pronoun is bound to the subject argument place of the source VP. From this we can continue the derivation completely in parallel to figure 10, and we end up with the meaning $\textsc{and}(\textsc{say}(\textsc{walk}\ \textsc{b})\textsc{b})\ (\textsc{say}(\textsc{walk}\ \textsc{j})\textsc{j})$. Crucially, here the pronoun was bound by a hypothetical antecedent. Of course it is also licit to bind

$$\dfrac{\dfrac{\dfrac{\dfrac{\textit{everybody}}{\text{EVERY} : q(N,S,S)}\,lex}{[x:N]_i}\,1 \quad \dfrac{\dfrac{\textit{said}}{\text{SAY} : N \backslash S/S}\,lex \quad \dfrac{\dfrac{\dfrac{\textit{he}}{[\lambda x.x : N|N]_i}\,lex}{x : N}\,|E \quad \dfrac{\textit{walked}}{\text{WALK} : N \backslash S}\,lex}{\text{WALK}\ x : S}\,\backslash E}{\text{SAY}(\text{WALK}\ x) : N \backslash S}\,/E}{\text{SAY}(\text{WALK}\ x)x : S}\,\backslash E}{\text{EVERY}(\lambda x.\text{SAY}(\text{WALK}\ x)x) : S}\,qE, 1$$

Fig. 8: Derivation of *Everybody said he walked*

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\textit{everybody}}{\begin{array}{c}q(n,s,s)\\ \text{EVERY}\end{array}}\,lex}{\begin{array}{c}[n]_i\\ y\end{array}}\,1 \quad \dfrac{{}'s}{\begin{array}{c}n \backslash n/cn\\ \text{OF}\end{array}}\,lex}{\begin{array}{c}n/cn\\ \text{OF}\ y\end{array}}\,\backslash E \quad \dfrac{\textit{mother}}{\begin{array}{c}cn\\ \text{MOTHER}\end{array}}\,lex}{\begin{array}{c}n\\ \text{OF}\ y\text{MOTHER}\end{array}}\,/E \quad \dfrac{\dfrac{\textit{loves}}{\begin{array}{c}n \backslash s/n\\ \text{LOVES}\end{array}}\,lex \quad \dfrac{\dfrac{\textit{him}}{\begin{array}{c}[n|n]_i\\ \lambda x.x\end{array}}\,lex}{\begin{array}{c}n\\ y\end{array}}\,|E}{\begin{array}{c}n \backslash s\\ \text{LOVE}\ y\end{array}}\,/E}{\begin{array}{c}s\\ \text{LOVE}\ y(\text{OF}\ y\text{MOTHER})\end{array}}\,\backslash E}{\begin{array}{c}s\\ \text{EVERY}(\lambda y.\text{LOVE}\ y(\text{OF}\ y\text{MOTHER}))\end{array}}\,qE, 1$$

Fig. 9: Everybody's mother loves him

to pronoun to the actual subject *John* and then doing ellipsis resolution, which results in the strict reading. The derivation of both readings is given in figure 11.
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.

(10) 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 [Sag, 1976], constructions like (10a) are ambiguous between a reading synonymous to (10b) and one synonymous to (10c). Under the present approach, reading (10b) 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

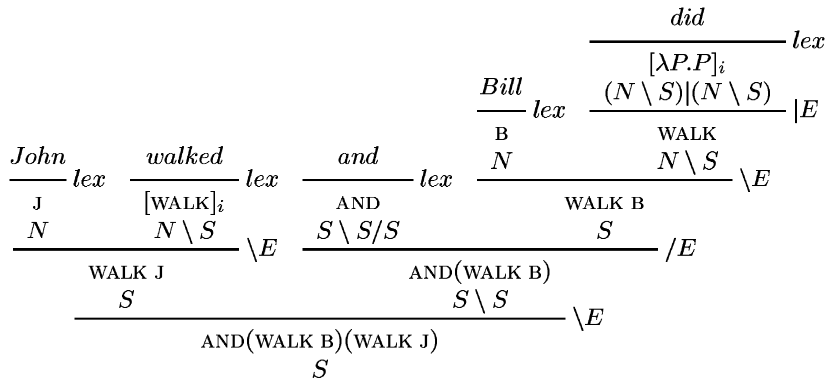

Fig. 10: John walked, and Bill did (too)

corresponding variable. Figure 12 gives the derivations for the source VPs of the two readings.

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 [Sag, 1976].

(11) 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 (11a) is *the man that Mary met before Bill met him*. Example (11b) 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.
Let me close this section with an example from [Gawron and Peters, 1990] that demonstrates that the ambiguity of bound versus coreferential interpretation of pronouns on the one hand and the strict/sloppy ambiguity on the other hand are independent phenomena:

(12) 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 (12a) has three reading (paraphrased in (12b-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

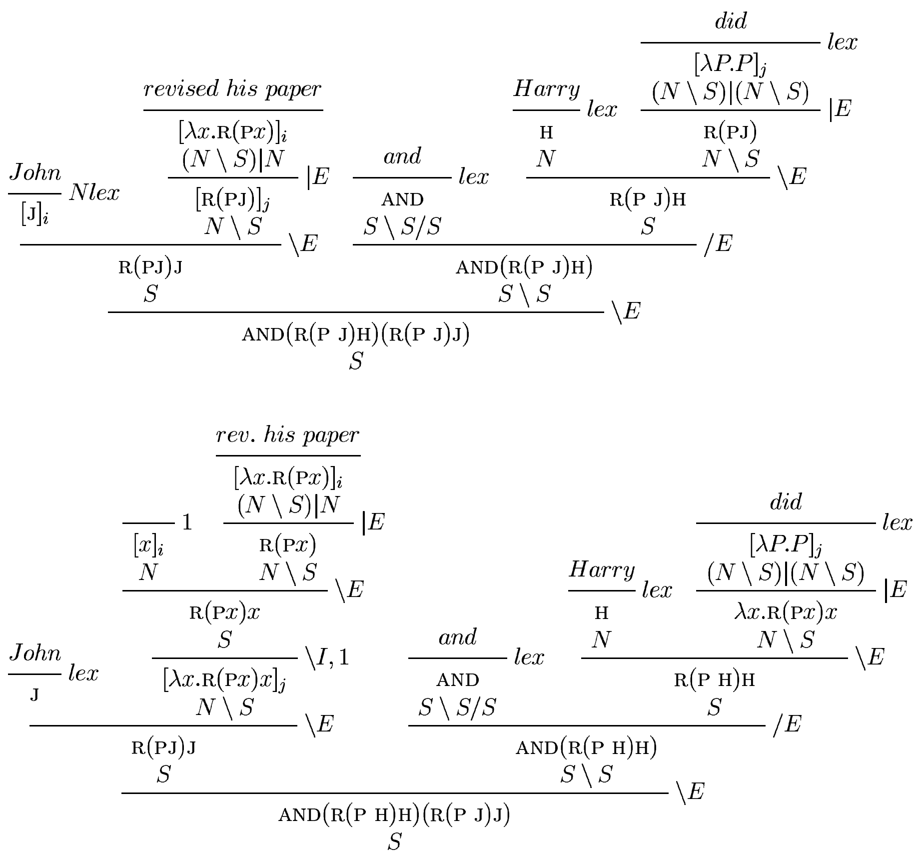

Fig. 11: Derivation of the strict and the sloppy reading of (9)

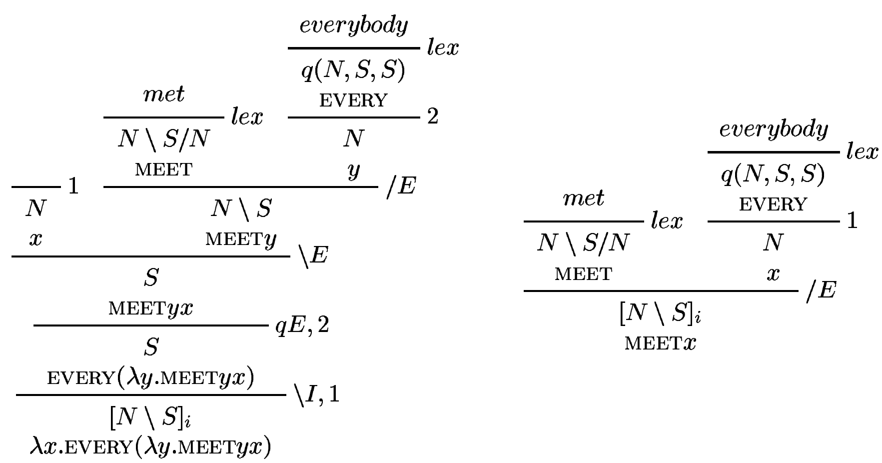

Fig. 12: Source VPs in (10a)

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. [Gawron and Peters, 1990]).
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 figure 5. Readings (12c,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 figure 11), and applying qE and thus replacing the hypothetical N by the quantifier.

4 VPE and Polymorphism

The approach to VP ellipsis presented in the last section belongs to the family of "identity-of-property" theories for VPE. Following basically [Sag, 1976], these theories assume that the source VP and the elliptical VP express the same *property* at some level of derivation or representation. This idea is in sharp contrast with theories like [Fiengo and May, 1994], where VPE is basically seen as involving identical syntactic structure that is not pronounced in the elliptical part. In the sequel we will discuss several problems for an identity-of-property approach that have been discussed in the literature, and we will demonstrate that an identity-of-meaning approach can be maintained if we admit a limited polymorphism in the lexicon, in a way akin to the standard categorial treatment of coordination.

4.1 The Hirschbühler problem

[Hirschbühler, 1982] notes that in the following example, the subject can take wide scope in both conjuncts.

(13) A Canadian flag was hanging in front of each window, and an American one was, too

In the preferred reading, there is one American and one Canadian flag per window. Hirschbühler considered the option that this reading arises because the object *each window* scopes over the whole construction, including the conjunction. This would render the example analogous to (10). However, such a solution would fail, as Hirschbühler points out. We observe a similar reading in (14).

(14) A Canadian flag was hanging in front of many windows, and an American one was, too

The preferred reading here is the one where the object takes scope over the subject in both conjuncts, but the conjunction still takes scope over both objects.
Identity-of-property approaches to VPE are unable to derive this reading. To see why, one has to consider what potential antecedent properties the source clause supplies here. The syntactic antecedent in the last example is *was hanging in front of many windows*. This VP is entirely unambiguous; the only meaning of type property that can be derived from it is the one where the object scopes over the VP:

$\lambda x.(\text{MANY WINDOWS}(\lambda y.\text{WAS\_HANGING\_IN\_FRONT\_OF}xy))$. Combining this meaning with either the source subject or the target subject yields inevitably the subject wide scope reading.
Even though several attempts have been undertaken to treat this kind of example within an identity-of-property approach, none of them was really successful. The Hirschbühler problem effectively falsifies this group of ellipsis theories.
It does not falsify a somewhat more general setup, something which has been called "identity-of-meaning" theories. It maintains the basic intuition that it is the meaning that is shared between source and target in VPE construction rather than syntactic structure, but it possibly gives up the assumption that this has to be a property. Under a flexible approach to meaning assignment, a phrase like *was hanging in front of many windows* may receive different meaning with different types. The key example for this more flexible treatment is [Kempson and Cormack, 1983]. They claim that the piece of meaning that is shared between source VP and ellipsis site is not a property of individuals but a property of quantifiers. A VP containing a quantified object will be ambiguous in this type, which in turn leads to the Hirschbühler ambiguity in ellipsis. To be somewhat more specific, the VP in question is ambiguous between the lifted properties $\lambda T.T(\lambda x.(\text{MANY WINDOWS}(\lambda y.Ryx)))$ and $\lambda T.\text{MANY WINDOWS}(\lambda y.T(\lambda x.Ryx))$, where R stands for the meaning of *hanging in front of*. The former meaning assignment leads to a reading where the subject has wide scope in both conjuncts, while the latter one gives the critical Hirschbühler reading.
Flexible meaning assignment is an essential aspect of any Categorial Grammar, so the Kempson/Cormack style treatment is easily to incorporate in the present theory of ellipsis resolution. To start with, even though categorial meaning assignment is flexible, the category-to-type correspondence between syntax and semantics is strict. So assigning the string *was hanging in front of many windows* a meaning of a higher type implies assignment of a more complex syntactic category. The obvious candidate is $(S/(N \setminus S)) \setminus S$, i.e. a functor that consumes a subject quantifier to its left to yield a clause. So the only adjustment that is necessary to adopt Kempson and Cormack's analysis is a modification of the lexical assignment for the auxiliary in VPE construction: instead of the identity function over properties, we assign it the identity function over properties of quantifiers, paired with the appropriate syntactic type. So the modified lexical entry is

(15) did/was – $\lambda x.x : ((S/(N \setminus S)) \setminus S)|((S/(N \setminus S)) \setminus S)$

The derivation of the lifted source VP of the Hirschbühler reading of (16) is given in figure 13.

(16) A doctor visited every patient, and a nurse did too.

It should be noted that due to the built-in flexibility of Type Logical Grammar, this approach overgenerates. The Hirschbühler examples admit scope inversion, but only if it occurs both in the source clause and the target clause. A reading where the subject takes wide scope in the source clause and narrow scope in the target clause is excluded. In the present setup, such crossed readings are derivable, however. This is due to the fact that argument lowering is a theorem of **L** (and thus of $\mathbf{L}_{|}$):

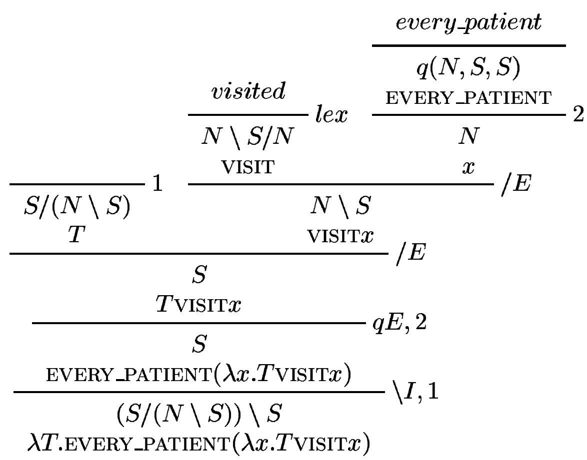

Fig. 13: *visited every patient* – object wide scope

$$\vdash x : (S/(N \setminus S)) \setminus S \Rightarrow \lambda y.x(\lambda z.zy) : N \setminus S$$

Now suppose we assign the source VP the object wide scope reading as in the sample derivation in figure 13, combine it with the source subject directly, but let the copy undergo argument lowering before we combine it with the target subject. This will result in a reading where the object has wide scope in the source but narrow scope in the target clause.

One might wonder though whether the kind of parallelism effects that we observe here should really be treated as a property of ellipsis resolution as such. Arguably, there is a parallelism constraint in coordinate construction anyway, quite independently of ellipsis. So possibly grammar in fact admits crossed readings, while pragmatics filters them out. This would lead to a kind of hybrid theory of VPE, where the syntax/semantics interface is fairly liberal as far as admissible readings are concerned, while pragmatic constraints that are basically independent of the ellipsis module are responsible for the fine tuning. Circumstantial evidence for such an architecture will be collected in the next subsection.

4.2 Non-subject sloppy readings

Even more problematic for an identity-of-property approach are cases where the antecedent for a sloppy pronoun is not the subject of the source VP. Possible antecedents can be

NPs embedded in the subject

(17) a. John's coach thinks *he* has a chance, and Bill's coach does too. ([Rooth, 1992])

b. People from LA adore *it* and people from NY do too. (after [Reinhart, 1983])

c. The policeman who arrested John failed to read *him his* rights, and the one who arrested Bill did too (after [Wescoat, 1989], cited from [Dalrymple et al., 1991])

NPs embedded in topic

(18) If Bill was having trouble in school, I would help *him*. If Harry was having trouble in school, I wouldn't (after [Hardt, 1993])

NPs from superordinated clauses

(19) I didn't know that Bill was a bigamist. Mary just said he's married to *her*, and Sally did, too. (from [Fiengo and May, 1994])

The sloppy pronouns are marked by italic font, and their antecedents by underlining. The first descriptive hypothesis about sloppy readings that comes to mind in view of these data is that the two antecedents of a sloppy pronoun must occupy structurally parallel positions in the source clause and the target clause (this is for instance assumed in [Fiengo and May, 1994]). However, this is shown to be too rigid by [Rooth, 1992] ((a) and (b)) and [Hardt, 1993] (c):

(20) a. First John told Mary that I was bad-mouthing *her*, and then Sue heard that I was.

b. Yesterday the guy John works for told *him* to shape up, and today Bill's boss did.

c. If John was having trouble in school, I would help *him*. On the other hand, if Bill was having trouble, I doubt if I would.

So apparently a notion of semantic rather than structural parallelism is called for, which may be enriched by some notion of "implicational bridging" ([Rooth, 1992]) to cover cases like (20a). This approach, however, turns out to be too narrow too, as the following example from [Fiengo and May, 1994] demonstrates.

(21) First John told Mary that I was bad-mouthing *her*, and then Sue behaved as though I would.

We do not have to offer a novel account of the structural/semantic/pragmatic relation here that has to hold between source and target in VPE here. What the examples above do show is that whatever governs the distribution of non-subject sloppy readings, it is certainly not determined by grammar in the narrow sense. The only (trivial) grammatical constraint seems to be that the sloppy pronoun has to find an antecedent in the pre-VP material of both clauses.

Even though an identity-of-property approach to VPE is incapable to cover any non-subject sloppy reading, these data are not overly problematic for an identity-of-meaning program if pronouns are analyzed in a variable free way. Let us take the intuition "the elliptical VP has to find an antecedent in the pre-VP material of both clauses" seriously.

To put this idea slightly differently, what is shared between source clause and target clause in a VPE construction is the meaning of a VP that may contain a series of pronouns which are bound inside the source clause and in the target clause respectively. (The source clause and the target clause need not be the local clauses, as the example (19) demonstrates.) Let us restrict the discussion to cases with one pronoun for a moment. Basically the category of a VP containing one pronoun is $(N \setminus S)|N$. Let us abbreviate this category with VP^1. To enforce binding of this pronoun within a superordinate clause, this type has to be lifted to

(22) $(S/\mathrm{VP}^1) \setminus S$

Note that after lifting, the VP in question does not contain any unresolved pronouns any longer. This can be generalized to an arbitrary number of pronouns in a simple way: Let us say that $\mathrm{VP}^0 = N \setminus S$ and $\mathrm{VP}^{n+1} = \mathrm{VP}^n|N$. The general type scheme for lifted VPs is then $(S/\mathrm{VP}^n) \setminus S$ for arbitrary natural numbers n. Accordingly, we assume a polymorphic lexical entry for the auxiliary, namely the identity function over all instances of lifted VPs.

(23) $\lambda x.x : ((S/\mathrm{VP}^n) \setminus S)|((S/\mathrm{VP}^n) \setminus S)$

Note that the proposal made in the last subsection is just a special case of this where $n = 0$.
To see how this proposal works, consider a simple example like

(24) John's father helps him, and Bill's father does too

The derivation of the source clause is given in figure 14. In an intermediate step of the derivation, the string *helps him* is assigned the lifted VP category $(S/\mathrm{VP}^1) \setminus S$, paired with the meaning $\lambda T.T\,\textsc{help}$. This piece of meaning serves as antecedent for ellipsis resolution. The derivation of the target clause runs completely in parallel, except for the fact that the lifted VP is not lexically founded but retrieved from the source clause via $|$-elimination. So the meaning of the target clause winds up being $(\textsc{help}\ \textsc{b}(\textsc{of}\ \textsc{b}\ \textsc{father}))$—*Bill's father helps Bill*.

5 Parallelism versus source ambiguity

On a somewhat less technical level, the lexical entry for the auxiliary given in (23)—paired with the general approach to anaphora presented in this paper—leads to one constraint on sloppy readings of VPE: a given sloppy pronoun has to find its antecedents in the pre-VP material of the source clause and the target clause (or clauses in the case of multiple ellipsis) respectively. Since this is a very mild constraint indeed, it is not very surprising that most examples that are discussed in the literature can be derived in such a system. A notoriously difficult one is due to [Dahl, 1974].

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

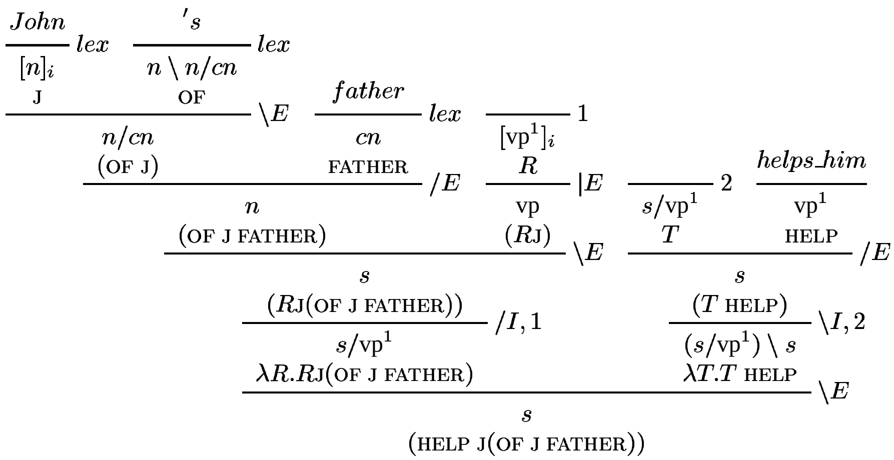

Fig. 14: Source clause of *John's father helps him, and Bill's father does too*

The critical reading is the one where John realizes that John is a fool, Bill fails to realize that Bill is a fool, but Bill's wife realizes that Bill is a fool. So apparently the second clause takes the first clause as antecedent and receives a sloppy reading, while the third clause is anaphoric to the second clause and strict. Under an identity-of-meaning theory, this configuration should be impossible. Another analysis is possible though. We may analyze both ellipses as taking the first clause as antecedent and receiving a sloppy construal. The second ellipsis is extremely sloppy because it takes the possessor of the subject as antecedent of the sloppy pronoun rather than the subject itself.

Liberal though the present theory may be, it is not entirely unconstrained. In particular, it predicts a fundamental asymmetry between VPE in coordination and subordination. In subordinative constructions, it is as constrained as the traditional [Sag, 1976] style theory.

To place this aspect into the right perspective, let us briefly return to the general issue: Does VPE involve identity of meaning? We have argued above that such a theory has to be paired with some theory of parallelism to cope with the problem of overgeneration. Given this, it is suggestive to totally trivialize the operation of ellipsis resolution ("fill in whatever gives you a sentence") and locate all interesting generalizations inside the parallelism module. This idea has been pursued by many authors, most prominently by [Dalrymple et al., 1991, Rooth, 1992], and [Shieber et al., 1996].

Reconsider a simple strict/sloppy ambiguity like

(26) John revised his paper, and Bill did too

An identity-of-meaning approach has to assume that the source VP is ambiguous between *to revise John's paper* and *to revise one's own paper*. Outside ellipsis construction, this ambiguity is spurious, but it leads to different truth conditions for the target clause in VPE.

A purely parallelism based approach can do without this kind of spurious ambiguity. Informally put, the mentioned theories require only that replacing *Bill* by *John* in the target clause leads to the same meaning as the source clause. Clearly both the strict and the sloppy reading fulfill this requirement, independently from the semantic derivation of the source clause. So a parallelism based theory does without the assumption of spurious ambiguity. Next to the fact that these theories are unified—only the parallelism constraint matters—this is another strong argument in their favor.
However, it can be argued that the assumption of a spurious ambiguity is unavoidable, as soon as we turn attention to subordination constructions. So an adequate account of VPE has to be hybrid between syntax/semantics and pragmatic to some degree.
Consider a comparative construction like

(27) John revised his paper faster than Bill did

The syntactic structure of this sentence, using traditional category labels, is given in figure 15.

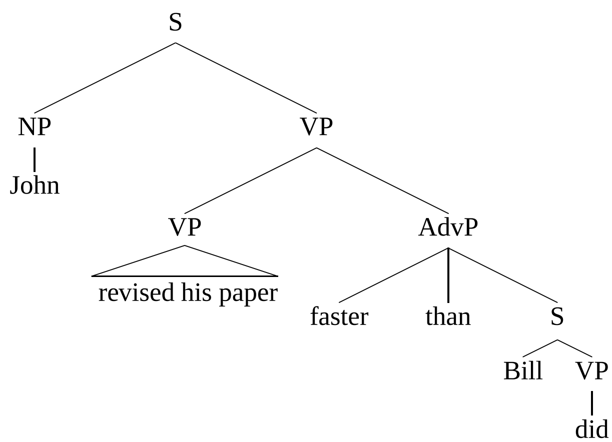

Fig. 15:

The fact that the comparative clause *faster than Bill did* cannot be attached to the matrix S node can be seen from the fact that it is impossible to give the comparative operator scope over the matrix subject. This is illustrated in the following example.

(28) Every math student revised his paper faster than every physics student

The only possible reading is the one where each math student is faster than each single physics student. If the matrix subject were in the scope of *faster*, we expect a reading where the only fastest math student must be faster than the fastest physics student to make this sentence true. Such a reading does not exist.
Given this, it is impossible to establish parallelism between source and target **clause** in (27), since the target clause is included in the source clause. So if parallelism plays

a role here, it can only be a parallelism between VPs, not between clauses. But this means that the meaning of source VP and target VP must be **identical**; the subjects are excluded from parallelism. The target VP is ambiguous between a strict and a sloppy reading, thus there must be a spurious ambiguity in the source VP.
Two conclusions have to be drawn from this. First, the syntax/semantics interface has to supply the option for a pronoun to be bound "sloppily" to the subject argument place of a superordinate verb before the overt subject is supplied. That much spurious ambiguity is inevitable. Second, since the parallelism constraint in whatever shape is unable to say anything about constructions like (27), but the space of possible interpretations there is neither totally free nor totally restricted there, we need a non-trivial theory of VP ellipsis beyond parallelism. The hybrid nature of VPE seems to be a *factum brutum*.
Now let us see what the present theory has to say about the sloppy reading of examples like (27). Reproducing the phrase structure given above in Categorial terms, we assume the lexical assignment

(29) FASTER : $(N \setminus S) \setminus (N \setminus S)/S$

for *faster than*. Binding the pronoun to *John* right away leads to the unproblematic strict reading. But we also correctly predict a sloppy reading. The construction requires that we derive successively two goal type for the source VP *revised his paper* while leaving the pronoun unresolved. First the unresolved VP has to be lifted to the ellipsis type $(S/\mathrm{VP}^1) \setminus S$ to supply an appropriate antecedent for the target clause. But after that, it has to be lowered to the ordinary VP type $N \setminus S$ to serve as argument of the operator *faster than*. There are two derivations for the first part, but they lead to the same result in the second part:

$$\begin{array}{ccccc} \mathrm{VP}^1 & \Rightarrow & (S/\mathrm{VP}^1) \setminus S & \Rightarrow & N \setminus S \\ R & & \left\{ \begin{array}{c} \lambda T.TR \\ \lambda T.T(\lambda x \lambda y.Ryx) \end{array} \right\} & & \lambda x.Rxx \end{array}$$

So for the matrix we derive the expected reading where John revises his own paper. As for the embedded clause, the subject *Bill* has to combine with the "copy" of the lifted VP to form a sentence. Here again both solutions for the lifted type lead to the same result:

$$\begin{array}{cccc} N & (S/\mathrm{VP}^1) \setminus S & \Rightarrow & S \\ \textsc{b} & \left\{ \begin{array}{c} \lambda T.TR \\ \lambda T.T(\lambda x \lambda y.Ryx) \end{array} \right\} & & R\textsc{bb} \end{array}$$

So we correctly predict there to be a sloppy reading in subordinating VPE constructions. Matters become more interesting if we combine the kind of non-subject sloppy scenario with subordination.

(30) John's lawyer defended him better than Bill's lawyer did

It goes without saying that this sentence has a strict reading where John was defended both by his own and by Bill's lawyer. We are interested in the reading where Bill's lawyer defended Bill. What we said about the two goal types of the source VP above applies here as well. If the pronoun *him* is not bound before ellipsis resolution (which leads to a strict reading), it can only be bound to the subject of the matrix verb. This leads to the subject-sloppy reading which is excluded here for binding theoretic reasons. There is no way to derive a genuine non-subject sloppy reading. And this reading in fact does not exist. In other words, we correctly predict sloppy readings in subordination construction to be strictly limited to subjects. Here the predictions derived with the "lifted" entry for the auxiliary are no different from the much simpler theory from subsection 3.2.

6 Summary

This paper proposed the logic $\mathbf{L}_|$ as a type logical reconstruction of Pauline Jacobson's treatment of anaphora in Categorial Grammar. It was shown that this logic preserves desirable meta-logical properties of the Lambek calculus like decidability, finite reading property and strong normalization. Furthermore we demonstrated that in combination with Moortgat's Type Logical approach to quantification, this system covers a considerable range of phenomena concerning the interaction of pronoun binding, quantification and VPE. In the second part of the paper we focused on some intricacies of VPE. We showed that a polymorphic treatment of this kind of ellipsis leads to a theory that correctly accounts for the asymmetry between coordinative and subordinative VPE constructions, the former ones being far more liberal in admitting sloppy readings than the latter.

Proof of theorem 3, Natural Deduction part For the notion of β-normalization in the associative Lambek calculus $\mathbf{L}$ see [Carpenter, 1998] and the references cited therein. $\mathbf{L}_|$ extends the notion of β-reduction by the following two redex configurations:

$$\dfrac{\begin{array}{c}\vdots\\ N:B\end{array}\quad\begin{array}{c}\vdots\\ O:C\end{array}\quad\dfrac{\dfrac{\begin{array}{c}\dfrac{}{x:B}1\\ \vdots\end{array}\quad\begin{array}{c}\dfrac{}{y:p}2\\ \vdots\end{array}\quad\begin{array}{c}\vdots\\ M:A\end{array}}{\langle x,y,M\rangle : B\bullet p\bullet A}\quad\dfrac{\begin{array}{c}\dfrac{}{x:B}3\\ \vdots\end{array}\quad\begin{array}{c}\vdots\\ M:A\end{array}}{\langle x,M\rangle : B\bullet A}}{\lambda x.M : A|B}|I}{\langle N,O,((\lambda x.M)N)\rangle : B\bullet C\bullet A}|E$$

$$\rightsquigarrow$$

$$\dfrac{\begin{array}{c}N:B\\ \vdots\end{array}\quad\begin{array}{c}O:C\\ \vdots\end{array}\quad\begin{array}{c}\vdots\\ M:A\end{array}}{\langle N,O,M[N/x]\rangle : B\bullet C\bullet A}$$

and

$$\dfrac{\begin{array}{c}\vdots\\ N:B\end{array} \qquad \dfrac{\dfrac{\begin{array}{c}\overline{x:B}^{\,1}\\ \vdots\end{array} \quad \begin{array}{c}\overline{y:p}^{\,2}\\ \vdots\end{array} \quad \begin{array}{c}\vdots\\ M:A\end{array}}{\langle x,y,M\rangle : B\bullet p\bullet A} \qquad \dfrac{\begin{array}{c}\overline{x:B}^{\,3}\\ \vdots\end{array} \quad \begin{array}{c}\vdots\\ M:A\end{array}}{\langle x,M\rangle : B\bullet A}}{\lambda x.M : A|B}\,|I}{\langle N,((\lambda x.M)N)\rangle : B\bullet A}\,|E$$

$$\rightsquigarrow$$

$$\dfrac{\begin{array}{c}N:B\\ \vdots\end{array} \quad \begin{array}{c}\vdots\\ M:A\end{array}}{\langle N, M[N/x]\rangle : B\bullet A}$$

It is obvious that these reduction steps, as well as all β-reduction steps in **L**, reduce the number of nodes in the proof tree. Since this number is always a positive integer, there are no infinite sequences of β-reduction steps. $\dashv$

Proof of theorem 3, sequent system part According to the Cut elimination proof given in [Jäger, 2000], every Cut elimination step either (a) eliminates a Cut altogether, (b) replaces a Cut by one or two Cuts of lower degree while keeping the number of distinct atoms in the proof constant (where two atomic formulas count as non-distinct iff they are identical under arbitrary uniform renaming of atoms), or (c) it replaces a Cut by up to three Cuts of possibly higher degree while reducing the number of atoms in the proof by one. (The latter case occurs when a principal Cut on the anaphora slash is eliminated.) Suppose a given Cut application has complexity n. Let us call 2^n the *exponential complexity* of this Cut application, and let us call the sum of the exponential complexities of all Cuts in a given proof Π the *Cut Complexity of* Π. Obviously, in the (a)- and (b)-steps of Cut elimination, this complexity measure is reduced. Let us say that a proof Π_1 is less complex than a proof Π_2 if either Π_1 contains fewer distinct atoms than Π_2 or they contain the same number of atoms and the Cut Complexity of Π_1 is lower than the Cut Complexity of Π_2. Under this complexity measure, every Cut elimination step transforms a given proof into a proof of lower complexity. Furthermore this complexity measure is well-founded (it is isomorphic to ω^2), thus there are no infinite sequences of Cut elimination step. $\dashv$

References

[Carpenter, 1998] Carpenter, B. (1998). *Type-Logical Semantics*. MIT Press.

[Dahl, 1974] Dahl, Ö. (1974). How to open a sentence. In *Logical Grammar Report*, number 12. University of Göteborg.

[Dalrymple et al., 1991] Dalrymple, M., Shieber, S. M., and Pereira, F. C. (1991). Ellipsis and higher-order unification. *Linguistics and Philosophy*, 14(4):399–452.

[Fiengo and May, 1994] Fiengo, R. and May, R. (1994). *Indices and Identity*. The MIT Press.

[Gawron and Peters, 1990] Gawron, J. M. and Peters, S. (1990). *Anaphora and Quantification in Situation Semantics*. CSLI, Stanford.

[Hardt, 1993] Hardt, D. (1993). *Verb Phrase Ellipsis: Form, Meaning, and Processing*. PhD thesis, University of Pennsylvania.

[Hepple, 1990] Hepple, M. (1990). *The Grammar and Processing of Order and Dependency: A Categorial Approach*. PhD thesis, University of Edinburgh.

[Hirschbühler, 1982] Hirschbühler, P. (1982). VP-deletion and Across-the-Board quantifier scope. In Pustejovsky, J. and Sells, P., editors, *Proceedings of NELS 12*, pages 132–139. GLSA, Amherst.

[Jacobson, 1999] Jacobson, P. (1999). Towards a variable-free semantics. *Linguistics and Philosophy*, 22(2):117–184.

[Jäger, 1999] Jäger, G. (1999). Deconstructing Jacobson's **Z**. In Dekker, P., editor, *Proceedings of the 12th Amsterdam Colloquium*, pages 133–138. University of Amsterdam.

[Jäger, 2000] Jäger, G. (2000). Anaphora and quantification in categorial grammar. In Moortgat, M., editor, *Proceedings of LACL 1998*, Springer Lecture Notes in Computer Science. Springer. to appear.

[Kempson and Cormack, 1983] Kempson, R. M. and Cormack, A. (1983). Type lifting rules and VP anaphora. In Barlow, M. T., Flickinger, D. P., and Wescoat, T., editors, *Proceedings of WCCFL 2*, pages 140–152. .

[Lambek, 1958] Lambek, J. (1958). The mathematics of sentence structure. *American Mathematical Monthly*, 65:154–170.

[Moortgat, 1996] Moortgat, M. (1996). Generalized quantification and discontinuous type constructors. In Sijtsma, W. and von Horck, A., editors, *Discontinuous Constituency*. De Gruyter, Berlin.

[Morrill et al., 1990] Morrill, G., Leslie, N., Hepple, M., and Barry, G. (1990). Categorial deduction and structural operations. In Barry, G. and Morrill, G., editors, *Studies in Categorial Grammar*, volume 5 of *Edinburgh Working Papers in Cognitive Science*, pages 1–21. University of Edinburgh.

[Pereira, 1990] Pereira, F. C. N. (1990). Categorial semantics and scoping. *Computational Linguistics*, 16(1):1–10.

[Reinhart, 1983] Reinhart, T. (1983). *Anaphora and Semantic Interpretation*. Croom Helm.

[Rooth, 1992] Rooth, M. (1992). Ellipsis redundancy and reduction redundancy. In Berman, S. and Hestvik, A., editors, *Proceedings of the Stuttgart Ellipsis Workshop*, Arbeitspapiere des SFB 340 "Sprachtheoretische Grundlagen für die Computerlinguistik", Nr. 29. IBM Heidelberg.

[Sag, 1976] Sag, I. A. (1976). *Deletion and Logical Form*. PhD thesis, MIT.

[Shieber et al., 1996] Shieber, S. M., Pereira, F. C., and Dalrymple, M. (1996). Interaction of scope and ellipsis. *Linguistics and Philosophy*, 19(5):527–552.

[Szabolcsi, 1989] Szabolcsi, A. (1989). Bound variables in syntax (are there any?). In Bartsch, R., van Benthem, J., and van Emde Boas, P., editors, *Semantics and Contextual Expressions*, pages 295–318. Foris.

[Wescoat, 1989] Wescoat, M. (1989). Sloppy readings with embedded antecedents. Ms., Stanford University.