Categorical formalisation of relativisation: pied piping, islands, and extaction sites

Glyn Morrill

Report LSI–92–23–R

Categorial Formalisation of Relativisation:
Pied Piping, Islands, and Extraction Sites[*]

Glyn Morrill

September 27th 1992

Departament de Llenguatges i Sistemes Informàtics
Universitat Politècnica de Catalunya
Edifici F I B
Pau Gargallo, 5
08028 Barcelona

This work formulates a grammar of English relativisation on the monostratal rule-to-rule design. Specification of a language model under this design is usually the interactive result of fixing a lexical categorisation and a notion of syntactic consequence (defined by specifying rules of formation). The language model is the closure of the lexicon under syntactic consequence. Our model however is specified in the formalism of *categorial logic*, by which we refer to enrichments of Lambek categorial grammar (Lambek 1958, 1961) with additional category-constructors. Lambek calculus proffers a system of categorisation in which the notion of syntactic consequence is fixed by defining an interpretation of category formulas. Thus the notion of syntactic consequence is defined *logically* (i.e. in virtue of interpretation), and rules of syntax just serve to *calculate*, not to *define*.

Organised as such Lambek calculus, and categorial logic, is the last word in lexicalism. Consequence is fixed model-theoretically and rules of formation are only used in proof-theoretic calculi (of which there can be many kinds) of the formalism.

For substantive background to this approach see Moortgat (1988b), van Benthem (1991) and Morrill (1992). The aim of the present paper is to demonstrate application of the framework developed more theoretically in these references to the formal specification of a 'significant' fragment. This is after all the point of a grammar

[*]This work was supported by a scholarship for *estancies temporales de cientificos y tecnologos extranjeros* from the *Ministerio de Educacion y Ciencia*, Madrid.

formalism; we hope to show that the approach cashes out as a practical, high level framework.

Relativisation provides an effective forum for this demonstration. As a long-distance construction, it exemplifies the basic challenge set to monostratal grammar of establishing non-local syntactic/semantic dependencies.

(1) a. a man who_i John likes e_i
b. a man who_i Mary believes John likes e_i
c. a man who_i Bill thinks Mary believes John likes e_i

Furthermore, it involves a significant number of interesting and well-known linguistic phenomena, including islands, extraction sites, and pied piping. The last of these is regarded as a critical problem for categorial grammar. In a little more detail, the characteristics of relativisation we are going to deal with are as follows.

Firstly, the relative pronoun *whom*, which might be thought of as being accusative, can bind a downstairs nominative position even though it can't bind an upstairs one:

(2) a. *a man whom_i e_i walks
b. a man whom_i John believes e_i walks

Thus we need to distinguish the positions that *whom* can and cannot bind, and this can't be done simply by a featural encoding of case.

Secondly, (some) relative pronouns can pied pipe, i.e. optionally draw material along to their fronted position:

(3) a. a man whom_i John votes for e_i
b. a man $[\text{for whom}]_i$ John votes e_i

(4) a. a man whom_i John knows the brother of e_i
b. a man $[\text{the brother of whom}]_i$ John knows e_i

The synonymy and optionality need to be dealt with, and the treatment of pied piping must also capture its long-distance character:

(5) a thesis [the height of the lettering on the cover of which]$_i$ e_i is prescribed by university regulations

Thirdly, certain domains are islands to relativisation; for example, sentential subjects and coordinate structures:

(6) a. *a man who$_i$ [that Mary likes e_i] is unfortunate
b. *a man who$_i$ [John likes e_i and Mary dislikes Bill]

Fourthly, not all positions are licensed for extraction:

(7) a. a tunnel which$_i$ John emerged out of e_i
b. *a tunnel [of which]$_i$ John emerged out e_i
c. a tunnel [out of which]$_i$ John emerged e_i

That (7a) is grammatical while (7b) is not suggests that we do not try to attribute the ungrammaticality of the latter to an island effect: for then the former should be ungrammatical also. We want to be able to represent positions as unlicensed for extraction without implying that they are islands.

Fifthly, some extraction sites cannot be canonically occupied (Kayne 1984):

(8) a. a man that John assures Mary to be reliable
b. *John assures Mary Bill to be reliable

We shall account for such data by extending the system of categorial classification with category-constructors increasing its expressivity, and we shall attempt to do this in a *clean* way. In writing grammars, as in writing computer programmes, it is not so easy to define good practice, but we suggest the following rules of thumb.

(9) • Lexical ambiguity is not used as an alternative to capturing generalisations;
• Analyses are encoded by locally refining the categorisation of participating elements, not by globally increasing the complexity of lexical categories;
• The desired properties are achieved by interpretation of the formalism, not by ad-hoc stipulations on the calculi;

The organisation of the paper is as follows. After this introduction, §1 treats pied piping. Constraints are discussed in §2; these are characterised in §3 (islands) and §4 (extraction sites). In §5 the economy of the grammar that has been developed is improved somewhat. The overall result is summarised in Appendix A. Appendix B indicates how the bracket operators of §3 constitute a completely general tool for the specification of contìnous and hierarchically organised domains, over and above the uses to which it is put in the main part of the paper.

Let us spell out the essential features of the approach taken by Lambek (1958). Assuming a set of atomic category formulas,[1] (say N for names, S for sentences), the category formulas classifying language objects are given by:

(10) If A is an atomic category formula,
A is a category formula;
If A and B are category formulas,
A\B, B/A and A·B are category formulas;

We may read the directional divisions / and \ as 'over' and 'under' respectively, and the product · as 'times'. The way we proceed is to define a notion of interpretation of the category formulas as sets of language objects. We call a mapping of category formulas to sets of language objects an 'inhabitation'; we call an inhabitation which respects the constraints of interpretation a language model.[2] A lexicon is an inhabitation, and it defines the language model which is its smallest extension satisfying the notion of interpretation.

To begin with language objects will be objects consisting of just a single dimension, of strings, though in general they will be multi-dimensional (with components for prosodics, semantics, ...); in the case of a two-dimensional symbol-

---

[1]What we call category formulas here are often referred to in linguistic contexts as just categories. In mathematical contexts the term 'type' is preferred, to avoid conflict with the terminology of category theory. We shall speak explicitly of 'formulas' in order to preempt this conflict, preserve the term types for semantic types, distinguish the symbolic objects (category formulas) from what they symbolise (categories, in the linguisic sense), and coincide with the usual terminology of logic.

[2]The notion of interpretation fixes possible valuations of category formulas as sets, in a way similar to that in which, say, the interpretation of classical propositional logic in the Boolean algebra formed by taking a power set fixes possible valuations of the formulas as sets. In all cases of model-theoretic semantics a notion of model-structure is specified (usually a set with relations and operations satisfying certain properties), along with constraints specifying possible valuations of formulas in relation to such structures (usually an inductive definition keyed on the construction of formulas). A model consists of a concrete model structure and a particular valuation. In our case a language model is a model in this sense, so that we use the term in the technical sense, as well as in the sense of a mathematical model of a linguistic empirical domain.

meaning interpretation we will speak of language objects as signs. Assume a set V (vocabulary items); this defines V^+, the closure under string concatenation $\oplus$ of the strings consisting of a vocabulary item. By V^* we denote the union of the set V^+ of non-empty strings over V with $\{e\}$ where e is the empty string. Then the structure $(V^*, \oplus, e)$ is an algebra such that $s_1\oplus(s_2\oplus s_3) = (s_1\oplus s_2)\oplus s_3$, i.e. $\oplus$ is associative, and $s\oplus e = e\oplus s = s$, i.e. e is the identity element for $\oplus$.

Each category formula A receives an interpretation as a set D(A) of language objects, for the time being subsets of the set V^+ of non-empty strings. In particular, where d is a mapping from atomic category formulas to such sets of strings:

(11)
$$\begin{aligned}
D(A) &= d(A) \text{ for atomic } A;\\
D(B/A) &= \{s \in V^+ \mid \forall s' \in D(A),\ s\oplus s' \in D(B)\};\\
D(A\backslash B) &= \{s \in V^+ \mid \forall s' \in D(A),\ s'\oplus s \in D(B)\};\\
D(A\cdot B) &= \{s \in V^+ \mid \exists s_1 \in D(A), \exists s_2 \in D(B), s = s_1\oplus s_2\};
\end{aligned}$$

As a result of such interpretation there is the following semantic notion of consequence: $\models A\Rightarrow B$ if and only if D(A) is a subset of D(B) in all models. The following, for instance, are valid:

(12)
$$\begin{aligned}
&A\cdot(A\backslash B) \Rightarrow B\\
&A \Rightarrow (A\cdot C)/C\\
&A \Rightarrow B/(A\backslash B)\\
&A\cdot(B\cdot C) \Leftrightarrow (A\cdot B)\cdot C
\end{aligned}$$

Lexical declaration specifies an inhabitation of categories by vocabulary items, and this defines a language model which is the closure of the lexical inhabitation under the notion of consequence. Hence for example, lexical assignments as in (13) induce assignments as in (14) (amongst others).

(13) John := N
walks := N\S

(14) John walks : S
John : (N·(N\S))/(N\S)
John : S/(N\S)

Lambek (1958) originally provided proof theory for this logic in the shape of the following 'categorical' formulation of statements $A \Rightarrow B$.

(15)

$$A \Rightarrow A \qquad \frac{A \Rightarrow B \quad B \Rightarrow C}{A \Rightarrow C}$$

$$\frac{A \Rightarrow C/B}{A \cdot B \Rightarrow C} \qquad \frac{B \Rightarrow A\backslash C}{A \cdot B \Rightarrow C}$$

$$\frac{A \cdot B \Rightarrow C}{A \Rightarrow C/B} \qquad \frac{A \cdot B \Rightarrow C}{B \Rightarrow A\backslash C}$$

He then showed that by extending the form of statements to $A_1, \ldots, A_n \Rightarrow A_0$, $n \geq 1$, a sequent presentation generating the same theorems is as follows.

(16)

$$A \Rightarrow A \qquad \frac{\Delta \Rightarrow B \quad \Gamma_1, B, \Gamma_2 \Rightarrow A}{\Gamma_1, \Delta, \Gamma_2 \Rightarrow A}\text{Cut}$$

$$\frac{\Delta \Rightarrow B \quad \Gamma_1, C, \Gamma_2 \Rightarrow A}{\Gamma_1, C/B, \Delta, \Gamma_2 \Rightarrow A}/L \qquad \frac{\Gamma, B \Rightarrow A}{\Gamma \Rightarrow A/B}/R$$

$$\frac{\Delta \Rightarrow B \quad \Gamma_1, C, \Gamma_2 \Rightarrow A}{\Gamma_1, \Delta, B\backslash C, \Gamma_2 \Rightarrow A}\backslash L \qquad \frac{B, \Gamma \Rightarrow A}{\Gamma \Rightarrow B\backslash A}\backslash R$$

$$\frac{\Gamma_1, A, B, \Gamma_2 \Rightarrow C}{\Gamma_1, A \cdot B, \Gamma_2 \Rightarrow C}\cdot L \qquad \frac{\Gamma_1 \Rightarrow A \quad \Gamma_2 \Rightarrow B}{\Gamma_1, \Gamma_2 \Rightarrow A \cdot B}\cdot R$$

A sequent $A_1, \ldots, A_n \Rightarrow A_0$ states that concatenating in order any strings in $A_1, \ldots, A_n$ yields a string in A_0, or equivalently, that the set-wise concatenation of $A_1, \ldots, A_n$ is contained in A_0. Apart from the identity axiom and the Cut rule, we find for each connective a rule of use and a rule of proof, these being notated as L(eft) rules and R(ight) rules respectively in the sequent format. Note that in the rules of proof /R and \R a sole antecedent cannot be condionalised, since we know from the interpretation clauses that division categories do not include the empty string. Finally, Lambek showed that the sequent calculus enjoys Cut-elimination, i.e. that for every proof there is a Cut-free counterpart, so that fully representative proof-search can be carried out on the basis of the Cut-free presentation. Since each refinement step decrements the total number of connective occurrences by one, this provides a decision procedure for the calculus (cf. Moortgat 1988b). Buszkowski (1986) provides a summary of completeness results; in fact few such results have currently been secured for categorial calculi. The $\{/, \backslash\}$ fragment is complete for the interpretation given above, but even for just $\{/, \backslash, \cdot\}$ completeness is still unproved for this interpretation.[3, 4]

Let us emphasise at this point that with equivalent categorical and sequent presentations, and other presentations being possible (quasi Hilbert style combinatory calculi: Zielonka 1981; Natural deduction: Morrill, Leslie, Hepple and Barry 1990, and below; Labelled deduction: Moortgat 1992) it is to an interpretation (model theory) of category formulas that we must look for the meaning and notion of consequence of which the various calculi are meant to be embodiments.

This approach is extendable to a modelling of language objects as pairs including a semantic component. The so-called Curry-Howard correspondence relates intuitionistic logic to typed lambda calculus by reading an implicational formula $A \rightarrow B$ as the space of functions from the space A to the space B. $A_1, ..., A_n \Rightarrow A_0$ is an intuitionistic

---

[3]In descriptive/empirical work we shall usually demonstrate grammaticality by means of a derivation in a (sound) proof theory, and ungrammaticality by reference to the model theory. Thus *completeness* of calculi is less important here than in e.g. parsing on the basis of a calculus, when failure to find a proof is to be equated with ungrammaticality. Completeness results are better for more general models, e.g. semigroups as opposed to the free semigroups of the main text, but for clarity we keep to the most specific models for the linguistic application in hand.

[4]We are dealing with Lambek's (1958) construal of categorial grammar as an *associative* system. His (1961) looks at the non-associative case in which the algebra of interpretation comprises binary trees, or equivalently binary bracketed strings $V^{[2]}$ over a vocabulary V. The interpretation clauses are the same as in (11) except with $V^{[2]}$ replacing V^+. The left hand sides of sequents (which we refer to as configurations) become binary-bracketed, and we write e.g. $\Gamma(\Delta)$ to indicate a configuration Γ with a distinguished (well-bracketed) subconfiguration occurrence Δ. Configuations are to be read as indicating the obvious setwise concatenation and bracketing. The sequent calculus is as follows.

(i)

$$A \Rightarrow A \qquad\qquad \frac{\Delta \Rightarrow B \quad \Gamma_1, B, \Gamma_2 \Rightarrow A}{\Gamma_1, \Delta, \Gamma_2 \Rightarrow A}\,\mathrm{Cut}$$

$$\frac{\Delta \Rightarrow B \quad \Gamma_1, C, \Gamma_2 \Rightarrow A}{\Gamma_1, C/B, \Delta, \Gamma_2 \Rightarrow A}\,/L \qquad\qquad \frac{\Gamma, B \Rightarrow A}{\Gamma \Rightarrow A/B}\,/R$$

$$\frac{\Delta \Rightarrow B \quad \Gamma_1, C, \Gamma_2 \Rightarrow A}{\Gamma_1, \Delta, B\backslash C, \Gamma_2 \Rightarrow A}\,\backslash L \qquad\qquad \frac{B, \Gamma \Rightarrow A}{\Gamma \Rightarrow B\backslash A}\,\backslash R$$

$$\frac{\Gamma_1, A, B, \Gamma_2 \Rightarrow C}{\Gamma_1, A{\cdot}B, \Gamma_2 \Rightarrow C}\,{\cdot}L \qquad\qquad \frac{\Gamma_1 \Rightarrow A \quad \Gamma_2 \Rightarrow B}{\Gamma_1, \Gamma_2 \Rightarrow A{\cdot}B}\,{\cdot}R$$

The non-associative calculus preserves constituent structure and does not of itself enable a treatment of non-local dependency.

theorem if and only if there is a lambda term of type A_0 with free variables of types A_1, ..., A_n. Since consequence in Lambek calculus entails that in intuitionistic logic, the type-functional interpretation can be inherited by the former. First, each atomic category formula A is associated with a semantic domain t_A, e.g.

(17) $t_N = E$ for some set E
$t_S = \{0, 1\}$ (truth values)
$t_{CN} = E \rightarrow \{0, 1\}$ (functions from E into {0, 1}, i.e. subsets of E)

Then t_A is extended to all category formulas A. In particular compound division and product categories are associated with spaces created by function-formation and Cartesian product:

(18) $t_{A\backslash B} = t_A \rightarrow t_B$
$t_{B/A} = t_A \rightarrow t_B$
$t_{A \cdot B} = t_A \times t_B$

Each category formula A will be interpreted as a subset D(A) of $V^{+} \times t_A$. Where d maps atomic category formulas to such subsets,

(19)
$$D(A) = d(A) \text{ for atomic } A;$$
$$D(B/A) = \{\langle s, m\rangle \mid \forall \langle s', m'\rangle \in D(A), \langle s \oplus s', m(m')\rangle \in D(B)\};$$
$$D(A\backslash B) = \{\langle s, m\rangle \mid \forall \langle s', m'\rangle \in D(A), \langle s' \oplus s, m(m')\rangle \in D(B)\};$$
$$D(A \cdot B) = \{\langle s, m\rangle \mid \exists \langle s_1, m_1\rangle \in D(A), \exists \langle s_2, m_2\rangle \in D(B), s = s_1 \oplus s_2, m = \langle m_1, m_2\rangle\};$$

In natural deduction format, inference for the product-free fragment with semantics is representable as follows. Deductions consist of annotated ordered tree structures generated from elementary deductions consisting of single category formulas by rules of use and proof labelled E(limination) and I(ntroduction) respectively. A distinct semantic variable is associated with each of the (ordered) assumptions, and semantic terms are built up step-wise towards the root where a lambda term with free variables the variables of undischarged assumptions represents the derivational semantic operation that is to be applied to lexical semantics:[5]

[5]Note again that the rules of proof (I rules in natural deduction) which discharge assumptions must be subject to the condition that A is not the sole assumption of the indicated derivation, for then we would invalidly derive assignment of the empty string to a category.

(20)

$$\dfrac{\begin{matrix}\vdots & \vdots \\ \chi\text{: } B/A & \phi\text{: } A\end{matrix}}{(\chi\,\phi)\text{: } B}/E \qquad \dfrac{\begin{matrix}\cdots\ \overline{x\text{: } A}^{i} \\ \vdots \\ \psi\text{: } B\end{matrix}}{(\lambda x\psi)\text{: } B/A}/I^{i}$$

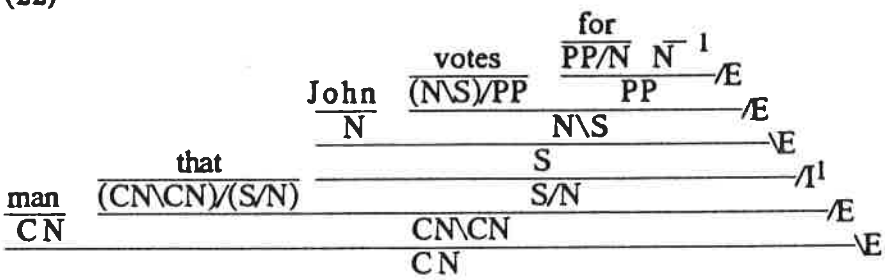
$$\dfrac{\begin{matrix}\vdots & \vdots \\ \phi\text{: } A & \chi\text{: } A\backslash B\end{matrix}}{(\chi\,\phi)\text{: } B}\backslash E \qquad \dfrac{\begin{matrix}\overline{x\text{: } A}^{i}\ \cdots \\ \vdots \\ \psi\text{: } B\end{matrix}}{(\lambda x\psi)\text{: } A\backslash B}/I^{i}$$

Let us illustrate with the beginnings of an account of object relativisation, this being based on Steedman's (1985) assignment to a relative pronoun of a higher-order type (CN\CN)/(S/N). Assume a lexicon as follows.

(21)
for	- **for**
	:= PP/N
John	- **j**
	:= N
likes	- **like**
	:= (N\S)/N
man	- **man**
	:= CN
that	- $(\lambda x(\lambda y(\lambda z[(y\ z) \wedge (x\ z)])))$
	:= (CN\CN)/(S/N)
the	- $(\lambda x[\iota y(x\ y)])$
	:= N/CN
thinks	- **think**
	:= (N\S)/S
votes	- **vote**
	:= (N\S)/PP
woman	- **woman**
	:= CN

Then an example of relativisation is derived as follows.

(22)



The derivational semantics is read directly off (22), according to (20), as being:

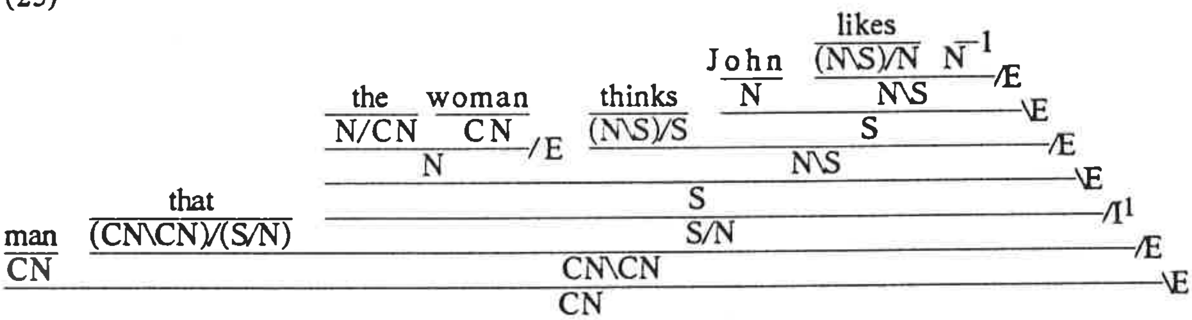
(23) $((x_{that}\,(\lambda x_1((x_{votes}\,(x_{for}\,x_1))\,x_{John})))\,x_{man})$

Substitution of the lexical semantics given in (21) followed by simplification yields:

(24) $(\lambda x[(\mathbf{man}\ x)\wedge((\mathbf{vote}\ (\mathbf{for}\ x))\ \mathbf{John})])$

A second example is given in (25).

(25)



The derivational semantics for *man that the woman thinks John likes* is (26).

(26) $((x_{that}\,(\lambda x_1((x_{thinks}\,((x_{likes}\,x_1)\,x_{John}))\,(x_{the}\,x_{woman})))))\,x_{man})$

Substitution of the lexical semantics followed by simplification yields (27), where an iota operator from the lexical semantics of the determiner is used to represent the definite description.

(27) $(\lambda x[(\mathbf{man}\ x)\wedge((\mathbf{think}\ ((\mathbf{like}\ x)\ \mathbf{j}))\ [\iota y(\mathbf{woman}\ y)])])$

This approach provides the basis of an account of relativisation and its unboundedness but, for example, does not generate medial object extractions such as that in (28), because the argument category S/N of the relative pronoun stands for sentences missing pronominals *at their right periphery*.

(28) the man that$_i$ John meets e_i today

Evidently, associative Lambek categorial grammar of itself both undergenerates, in that it doesn't enable medial extraction, and overgenerates, in that it offers no potential to capture constraints. Accordingly, the tradition of Combinatory Categorial Grammar (CCG; Steedman 1987) seeks notions of syntactic consequence which are both stronger (e.g. in allowing 'mixed composition': A/B, A\C ⇒ C/B, in the Lambek notation) and weaker (e.g. in having only a handful of rule schema). In general, adding such rules as mixed composition to Lambek calculus precipitates permutation closure (see Moortgat 1988a). By keeping to a small set of rule schema, CCG avoids such collapse. But then the theory of syntax is not logical, in the sense of being the reflection of an interpretation of category formulas but, as in e.g. HPSG, DCG and CF-PSG, a deductive system receiving definition in terms of non-logical axioms and rules. Within the logical school, an alternative tradition has arisen which meets the data not by adjusting the calculus, but by increasing the expressivity of categorisation by means of new category-constructors. We consider this first in relation to medial extraction.

The problem is to enable relative pronouns to bind medial as well as peripheral gaps without adding non-logical rules such as mixed composition. One approach arises from the fact that Lambek calculus can be located in the field of substructural logics: logics obtained by dropping structural rules (permutation, contraction, etc.) from intuitionistic logic, e.g. relevance logic, linear logic, associative Lambek calculus, and non-associative Lambek calculus. (See e.g. van Benthem 1989a, Moortgat 1988b, Morrill, Leslie, Hepple and Barry 1990). A notion arising in this field is that of structural modalities (or: exponentials), connectives that license a structural rule on formulas bearing them, though the structural rule is not generally valid (see Došen 1990a). This scheme originated with the modalities ! ('of course') and ? (why not?) of linear logic. Drawing on this, Morrill, *et al.* (1990) and Barry *et al.* (1991) propose permutation structural modalities for medial relativisation. An approach to the interpretation is offered in Morrill (1992, Chaper VI), and pursued in Venema (1992) but logic of structural modalities in general is not yet a resolved issue.

Here we choose a simpler approach, proceeding with a binary operator $\uparrow$ proposed by Moortgat (1988b):

(29) If A and B are category formulas,
$A\uparrow B$ is a category formula;

(30)
$$t_{A\uparrow B} = t_B \rightarrow t_A;$$
$$D(A\uparrow B) = \{\langle s, m\rangle \mid \exists s_1 s_2 \in V^*, s = s_1 \oplus s_2, \forall \langle s', m'\rangle \in D(B), \langle s_1 \oplus s' \oplus s_2, m(m')\rangle \in D(A)\};$$

In sequent format, the rule of proof is thus:

(31)
$$\frac{\Gamma_1, A, \Gamma_2 \Rightarrow B}{\Gamma_1, \Gamma_2 \Rightarrow B\uparrow A}\uparrow R$$

The rule of proof in natural deduction format is as follows.

(32)
$$\frac{\begin{array}{c}\cdots \overline{y\colon B}^{\,i} \cdots \\ \vdots \\ \phi\colon A\end{array}}{(\lambda y \phi)\colon A\uparrow B}\uparrow I^i$$

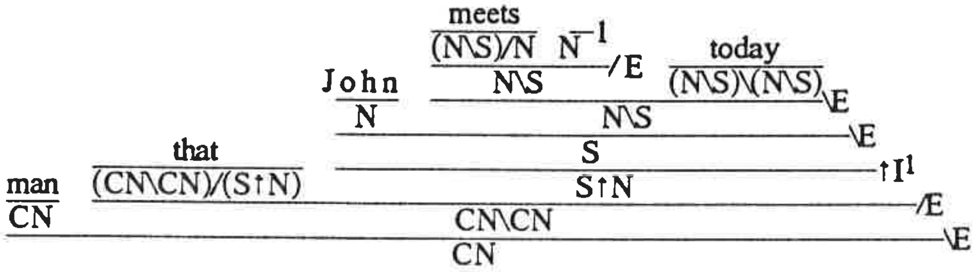
A rule of use is more difficult to give. As in Moortgat (1988b) we limit ourselves to a partial logic of $\{/, \backslash, \cdot, \uparrow\}$, one in which there is only a rule of proof for $\uparrow$, i.e. only a rule of inference for 'positive' occurrences of the connective in statements $A_1, \ldots, A_n \Rightarrow B$. We define the positive occurrences of connectives as follows; if a connective occurrence is not positive, it is negative.

(33)
- A connective occurs posivitely (negatively) in $A_1, \ldots, A_n \Rightarrow B$ if it occurs posivitely (negatively) in B ($A_1, \ldots, A_n$);
- A connective occurs positively (negatively) in A/B, $B\backslash A$, $A\cdot B$, $B\cdot A$, $A\uparrow B$ if it occurs positively (negatively) in A;
- A connective occurs negatively (positively) in A/B, $B\backslash A$, $A\uparrow B$ if it occurs positively (negatively) in B;
- The principle connective of a formula occurs positively in that formula;

With the relative pronoun assignment (34), medial extraction can be derived as shown in (35).

(34) that - (λx(λy(λz[(y z)∧(x z)])))
:= (CN\CN)/(S↑N)

(35)



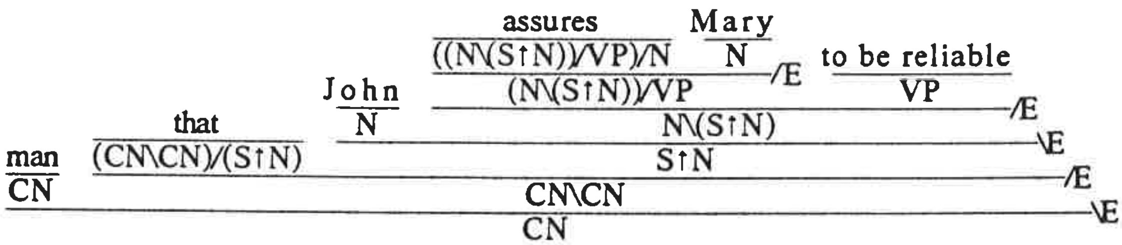
We may see the operator being used here rather like the 'slash' of GPSG (Gazdar, Klein, Pullum and Sag 1985).

Note now the observation by Kayne (1984) that e.g. *assure* in English allows extraction of the subject of its infinitival complement, but does not have a canonical, i.e. non-extracted, form:

(36) a. a man that John assures Mary to be reliable
b. *John assures Mary Bill to be reliable

We seek a weak category for *assures* that can *only* allow extraction; such is provided by ((N\(S↑N))/VP)/N. For on the one hand (36a) can be derived as follows.

(37)



On the other hand however, the sequent (38), corresponding to (36b), is not valid, and nor is any variant where the underlined N corresponding to *Bill* is relocated.

(38) $N, ((N\backslash(S\uparrow N))/VP)/N, N, \underline{N}, VP \Rightarrow S$

The problem is not that a rule of use has not been given for ↑, but that the negative occurrence of ↑ does not yield S model-theoretically from any particular location of the N. Continuing the analogy with PSG, it is as if the argument had been put on the 'slash stack' in the lexicon; that is presumably how (34) would be covered in HPSG (Pollard and Sag 1992).

1 *Pied-Piping*

Historically, pied piping has played a crucial rôle in the promotion of feature percolation and phrase structural approaches (Gazdar *et al.* 1985; Pollard and Sag 1987, 1992) over categorial grammar. Pollard (1988, p. 412) for example regards it as a critical inadequacy:

(39) "Evidently, there is no principled analysis of pied piping in an extended categorial framework like Steedman's without the addition of a feature-passing mechanism for unbounded dependencies."

On the phrase structural view, a relative pronoun introduces information which may percolate up normal constituent structure to endow larger phrases with the relativisation property of occurring fronted and binding a gap of the same category as the entire fronted constituent. Cases in which there is no pied piping are, convincingly, obtained as the special case where the fronted constituent comprises only the relative pronoun.

Note that while a relative pronoun such as *whom* can pied pipe, *that* cannot:

(40) a. a man whom John votes for
b. a man for whom John votes

(41) a. a man that John votes for
b. *a man for that John votes

Note also that (40a), (40b) and (41a) are synonymous. The category $(CN\backslash CN)/(S\uparrow N)$ we are presently using would predict for *that* that it cannot pied pipe, so the data given might be interpreted as indicating that e.g. *whom* should have some additional, pied piping, lexical entry over and above the one shared with *that*. For the particular instance in (40b) we could assume a category (PP/N)\((CN\CN)/(S↑PP)). But in (42) for example the relative pronoun is medial in the fronted constituent and would need to be a functor over both the preceeding and the following material.

(42) a man a brother of whom from Brazil appeared on television

Furthermore it would need to be arranged that *from Brazil* modify *brother*. And it seems clear that any number of such examples, each requiring individual treatment, could be found. Thus, CCG not only falls short of the phrase structure grammar characterisation of no pied piping as null pied piping, but also falls short of rescuing adequacy through lexical ambiguity.

Pied piping belongs to a family of *in situ* binding phenomena (PSG 'FOOT' feature phenomena) such as reflexivisation and quantification, each of which involves an element taking semantic scope at the level of a superordinate phrase in which it occurs. Moortgat (1990) presents a binary category-constructor such that A⇑B signifies elements that occupy an A position, and take semantic scope at an embedding B, e.g. quantifier phrases would be N⇑S. But pied piping also involves a syntactic difference (and that is why it is more obviously a problem for CCG): the category of a phrase in which a pied piping relative pronoun occurs is changed from a B to that of a filler for B. Moortgat (1991) presents a ternary generalisation of his 1990 proposal: informally, Q(A, B, C) means something that occurs in an A position in a B resulting in a C. A⇑B could be defined as Q(A, B, B). We will treat pied piping by means of Moortgat's ternary connective, an application of which he happenned to be unaware.

The operator definitions are as follows:

(43) If A, B and C are category formulas,
Q(A, B, C) is a category formula;

(44) $t_{Q(A, B, C)} = (t_A \rightarrow t_B) \rightarrow t_C$;

(45) $D(Q(A, B, C)) = \{\langle s, m\rangle \mid \forall s_1, s_3 \in V^*, \forall m' \in t_A \rightarrow t_B, [\forall \langle s_2, m_2\rangle \in D(A), \langle s_1 \oplus s_2 \oplus s_3, m'(m_2)\rangle \in D(B)] \rightarrow \langle s_1 \oplus s \oplus s_3, m(m')\rangle \in D(C)\}$

The lexical categorisation of *whom* in (46) states that it occupies an N position within a PP with the result being a (CN\CN)/(S↑PP), i.e. a relative clause filler for category PP.

(46) whom - $(\lambda x(\lambda y(\lambda z(\lambda w[(z\ w) \wedge (y\ (x\ w))]))))$
$:= Q(N, PP, (CN\backslash CN)/(S\uparrow PP))$

We give inference for just the negative occurrences of the new connective:

(47)

$$\frac{\Delta(A) \Rightarrow B \qquad \Gamma(B) \Rightarrow D}{\Gamma(\Delta(Q(A, B, C))) \Rightarrow D}\,QL$$

(48)

$$\begin{array}{c} \vdots \\ \dfrac{\xi\text{: } Q(A, B, C)}{x\text{: } A}\,QE^{i} \\ \vdots \\ \dfrac{\psi\text{: } B}{(\xi\ (\lambda x \psi))\text{: } C}\,{}^{i} \end{array}$$

In the semantics of (41), the first argument will be the meaning of the fronted material abstracted over that of the position occupied by the relative pronoun, and the second argument will be the meaning of the embedded sentence abstracted over the meaning of the PP gap; it is the functional composition of these that defines the predicate that is intersected with the meaning of the modified noun. To see this in action, recall in addition to (46) the assignments in (21). A derivation for (40b) is as follows:

(49)

$$\dfrac{\dfrac{\text{man}}{CN} \qquad \dfrac{\dfrac{\dfrac{\dfrac{\text{for}}{PP/N} \qquad \dfrac{\dfrac{\text{whom}}{Q(N, PP, (CN\backslash CN)/(S\uparrow PP))}}{N}\,QE^{1}}{PP}\,/E}{(CN\backslash CN)/(S\uparrow PP)}\,{}^{1} \qquad \dfrac{\dfrac{\dfrac{\text{John}}{N} \qquad \dfrac{\dfrac{\text{votes}}{(N\backslash S)/PP} \qquad \overline{PP}^{2}}{N\backslash S}\,/E}{S}\,\backslash E}{S\uparrow PP}\,\uparrow I^{2}}{CN\backslash CN}\,/E}{CN}\,\backslash E$$

This delivers as derivational semantics (50), read directly off the natural deduction derivation.

(50) $(((x_{whom}\ (\lambda x_1(x_{for}\ x_1)))\ (\lambda x_2((x_{votes}\ x_2)\ x_{John})))\ x_{man})$

Substitution of the lexical semantics followed by simplification yields (51).

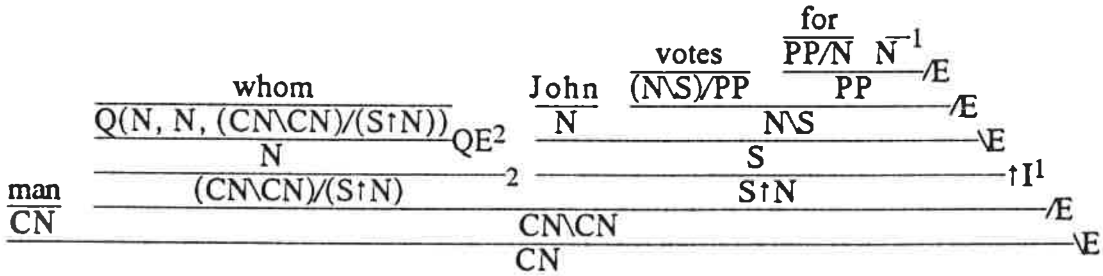
(51) $(\lambda w[(\mathbf{man}\ w) \wedge ((\mathbf{vote}\ (\mathbf{for}\ w))\ \mathbf{j})])$

As required, this is the same semantics (24) as that obtained for (41a) *man that John votes for* under the non pied piping assignment to *that.* It is clear that the category assignment (52) for nominal as opposed to prepositional pied piping will allow e.g. *the man a picture of whom appeared in the newspaper* and also medial cases such as (42).

(52) whom - $(\lambda x(\lambda y(\lambda z(\lambda w[(z\ w) \wedge (y\ (x\ w))]))))$
$:= Q(N, N, (CN\backslash CN)/(S{\uparrow}N))$

What is less obvious is that the assignment (52) also correctly generates non-pied piping cases such as (40a) *the man whom John votes for.*

(53)



The derivational semantics for this is (54).

(54) $(((x_{whom}\ (\lambda x_2 x_2))\ (\lambda x_1((x_{votes}\ (x_{for}\ x_1))\ x_{John})))\ x_{man})$

The one-step inference from $Q(N, N, (CN\backslash CN)/(S{\uparrow}N))$ to $(CN\backslash CN)/(S{\uparrow}N)$ by QE involves applying the lexical semantics of *whom* to the identity function, resulting in the usual non pied piping semantics for a relative pronoun, so that the semantics obtained for (40a) is the same as that for the synonomous (40b) and (41a).

Constraints

The approach as it stands predicts that the body of a relative clause can be any sentence with a gap at a name position. But the phenomenon of relativisation is subject to further constraints. Thus according to the Coordinate Structure Constraint (CSC) of Ross (1967), coordinate structures are islands.[6, 7]

(55) a. *the man $that_i$ [e_i walks and John sings]

b. *the man $that_i$ [Mary walks and e_i sings]

c. *the man $that_i$ [Mary likes e_i and John dislikes Bill]

d. *the man $that_i$ [Mary likes Fred and John dislikes e_i]

Ross (1967) also proposed a Sentential Subject Constraint (SSC) to block such cases as (56).

(56) a. *a topic $which_i$ [that John talks about e_i] annoys Mary

b. *the pleasures $which_i$ [for you to give up e_i] would be a pity

Extraction from the 'subject' of a possessive as in (57) is also bad.

(57) a. *the man who_i John read [the brother of e_i's] book

b. *the man who_i I read [e_i's] book

And, the last of what we consider 'strong' island effects, according to the Complex Noun Phrase Constraint of Ross (1967), relative clauses themselves are islands:

(58) *the man $that_i$ John met a woman [$that_j$ $e_{i/j}$ loves $e_{i/j}$]

Complex noun phrases here are those with sentential modifiers, i.e. the constraint deems the following ungrammatical.

---

[6]Across-The-Board (ATB) exceptions such as (i) are predicted under coordination of like categories (in the present case $S\uparrow N$) in a manner essentially equivalent to that of Gazdar (1981).

(ii) the man $that_i$ [Mary likes e_i and John dislikes e_i]

To generate the full range of constituent and non-constituent Boolean coordination in relation to which categorial grammar is so promising (cf. Steedman 1985, Dowty 1988) it is tempting to assume a second-order polymorphic category $\forall A[(A\backslash A)/A]$ for coordinators with semantics *n*-ary generalised coordination for A an *n*-ary relational category. But note for example that this would predict the classically erroneous wide-scope *and* reading for a *man walks and talks* under the instanciation A = $(S/(N\backslash S))\backslash S$. Rather, some restricted second-order polymorphic category $\forall A{:}R(A)[(A\backslash A)/A]$ is required where the precise empirical and technical formulation of R is still very much an open issue.

[7]All grammaticality judgements are ours, and not necessarily those of authors cited.

(59) ?the man whom$_i$ Mary acquired a belief that I dislike e_i

Such cases are less clearly unacceptable. In view of such data as (60), Chomsky (1973) formulated the Subject Condition (SC) which subsumes the SSC, asserting that all subjects are islands.

(60) ?a woman whom$_i$ [a picture of e_i] appeared in the newspaper

And in order to capture such examples as (61) Bach and Horn (1976) go even further, proposing an NP Constraint deeming all noun phrases islands.

(61) a. *the person who$_i$ John destroyed [a book about e_i]
b. *the spy who we saw Mary's picture of

Structurally equivalent examples however are be semi-acceptable or acceptable:

(62) ?the man who$_i$ I lost [a picture of e_i]

(63) a. the programme which$_i$ I missed [the end of e_i]
b. the town which$_i$ I bought [a ticket to e_i]

It seems to us that in fact all the examples since (59) are grammatical in context. Consider for example (61a) where there were various individuals each of whom destroyed a book on a different person. With stress on *John* we have a perfectly grammatical identification of someone. Likewise, consider (61b) in a context where various groups saw just one of several pictures of spies that Mary has. Then with stress on *we*, the example is fine, with continuations such as ... *turned up in Jamaica*.

There are other constraints on extraction domains, positing e.g. factive verb complement islands, *wh*-islands and adverbial islands. The present paper develops a general *formalism* for the specification of labelled bracketed strings, as described immediately below and in appendix B, and this may be used to define all kinds of phonological and syntactic domains. It should be clear how this apparatus would be used to capture various alleged island domains.

But in addition to introducing the formalism, we wish to make a theoretical proposal with respect to the strong island effects we have noted for coordinate structures, relative clauses, sentential subjects and the subjects of possessives. This is that although prosodic phrasing is quite free, it cannot cross-cut these domains, and that it is this prosodic rigidity which underlies their islandhood:

(64) *Prosodic Island Constraint*
Domains which cannot be broken by prosodic phrasing are strong islands.

Consider the following prosodic phrasings:

(65) a. (Bill) (thinks John walks)
b. (Bill thinks) (John walks)
c. (Bill thinks John) (walks)

The sentence allows all divisions into two prosodic constituents. In general many prosodic phrasings are well-formed, but not all are. Coordinate structures, for instance, cannot be interrupted in this way:

(66) a. (thinks) ([John walks and Mary talks])
b. *(thinks [John) (walks and Mary talks])
c. *(thinks [John walks) (and Mary talks])
d. *(thinks [John walks and) (Mary talks])
e. *(thinks [John walks and Mary) (talks])

Likewise, prosodic phrasing cannot cross-cut relative clauses:

(67) a. (a rat) ([that John chased])
b. *(a rat [that) (John chased])
c. *(a rat [that John) (chased])

For sentential subjects we have:

(68) a. ([that John likes physics]) (annoys Mary)
b. *([that John likes) (physics] annoys Mary)
c. *([that John) (likes physics] annoys Mary)
d. *([that) (John likes physics] annoys Mary)

And for the subjects of possessives:

(69) a. (John found) ([the tall man's] book)

b. *(John found [the) (tall man's] book)

c. *(John found [the tall) (man's] book)

d. *(John found [the tall man)'(s] book)

e. (John found [the tall man]'s) (book)

That the Lambek calculus we are using is associative can be taken as indicating that all prosodic phrasings are possible (cf. Moortgat 1988b): that a string is generated can be taken to indicate that each of its equivalence class (under associativity) of bracketings is a well-formed prosodic phrasing. This bypasses problems of prosodic/syntactic mismatch, but constitutes the null hypothesis that all phrasings are possible. We shall arrange generation of (partially) bracketed strings where the brackets constitute prosodic phrasing constraints and, as we have indicated, the relevant domains will also be islands to extraction.

Apart from constraints on extraction domains there are constraints on extraction sites. Thus while e.g. all the object positions of verbs are possible extraction sites in English, subjects positions of complementised embedded sentences are not (the 'fixed subject constraint', '*that*-trace filter' or 'empty subject filter'; Bresnan 1972; Chomsky and Lasnik 1977):

(70) a. *the man who$_i$ John thinks that e_i walks

b. the man who$_i$ John thinks e_i walks

That a position is not a possible extraction site does not necessarily mean that it constitutes an island. Thus the following indicates that while in English the PP complement of a preposition is not extractable, it does not form an island domain:

(71) a. the tunnel out of which$_i$ John emerged e_i

b. *the tunnel of which$_i$ John emerged out e_i

c. the tunnel which$_i$ John emerged out of e_i

In what follows we shall formulate syntactic calculus for these constraints on relativisation.

Islands

For islands we are going to propose category-constructors which directly project bracketing-off of the relevant domains. We propose a calculus of well-bracketed strings (i.e. trees) in which bracketing will mark off islands.[8] Let b be the operation of adding outermost matching brackets to a string. A set V of vocabulary items generates the set $V^{+,[]}$ which is the closure of unit strings on V under $\oplus$ *and* b. By $V^{*,[]}$ we denote the union of $V^{+,[]}$ with the empty string. Category formulas constructed by /, \, ·, ↑ and Q are interpreted as before, but with reference to $V^{+,[]}$ and $V^{*,[]}$ rather than V^* and V^+, i.e. each category A is interpreted as a subset of $V^{+,[]} \times t_A$ and, for example:

(72) $D(A\uparrow B) = \{\langle s, m\rangle \mid \exists s_1 s_2 \in V^{*,[]}, s = s_1 \oplus s_2, \forall \langle s', m'\rangle \in D(B), \langle s_1 \oplus s' \oplus s_2, m(m')\rangle \in D(A)\};$

In addition we have:

(73) If A is a categorial type formula,
$[\,]A$ and $[\,]^{-1}A$ are categorial type formulas;

Where

(74) $t_{[]A} = t_{[]^{-1}A} = t_A;$

And

(75) $D([\,]A) = \{\langle s, m\rangle \mid \exists \langle s', m\rangle \in D(A), s = b(s')\};$
$D([\,]^{-1}A) = \{\langle s, m\rangle \mid \exists \langle s', m\rangle \in D(A), s' = b(s)\};$

Inhabitants of []A are all and only the results of bracketing inhabitants of A. For example, if A is {John walks, [that it rains] is obvious, [John walks and it rains]}, []A is {[John walks], [[that it rains] is obvious], [[John walks and it rains]]}. Inhabitants of A and []A stand in a one-one relation: the cardinality of the sets is the same. The inhabitants of $[\,]^{-1}A$ are all and only the results of debracketing an inhabitant of A, so given A as above, $[\,]^{-1}A$ is {John walks and it rains}. There are only inhabitants of []-

[8]Morrill (1990c) treats islands this way, using a mixed associative/nonassociative system which can project brackets from functors. But it is of lower expressivity then the present system, unable, for example, to project brackets onto arguments, as for the SSC. We indicate in the appendix that the present system is sufficiently powerful to embed the nonassociative calculus.

^{1}A corresponding to those inhabitants of A which are bracketed outermost, so that the cardinality of $[\,]^{-1}A$ is less then or equal to that of A.

Sequents will now contain bracketing [...] and antibracketing $[...]^{-1}$. We refer to the left hand sides of sequents as configurations, and $\Gamma(\Delta)$ represents a configuration Γ with a distinguished well-bracketed subconfiguration occurrence Δ (cf. fn. 4). A configuation represents an operation comprising concatenation, bracketing and unbracketing in a transparent manner, and a sequent $\Gamma \Rightarrow A$ now asserts that applying the relevant set-wise generalisation of this operation to the antecedent types yields a subset of A. Before presenting the rules for the connectives themselves, it will already be seen that the following 'structural rules' are valid.

(76)

$$\frac{\Delta(\Gamma) \Rightarrow A}{\Delta([[\Gamma]]^{-1}) \Rightarrow A}AB \qquad \frac{\Delta(\Gamma) \Rightarrow A}{\Delta([[\Gamma]^{-1}]) \Rightarrow A}BA$$

The validity of the first of these follows from the fact that the result of debracketing the bracketing of a set of strings is identical to, and thus an (improper) subset of, the original set of strings. The validity of the second follows from the fact that the result of bracketing the debracketing of a set of strings is a (generally proper) subset of the original set of strings. The logical rules we propose in sequent format are:

(77)

$$\frac{\Gamma([A]) \Rightarrow B}{\Gamma([\,]A) \Rightarrow B}[\,]L \qquad \frac{\Gamma \Rightarrow A}{[\Gamma] \Rightarrow [\,]A}[\,]R$$

$$\frac{\Gamma([A]^{-1}) \Rightarrow B}{\Gamma([\,]^{-1}A) \Rightarrow B}[\,]^{-1}L \qquad \frac{[\Gamma] \Rightarrow A}{\Gamma \Rightarrow [\,]^{-1}A}[\,]^{-1}R$$

By way of illustration, observe how the calculus derives the three validities $[\,]([\,]^{-1}A) \Rightarrow A$, $[\,]^{-1}([\,]A) \Rightarrow A$, $A \Rightarrow [\,]^{-1}([\,]A)$, and fails to derive the invalidity $A \Rightarrow [\,]([\,]^{-1}A)$:

(78)

a.
$$\dfrac{\dfrac{\dfrac{A \Rightarrow A}{[[\, A\,]^{-1}] \Rightarrow A}\,\text{BA}}{[\,[\,]^{-1}A\,] \Rightarrow A}\,[\,]^{-1}\text{L}}{[\,]([\,]^{-1}A) \Rightarrow A}\,[\,]\text{L}$$

b.
$$\dfrac{\dfrac{\dfrac{A \Rightarrow A}{[[\, A\,]]^{-1} \Rightarrow A}\,\text{AB}}{[\,[\,]A\,]^{-1} \Rightarrow A}\,[\,]\text{L}}{[\,]^{-1}([\,]A) \Rightarrow A}\,[\,]^{-1}\text{L}$$

c.
$$\dfrac{\dfrac{A \Rightarrow A}{[\, A\,] \Rightarrow [\,]A}\,[\,]\text{R}}{A \Rightarrow [\,]^{-1}([\,]A)}\,[\,]^{-1}\text{R}$$

d.
$$\dfrac{}{A \Rightarrow [\,]([\,]^{-1}A)}\,{*}[\,]\text{R}$$

We propose the inference rules (80) in natural deduction format, with rules (79) corresponding to (76).

(79)

$$\dfrac{\begin{matrix}\vdots\\ \phi\!: A\end{matrix}}{[\,[\,\phi\!: A\,]\,]^{-1}} \qquad\qquad \dfrac{\begin{matrix}\vdots\\ \phi\!: A\end{matrix}}{[\,[\,\phi\!: A\,]^{-1}\,]}$$

(80)

$$\dfrac{\begin{matrix}\vdots\\ [\,\phi\!: [\,]A\,]^{-1}\end{matrix}}{\phi\!: A}\,[\,]\text{E} \qquad\qquad \dfrac{\begin{matrix}\vdots\\ [\,\phi\!: A\,]\end{matrix}}{\phi\!: [\,]A}\,[\,]\text{I}$$

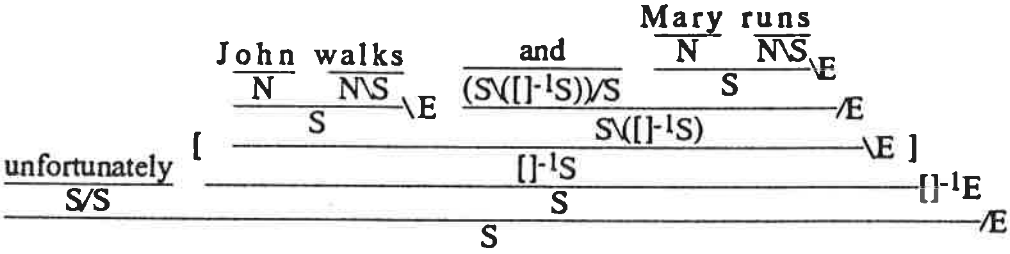
$$\dfrac{\begin{matrix}\vdots\\ [\,\phi\!: [\,]^{-1}A\,]\end{matrix}}{\phi\!: A}\,[\,]^{-1}\text{E} \qquad\qquad \dfrac{\begin{matrix}[^{i} \quad]^{-i}\\ \vdots\\ \phi\!: A\end{matrix}}{\phi\!: [\,]^{-1}A}\,[\,]^{-1}\text{I}^{i}$$

Consider the effect of assigning a sentential coordinator to the category $(S\backslash([\,]^{-1}S))/S$:

(81) and - $(\lambda x(\lambda y[x \wedge y]))$
$:= (S\backslash([\,]^{-1}S))/S$

A sentence such as *unfortunately* [*John walks and Mary runs*] can now only be recognised with the indicated bracketing. With no bracketing, or with the bracketing elsewhere, the crucial inference using $[\,]^{-1}$ cannot take place, and any excess bracketing would also block derivation.

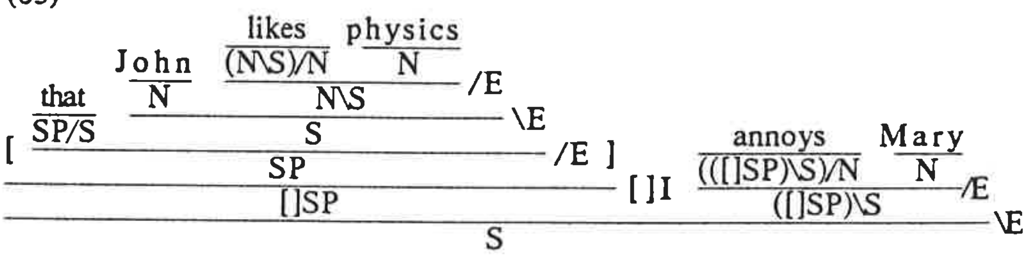
(82)



In the case of *the man who* [*John likes and Mary sings*] then, the bracketing would have to be as shown. But then [*John likes and Mary sings*] does not belong to the relative pronoun argument category S↑N; for this is the category of expressions that can be partitioned into two members of $V^{*,[]}$ which when interpolated by an N yield an S - recall (72) - and [*John likes* and *and Mary sings*] are not members of $V^{*,[]}$, all the members of which are well-bracketed.

In a similar way, assignment of category (([]$^{-1}$(CN\CN))/(S↑N) to a relative pronoun marks off a relative clause itself as an island. Thus coordinate structures and relative clauses are characterised as islands by projecting from []$^{-1}$ on the value of functors. In the case of the Sentential Subject Condition something slightly different is required. For *that John likes physics annoys Mary* we might try to specify the subject island by assigning *that* category ([]$^{-1}$SP)/S. But *that*-clauses are not in general islands. Rather, we want the verb phrase taking the sentential subject to mark the latter as an island. This is what is done by assigning *annoys* to category ((([]SP)\S)/N:

(83)



And again, *(*a topic*) *which*$_i$ [*that John likes* e_i] *annoys Mary* is blocked since [*that John likes*] *annoys Mary* can't be partitioned into well-bracketed strings that give an S when interpolated by an N.

Extraction sites

One approach to constraints in Combinatory Categorial Grammar is to disqualify certain categories (e.g. *NP/NP for the NP Constraint) or to disqualify certain rule schema instances (e.g. A/B, B/C $\Rightarrow$ A/C, provided A/C $\not\Rightarrow$ NP/NP for the NP Constraint). But limitation to a few rule schema already loses the logical foundation, and restricted category formation or rule instanciation only entrenches this deficit.

A logical treatment of locality domains is iniciated in Morrill (1989, 1990b) in terms of a modal categorial calculus. These proposals are oriented in the first place towards a categorial treatment of intensionality, i.e. semantic domains, with syntactic locality constraints being regarded as a potential spin-off. The approach illustrates how, in contrast to the CCG *restriction* of rules of inference for say a {/, \} categorial language, we can *extend* the categorial expressivity by means of additional operators.

(84) If A is a category formula,
$\Box$A is a category formula;
If A is a category formula,
$\Diamond$A is a category formula;

In modal languages interpretation is relativised Kripke-style to points in a set I on which an accessibility relation R is defined. As for various other category-constructors (conjunction/disjunction, universal/existential quantification; see Morrill 1990a, 1992) some of which we see later, there are both semantically potent and non-potent versions.[9] For presentational purposes we avoid semantics at first. A category formula A is now interpreted as a set $D(A)^i$ of strings relative to each point i (equivalently each category formula has a single interpretation as a set of point-string pairs).

(85) $D(\Box A)^i = \{s \mid \forall jRi,\ s \in D(A)^j\}$
$D(\Diamond A)^i = \{s \mid \exists jRi,\ s \in D(A)^j\}$

The interpretation of other connectives is fixed 'point-wise':

(86) $D(B/A)^i = \{s \mid \forall s' \in D(A)^i,\ s{\oplus}s' \in D(B)^i\}$
$D(A\backslash B)^i = \{s \mid \forall s' \in D(A)^i,\ s'{\oplus}s \in D(B)^i\}$

[9]Modal operators are in effect restricted cases of (perhaps higher-order) universal and existential quantification, which are in turn generalisations of conjunction and disjunction. This is why there are family resemblances.

Various modal logics are defined by setting conditions on the accessibility relation R:

(87) iRi — reflexivity

iRj & jRk → iRk — transitivity

iRj → jRi — symmetry

For instance, in the model logic K there are no conditions; in S4 reflexivity and transitivity are imposed; and in S5 symmetry is also required.

(88)

K

↓ reflexivity, transitivity

S4

↓ symmetry

S5

Sequent calculus rules for S4 are as follows, where $\Box\Gamma$ ranges over sequences $\Box A_1$, ..., $\Box A_n$ of $\Box$-ed formulas.

(89)

$$\frac{\Box\Gamma \Rightarrow A}{\Box\Gamma \Rightarrow \Box A}\Box R \qquad \frac{\Gamma_1, A, \Gamma_2 \Rightarrow B}{\Gamma_1, \Box A, \Gamma_2 \Rightarrow B}\Box L$$

$$\frac{\Gamma \Rightarrow A}{\Gamma \Rightarrow \Diamond A}\Diamond R \qquad \frac{\Box\Gamma_1, A, \Box\Gamma_1 \Rightarrow \Diamond B}{\Box\Gamma_1, \Diamond A, \Box\Gamma_1 \Rightarrow \Diamond B}\Diamond R$$

A number of possibilities for capture of constraints are created by modalisation. For example, suppose (for S4 or S5) that in general lexical types are $\Box$-ed outermost and that sentence-embedding elements are functors over $\Box S$. Then assignment of a relative pronoun category $\Box(R/(S/(\Box N)))$ would allow long distance relativisation, but $\Box(R/(S/N))$ only local relativisation. Similarly, a reflexive category $\Box(((N\backslash S)/(\Box N))\backslash(N\backslash S))$ would allow long-distance reflexivisation, but $\Box(((N\backslash S)/N)\backslash(N\backslash S))$ only local reflexivisation (see Morrill 1989, 1990b).

For semantically potent modalities we have type maps $t_{\Box A} = I \rightarrow t_A$, and $t_{\Diamond A} = I \times t_A$. Then, adding the semantic dimension:

(90) $D(\Box A)^i = \{\langle s, m\rangle | \forall jRi, \langle s, m(j)\rangle \in D(A)^j\}$

$D(\Diamond A)^i = \{\langle s, m\rangle | \exists jRi, m = \langle j, m'\rangle, \langle s, m'\rangle \in D(A)^j\}$

For the semantically non-potent case, $t_{\Box A} = t_{\Diamond A} = t_A$, and we have:

(91) $D(\Box A)^i = \{\langle s, m\rangle | \forall jRi, \langle s, m\rangle \in D(A)^j\}$

$D(\Diamond A)^i = \{\langle s, m\rangle | \exists jRi, \langle s, m\rangle \in D(A)^j\}$

For the modal logic S5, interpretation is particularly simple: because the accessibility relation is universal we can ignore it and just quantify over the set of points. For the semantically potent case:

(92) $D(\Box A)^i = \{\langle s, m\rangle | \forall j, \langle s, m(j)\rangle \in D(A)^j\}$

$D(\Diamond A)^i = \{\langle s, m\rangle | \exists j, m = \langle j, m'\rangle, \langle s, m'\rangle \in D(A)^j\}$

And for the semantically non-potent case:

(93) $D(\Box A)^i = \{\langle s, m\rangle | \forall j, \langle s, m\rangle \in D(A)^j\}$

$D(\Diamond A)^i = \{\langle s, m\rangle | \exists j, \langle s, m\rangle \in D(A)^j\}$

More than one modal operator of a given kind may be used. Morrill (1992) uses bimodal □ and ■ in a treatment of intensionality in terms of times and worlds. Hepple (1990) even attempts treatment of syntactic domains by the polymodal boxes with an inclusion ordering. Aside from technical complications of polymodality, sensitisation by universal modality has the drawback that a *global* modalisation is required: in general *all* category formulas need to be beefed up by one or more modal operators. This is a failure of economy, unless the modality has an independent motivation, e.g. as encoding a dimension of intensional semantics, as in Morrill (1990b). Our own present view is that while semantically potent modality is the right tool for intensionality, and that certain locality conditions may be tied to semantic domains of reference as a spin-off effect (e.g. ■ for tensed S constraints), semantically non-potent universal modality should not be used obsessively for syntactic domains, precisely because a global modalisation is required to deal with constraints that are empirically tied to specific elements.

Slightly ironically however, we consider that semantically non-potent ◇ *is* appropriate to specification of extraction *sites* (as opposed to domains). We are going to capture the idea of licensing extraction by arranging that the 'gap' subcategory of a fronted constituent is of the form ◇A. Then only argument positions marked ◇A will be extraction sites. Thus, a relative pronoun category R/(S↑(◇N)) will only bind positions which are ◇-ed. Subject-object asymmetry will be embodied, for example, in an assigment (N\S)/(◇N) to transitive verbs. Lexically realised elements will be able to occupy subject position as usual, and also object position since N ⇒ ◇N is valid. A relative pronoun will be able to bind an object position, but not a subject position, because ◇N ⇒ N is not valid. This strategy of modalisation does not have the shortcoming of being global, involving as it does only fillers and extraction sites themselves.

We extend the category system with semantically non-potent S5 ◇. The rule of proof is:

(94)

$$\frac{\Gamma \Rightarrow A}{\Gamma \Rightarrow \Diamond A}\Diamond R$$

(95)

$$\frac{\begin{array}{c}\vdots\\ \phi: A\end{array}}{\phi: \Diamond A}\Diamond I$$

We will not attempt to give a rule of use for ◇ in the {/, \, ·, ↑, []$^{-1}$, [], Q, ◇} categorial system we are using,[10] but note however that ◇A ⇒ A is not generally valid since there is no necessity that an inhabitant of ◇A (at a point i) is an inhabitant of A (at i).

In relation to constraints on extraction sites, recall the paradigm (7, 71), repeated here:

[10] S5 is simpler model-theoretically, but more difficult proof-theoretically, than S4. Like S4, a complete logic for S5 ◇ would need to make reference to its dual universal modality.

(96) a. the tunnel out of which$_i$ John emerged e_i

b. *the tunnel of which$_i$ John emerged out e_i

c. the tunnel which$_i$ John emerged out of e_i

This shows that while extraction from the PP complement of *out* is possible, extraction of that complement itself is not. In order to define the argument positions as being possible extraction sites or not we distribute the operator ◊ thus:

(97) which - (λx(λy(λz(λw[(z w)∧(y (x w))]))))
:= Q(N, N, (CN\CN)/(S↑(◊N)))
which - (λx(λy(λz(λw[(z w)∧(y (x w))]))))
:= Q(N, PP, (CN\CN)/(S↑(◊PP)))

(98) emerged - emerge
:= (N\S)/(◊PP)
John - j
:= N
of - of
:= PP/(◊N)
out - out
:= PP/PP

Then (96a) for example is derived thus:

(99) a.

```
                                 which
                      ------------------------------
                      Q(N, PP, (CN\CN)/(S↑(◊PP)))
                      ------------------------------ QE¹
             of                     N
          --------    ------------------------------ ◊I
  out     PP/(◊N)                  ◊N
-------   ------------------------------------------- /E
PP/PP                        PP
----------------------------------------------------- /E
                          PP
----------------------------------------------------- 1
                 (CN\CN)/(S↑(◊PP))
```

b.

```
                                        emerged
                            John    ------------  ----
                            ----    (N\S)/(◊PP)   ◊PP²
                              N     ------------------ /E
                                          N\S
     out of which           ---------------------- \E
-----------------------               S
(CN\CN)/(S↑(◊PP))           ---------------------- ↑I²
                                   S↑(◊PP)
---------------------------------------------------- /E
                    CN\CN
```

In relation to (96b) however, there is no derivation: the conflict between *out*, that requires a PP argument, and ***of which*** that binds a ◊PP position forces ungrammaticality.

In this way, stranding of PP-complement prepositions in English is blocked. In Romance languages where prepositions in general cannot be stranded, rather than, for example, *of*: PP/(◊N) we would have *de*: PP/N.

For sentence embedding we assume assignments such as (100).

(100) that - (λxx)
:= CP/S
thinks - **think**
:= (N\S)/CP

Then for instance the verb phrase 'thinks that Mary runs' has semantics (**think** (**run m**)). Subject-object asymmetry is reflected in assignment of transitive verbs to (N\S)/(◊N), capturing the fixed subject constraint.

(101) a. the man whom John thinks that Mary likes
b. *the man whom John thinks that likes Bill

This same asymmetry between subject and non-subject arguments enables capture of the following contrast

(102) a. the one who/that/whose CN/which walks
b. *the one whom walks

A category Q(N, N, (CN\CN)/(S↑(◊N))) for *whom* would correctly fail to generate (102b) since it can only bind diamonded positions, so it remains to find a category for *who, that, whose* CN, and *which* generating (102a): (CN\CN)/(N\S) suffices.[11] Note that an attempt to capture the contrast in (102) by restricting *whom* to accusitive case would fail to generate (103b) where *whom* binds the nominative subject position of a non-complementised embedded sentence:

[11] We shall show in the next section how to integrate these categories.

(103) a. John thinks she/*her likes Bill

b. the man whom$_i$ John thinks e_i likes Bill

An assignment of (N\S)/S to *thinks* would also fail to generate examples such as (103b). Such subject extractions can be licensed however by assuming that sentence-embedding verbs take as their argument (◊N)·(N\S), i.e. an extractable nominal followed by a verb phrase.[12]

(104) thinks - $(\lambda x(\mathbf{think}\ ((\pi_1 x)\ (\pi_2 x))))$

$:= (N\backslash S)/((\diamond N)\cdot(N\backslash S))$

The semantics is such as to make e.g. *thinks Mary runs* synonymous with *thinks that Mary runs*. A derivation of *thinks Mary runs* is as follows.

(105)

$$\dfrac{\dfrac{\text{thinks}}{(N\backslash S)/((\diamond N)\cdot(N\backslash S))} \qquad \dfrac{\dfrac{\dfrac{\text{Mary}}{N}}{\diamond N}\,\diamond I \quad \dfrac{\text{runs}}{N\backslash S}}{(\diamond N)\cdot(N\backslash S)}\,\cdot I}{N\backslash S}\,/E$$

The derivational semantics is (106a) which evaluates to (106b) with the lexical semantics.

(106) a. $(x_{\text{thinks}}\ (x_{\text{Mary}}, x_{\text{runs}}))$

b. $(\mathbf{think}\ (\mathbf{run}\ \mathbf{m}))$

Subject extraction of the kind in (103b) is derived thus:

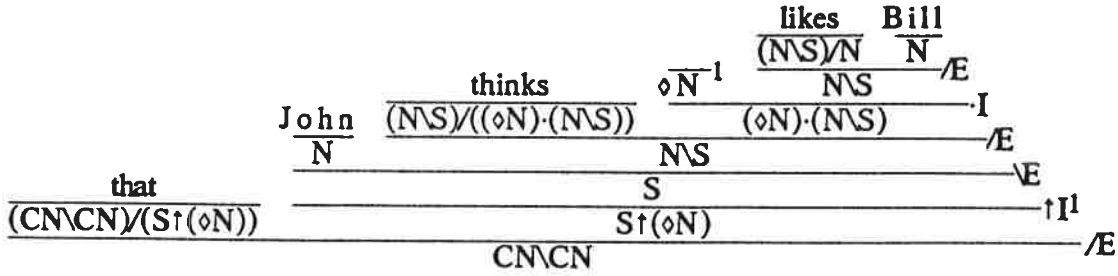
[12] In associative calculus, this can be equivalently defined curried:

(iii) thinks - $(\lambda x(\lambda y(\mathbf{think}\ (y\ x))))$

$:= ((N\backslash S)/(N\backslash S))/(\diamond N)$

Our use of product enables us to illustrate a little semantics with pairing and projection, for which the essential evaluation clauses are that $\pi_1(\phi, \psi)$ equals the first coordinate ϕ and $\pi_2(\phi, \psi)$ equals the second coordinate ψ.

(107)



Note that the categorisation (104) is the *only* one we need to assume for embedding of uncomplementised sentences, covering both canonical structures and subject extraction. We presently need to assume another entry however for the complementised case. These details are ironed out in appendix A which exploits economic devices presented in the next section.

Refinements

In this section we will collate properties of relative pronouns and extend a little the apparatus for their codification. In particular, we shall employ conjunctive and disjunctive type-constructors. For these connectives see Morrill (1990a, 1992).

The first category constructor we consider is a conjunction $\wedge$ thus:

(108) If A and B are category formulas,
$A \wedge B$ is a category formula;

(109) $t_{A \wedge B} = t_A \times t_B$;

(110) $D(A \wedge B) = \{\langle s, m \rangle \mid \exists \langle s, m_1 \rangle \in D(A), \exists \langle s, m_2 \rangle \in D(B), m = \langle m_1, m_2 \rangle\}$;

Thus an inhabitant of $A \wedge B$ is a sign the meaning of which is an ordered pair, and such that the symbol with meaning the first coordinate of that pair is a sign in A, and the symbol with meaning the second coordinate of that pair is a sign in B. Rules of use are as follows:

(111) a.

$$\frac{\Gamma(A) \Rightarrow D}{\Gamma(A \wedge B) \Rightarrow D} \wedge L_a$$

b.

$$\frac{\Gamma(B) \Rightarrow D}{\Gamma(A \wedge B) \Rightarrow D} \wedge L_b$$

(112) a.

$$\frac{\begin{matrix}\vdots \\ \chi : A \wedge B\end{matrix}}{(\pi_1 \chi) : A} \wedge E_a$$

b.

$$\frac{\begin{matrix}\vdots \\ \chi : A \wedge B\end{matrix}}{(\pi_2 \chi) : B} \wedge E_b$$

This conjunction has the property of considering different meanings in the component categories A and B; it is semantically potent. An alternative version may be used for signs that belong to two categories with the same meaning (in the same semantic domain). We use $\underline{\wedge}$ for this semantically non-potent, intersection, type-constructor. Since the meaning of a sign in $A \underline{\wedge} B$ must also be a meaning for signs in both A and B, the associated semantic type is defined as $\{\}$ when $t_A \neq t_B$. Consequently $D(A \underline{\wedge} B)$, a pairing of symbols with elements from the corresponding semantic domain, will be empty when this is the case.

(113) If A and B are category formulas,
$A \underline{\wedge} B$ is a category formula;

(114) $t_{A \underline{\wedge} B} = t_A$ if $t_A = t_B$, else $\{\}$;

(115) $D(A \underline{\wedge} B) = \{\langle s, m\rangle | \langle s, m\rangle \in D(A) \text{ and } \langle s, m\rangle \in D(B)\}$

Rules of inference are thus:

(116) a.

$$\frac{\begin{matrix}\vdots \\ \chi : A \underline{\wedge} B\end{matrix}}{\chi : A} \underline{\wedge} E_a$$

b.

$$\frac{\begin{matrix}\vdots \\ \chi : A \underline{\wedge} B\end{matrix}}{\chi : B} \underline{\wedge} E_b$$

Dual to conjunction is disjunction.

(117) If A and B are category formulas,
$A \vee B$ is a category formula;

(118) $t_{A \vee B} = t_A \cup t_B$

(119) $D(A\vee B) = \{\langle s, m\rangle \mid m = \langle 1, m'\rangle$, and $\langle s, m'\rangle \in D(A)$, or $m = \langle 2, m'\rangle$, and $\langle s, m'\rangle \in D(B)\}$

The semantic type for A∨B is the disjoint union $t_A \uplus t_B$, $= (\{1\}\times t_A)\cup(\{2\}\times t_B)$. An inhabitant of A∨B is a sign the meaning of which is an ordered pair such that if the first coordinate is 1, the symbol with the second coordinate is a sign in A, and if the first coordinate is 2, the symbol with the second coordinate is a sign in B.[13] Rules of proof are thus:

(120) a.

$$\frac{\Gamma \Rightarrow A}{\Gamma \Rightarrow A\vee B}\vee R_a$$

b.

$$\frac{\Gamma \Rightarrow A}{\Gamma \Rightarrow A\vee B}\vee R_b$$

(121) a.

$$\frac{\begin{array}{c}\vdots\\ \phi: A\end{array}}{(\iota_1\phi): A\vee B}\vee I_a$$

b.

$$\frac{\begin{array}{c}\vdots\\ \psi: B\end{array}}{(\iota_2\psi): A\vee B}\vee I_b$$

There is a corresponding semantically non-potent union type-constructor. An inhabitant of $A\underline{\vee}B$ is either an inhabitant of A or an inhabitant of B. Again, $t_{A\underline{\vee}B}$, and consequently the category, is empty when the types for the component categories are not the same.

(122) If A and B are category formulas,
$A\underline{\vee}B$ is a category formula;

(123) $t_{A\underline{\vee}B} = t_A$ if $t_A = t_B$, else $\{\}$;

(124) $D(A\underline{\vee}B) = \{\langle s, m\rangle \mid \langle s, m\rangle \in D(A)$ or $\langle s, m\rangle \in D(B)\}$;

Rules of proof are thus:

[13] The associated computational operation is branching by case. We represent a branch statement by $(\chi \rightarrow x.\delta_1; y.\delta_2)$ and first and second injections by $(\iota_1\phi)$ and $(\iota_2\phi)$. Then $((\iota_1\phi)\rightarrow x.\delta_1; y.\delta_2)$ evaluates to δ_1 with ϕ substituted for free occurrences of x, and $((\iota_2\phi)\rightarrow x.\delta_1; y.\delta_2)$ evaluates to δ_2 with ψ substituted for free occurrences of y.

(125) a. b.

$$\frac{\begin{array}{c}\vdots\\ \phi: A\end{array}}{\phi: A\underline{\vee}B}\vee I_a \qquad \frac{\begin{array}{c}\vdots\\ \psi: B\end{array}}{\psi: A\underline{\vee}B}\vee I_b$$

Properties of the relative pronouns of English are set out in (126). The first row indicates whether the pronoun (or its nominal projection in the case of the determiner *whose*) can appear as the main subject in the body of a relative clause. The second indicates whether it (or its projection) can pied pipe.

(126)

	who	that	whom	which	whose
main subject	1	1	0	1	1
pied piping	0	0	1	1	1

The pronouns *who* and *that* can occur as main subject, but cannot pied pipe. The standard relativisation category (CN\CN)/(S↑(◊N)) and the main subject category (CN\CN)/(N\S) can be integrated by taking advantage of the fact that the semantics is the same in the two cases -- as required, $t_{S\uparrow(\Diamond N)} = t_{N\backslash S}$ (= E→{0, 1}) -- and defining a single functor over (S↑(◊N)∨(N\S): [14]

(127) who - (λx(λy(λz[(y z)∧(x z)])))
:= (CN\CN)/((S↑(◊N))$\underline{\vee}$(N\S))
that - (λx(λy(λz[(y z)∧(x z)])))
:= (CN\CN)/((S↑(◊N))$\underline{\vee}$(N\S))

The pronoun *whom* cannot occur as main subject, but can pied pipe in nominals and prepositional phrases. Where we assume that t_{PP} = E = t_N, the semantics for the two cases is the same, and use of $\underline{\wedge}$ obviates the need to specify it twice:[15]

(128) whom - (λx(λy(λz(λw[(z w)∧(y (x w))]))))
:= Q(N, PP, (CN\CN)/(S↑(◊PP)))$\underline{\wedge}$Q(N, N, (CN\CN)/(S↑(◊N)))

The pronoun *which* can both pied pipe in nominals and PPs, and occur as main subject. Since the semantics for the latter however is not the same as that as the former,

[14]We do not trouble at the moment to mark the islandhood of relative clauses.

[15]An appropriate featural discrimination between N and PP, defining them as e.g. NonVerb(+Noun) and NonVerb(-Noun) might provide for further appropriate simplification, by undersecification, i.e. universal quantification.

we need to use the semantically potent $\underline{\wedge}$ to collapse the properties to a single category, the semantics of which is an ordered pair comprising the pied piping semantics and the standard semantics. Note that in this case no real economy is gained: we might just about as well have two lexical entries.

(129) which $- ((\lambda x(\lambda y(\lambda z(\lambda w[(z\ w)\wedge(y\ (x\ w))])))), (\lambda x(\lambda y(\lambda z[(y\ z)\wedge(x\ z)]))))$

$$:= (Q(N, PP, (CN\backslash CN)/(S{\uparrow}(\Diamond PP)))\underline{\wedge}Q(N, N, (CN\backslash CN)/(S{\uparrow}(\Diamond N))))$$

$$\wedge((CN\backslash CN)/(N\backslash S))$$

The phrase formed by *whose* when it combines with a common noun is like *which* in that it can pied pipe and can be main subject. Its category is therefore a functor over CN into the same category as *which.*

(130) whose $- (\lambda s((\lambda x(\lambda y(\lambda z(\lambda w[(z\ w)\wedge(y\ (x\ [\iota t[(s\ t)\wedge(\mathbf{poss}\ (t, w))]]))])))),$

$$(\lambda x(\lambda y(\lambda z[(y\ z)\wedge(x\ [\iota t[(s\ t)\wedge(\mathbf{poss}\ (t, z))]])])))))$$

$$:= ((Q(N, PP, (CN\backslash CN)/(S{\uparrow}(\Diamond PP)))\underline{\wedge}Q(N, N, (CN\backslash CN)/(S{\uparrow}(\Diamond N))))$$

$$\wedge((CN\backslash CN)/(N\backslash S)))/CN$$

This concludes our explication of the categorial formalism and its treatment of the data. The grammar is summarised in appendix A, and appendix B indicates further applicability of the bracket operators. The grammar is included to emphasise how all the characterisations we have considered coexist. Treatment of features, distinguishing e.g. verb inflection and preposition forms, is conspicuously absent. Morrill (1990a, 1992) shows how such polymorphism would be treated by means of first-order quantified category formulas; see also van Benthem (1989b) on the second-order case. Implementation of the associated inference proceeds naturally by unification; in this way the relation with unification grammar is established.

Unification grammar still enjoys virtual identification with formal grammar. We therefore feel justified in observing, without studying the issues in depth here, that for example, an HPSG cover grammar of appendix A would be considerably less parsimonious. Means of indirection such as aliases and templates could improve economy, but equally they could improve economy further here: assembly language does not become high level because it has text macros. We hope the present paper has suggested that unification grammar is no longer the only, or even the best, formalism for implementation of significant grammar fragments.

Appendix A: Summary grammar

The set F of category formulas is defined in terms of the set $P = \{N, S, PP, CP, VP\}$ of atomic category formulas thus:

$$\text{(A1)} \quad F = P \mid (F/F) \mid (F\backslash F) \mid (F{\cdot}F) \mid (F{\uparrow}F) \mid Q(F, F, F) \mid ([\,]F) \mid ([\,]^{-1}F) \mid ({\diamond}F) \mid (F{\wedge}F) \mid (F{\vee}F) \mid (F\underline{\wedge}F) \mid (F\underline{\vee}F)$$

A language model is a quadruple $(\{v, \ldots,], [\}, I, \{\tau_A\}_{A \in P}, d)$. $\{v, \ldots\}$ is a set of vocabulary items. This defines the set $V^{+, []}$ which is.the closure of unit strings on vocabulary items under the binary operation of concatenation $\oplus$ and the unary operation b of delimitation by outermost matching brackets (i.e. $b(s) = [\oplus s \oplus]$). I is a set (points). $\{t_A\}_{A \in P}$ is a family of non-empty basic semantic domains, one for each atomic category formula. This type map is extended to all category formulas thus:

$$
\begin{aligned}
\text{(A2)} \quad & t_{A\backslash B} = t_A \rightarrow t_B;\\
& t_{B/A} = t_A \rightarrow t_B;\\
& t_{A \cdot B} = t_A \times t_B;\\
& t_{B \uparrow A} = t_A \rightarrow t_B;\\
& t_{Q(A, B, C)} = (t_A \rightarrow t_B) \rightarrow t_C;\\
& t_{[\,]A} = t_{[\,]^{-1}A} = t_A;\\
& t_{\diamond A} = t_A;\\
& t_{A \wedge B} = t_A \times t_B;\\
& t_{A \underline{\wedge} B} = t_A \text{ if } t_A = t_B, \text{ else } \{\};\\
& t_{A \vee B} = t_A \cup t_B;\\
& t_{A \underline{\vee} B} = t_A \text{ if } t_A = t_B, \text{ else } \{\};
\end{aligned}
$$

Relative to each $i \in I$, the interpretation function d maps each atomic category formula A to a subset $d(A)^i$ of $V^{+, []} \times \tau_A$. This is extended to an interpretation relative to each i of every category formula A as a subset $D(A)^i$ of $V^{+, []} \times \tau_A$ thus:

$$
\begin{aligned}
\text{(A3)} \quad & D(A)^i && = d(A)^i \text{ for atomic A};\\
& D(B/A)^i && = \{\langle s, m\rangle \mid \forall \langle s', m'\rangle \in D(A)^i, \langle s \oplus s', m(m')\rangle \in D(B)^i\};\\
& D(A\backslash B)^i && = \{\langle s, m\rangle \mid \forall \langle s', m'\rangle \in D(A)^i, \langle s' \oplus s, m(m')\rangle \in D(B)^i\};\\
& D(A \cdot B)^i && = \{\langle s, m\rangle \mid \exists \langle s_1, m_1\rangle \in D(A)^i, \exists \langle s_2, m_2\rangle \in D(B)^i,\\
& && \qquad s = s_1 \oplus s_2, m = \langle m_1, m_2\rangle\};\\
& D(A \uparrow B)^i && = \{\langle s, m\rangle \mid \exists s_1 s_2 \in V^{*, []}, s = s_1 \oplus s_2,\\
& && \qquad \forall \langle s', m'\rangle \in D(B)^i, \langle s_1 \oplus s' \oplus s_2, m(m')\rangle \in D(A)^i\};
\end{aligned}
$$

$$D(Q(A, B, C))^i = \{\langle s, m\rangle \mid \forall s_1, s_3 \in V^{*, []}, \forall m' \in t_A \to t_B, [\forall\langle s_2, m_2\rangle \in D(A)^i, \langle s_1 \oplus s_2 \oplus s_3, m'(m_2)\rangle \in D(B)^i] \to \langle s_1 \oplus s \oplus s_3, m(m')\rangle \in D(C)^i\};$$

$$D([\,]A)^i = \{\langle s, m\rangle \mid \exists\langle s', m\rangle \in D(A)^i, s = b(s')\};$$

$$D([\,]^{-1}A)^i = \{\langle s, m\rangle \mid \exists\langle s', m\rangle \in D(A)^i, s' = b(s)\};$$

$$D(\diamond A)^i = \{\langle s, m\rangle \mid \exists j, \langle s, m\rangle \in D(A)^j\}$$

$$D(A \wedge B)^i = \{\langle s, m\rangle \mid \exists\langle s, m_1\rangle \in D(A)^i, \exists\langle s, m_2\rangle \in D(B)^i, m = \langle m_1, m_2\rangle\};$$

$$D(A \underline{\wedge} B)^i = \{\langle s, m\rangle \mid \langle s, m\rangle \in D(A)^i \text{ and } \langle s, m\rangle \in D(B)^i\}$$

$$D(A \vee B)^i = \{\langle s, m\rangle \mid m = \langle 1, m'\rangle, \text{ and } \langle s, m'\rangle \in D(A)^i, \text{ or } m = \langle 2, m'\rangle, \text{ and } \langle s, m'\rangle \in D(B)^i\}$$

$$D(A \underline{\vee} B)^i = \{\langle s, m\rangle \mid \langle s, m\rangle \in D(A)^i \text{ or } \langle s, m\rangle \in D(B)^i\};$$

A lexicon summarising the coverage and machinary of the paper is as follows.

(A4)

(A4)	and	- $(\lambda x(\lambda y[x \wedge y]))$
		:= $(S\backslash([\,]^{-1}S))/S$
	annoys	- **annoy**
		:= $(([\,]CP\backslash S)/N$
	assures	- **assure**
		:= $((N\backslash(S\uparrow N))/VP)/N$
	believes	- $(\lambda x(\textbf{believe}\ (x \to y.y; z.((\pi_2 z)(\pi_1 z)))))$
		:= $(N\backslash S)/(CP \vee ((\diamond N)\cdot(N\backslash S)))$
	brother	- $(\lambda x(x\ \textbf{brother}))$
		:= $CN/(\diamond PP)$
	emerges	- **emerge**
		:= $(N\backslash S)/(\diamond PP)$
	for	- **for**
		:= $PP/(\diamond N)$
	John	- **j**
		:= N
	knows	- know
		:= $(N\backslash S)/(\diamond N)$
	likes	- **like**
		:= $(N\backslash S)/(\diamond N)$
	man	- **man**
		:= CN

of	- **of**
	$:= PP/(\diamond N)$
out	- **out**
	$:= PP/PP$
physics	- **physics**
	$:= N$
's	$- (\lambda x(\lambda y[\iota z[(y\ z) \wedge (\mathbf{poss}\ (z, x))]]))$
	$:= N\backslash([\,]^{-1}(N/CN))$
subject	- **subject**
	$:= CN$
that	$- (\lambda xx)$
	$:= CP/S$
that	$- (\lambda x(\lambda y(\lambda z[(y\ z) \wedge (x\ z)])))$
	$:= ([\,]^{-1}(CN\backslash CN))/((S\uparrow(\diamond N))\underline{\vee}(N\backslash S))$
the	$- (\lambda x[\iota y(x\ y)])$
	$:= N/CN$
today	- **today**
	$:= (N\backslash S)\backslash(N\backslash S)$
tunnel	- **tunnel**
	$:= CN$
votes	- **vote**
	$:= (N\backslash S)/(\diamond PP)$
walks	- **walk**
	$:= N\backslash S$
which	$- ((\lambda x(\lambda y(\lambda z(\lambda w[(z\ w) \wedge (y\ (x\ w))])))), (\lambda x(\lambda y(\lambda z[(y\ z) \wedge (x\ z)]))))$
	$:= (Q(N, PP, ([\,]^{-1}(CN\backslash CN))/(S\uparrow(\diamond PP))) \underline{\wedge} Q(N, N, ([\,]^{-1}(CN\backslash CN))/(S\uparrow(\diamond N)))) \wedge (([\,]^{-1}(CN\backslash CN))/(N\backslash S))$
who	$- (\lambda x(\lambda y(\lambda z[(y\ z) \wedge (x\ z)])))$
	$:= ([\,]^{-1}(CN\backslash CN))/((S\uparrow(\diamond N))\underline{\vee}(N\backslash S))$
whom	$- (\lambda x(\lambda y(\lambda z(\lambda w[(z\ w) \wedge (y\ (x\ w))]))))$
	$:= Q(N, PP, ([\,]^{-1}(CN\backslash CN))/(S\uparrow(\diamond PP))) \underline{\wedge} Q(N, N, ([\,]^{-1}(CN\backslash CN))/(S\uparrow(\diamond N)))$

whose - $(\lambda s((\lambda x(\lambda y(\lambda z(\lambda w[(z\ w)\wedge(y\ (x\ [\iota t[(s\ t)\wedge(\text{poss}\ (t, w))]]))])))),$

$(\lambda x(\lambda y(\lambda z[(y\ z)\wedge(x\ [\iota t[(s\ t)\wedge(\text{poss}\ (t, z))]])]))))$

$:= ((Q(N, PP, ([\,]^{-1}(CN\backslash CN))/(S\uparrow(\diamond PP)))$

$\underline{\wedge}Q(N, N, ([\,]^{-1}(CN\backslash CN))/(S\uparrow(\diamond N))))$

$\wedge(([\,]^{-1}(CN\backslash CN))/(N\backslash S)))/CN$

Appendix B: Bracket Operators, Structured Consequence, and Finer Discrimination

In this appendix we make three further remarks on the bracketing logic, in relation to embedding, bounded structural resource and varying penetrative power, and logic of labelled bracketing.

Girard (1987) shows that the structural modalities of linear logic are of interest in that they enable *travel* between logics; in particular translations can be given of sequents of intuitionistic formulas into sequents of linear formulas which preserve theoremhood.[16] This is discussed for substructural logics more generally in Došen. (1990b). In these cases the relevant operators provide a structural flexibility to the formulas that bear them enabling embedding of a strong logic in a weak one with structurally *facilitating* operators. We suggest here that the bracket operators provide an instance of the converse mapping: embedding of a weak logic in a strong one with structurally *inhibiting* operators. In particular, we conjecture that the following translation of non-associative Lambek formulas into associative formulas with $[\,]^{-1}$ preserves theoremhood.[17]

(B1)
$$|A| = A \text{ for atomic } A$$
$$|A\backslash B| = |A|\backslash([\,]^{-1}|B|)$$
$$|B/A| = ([\,]^{-1}|B|)/|A|$$

The account so far defines islands in an all-or-nothing manner. We have seen in §2 however that island status is not so clear-cut. It could be argued, for instance, that left

---

[16]For the implicational case the mapping is $|.|$ in (iv).

(iv)
$$|A_1, \ldots, A_n \rightarrow A| = !|A_1|, \ldots, !|A_n| \rightarrow |A|$$
$$|A| = A \text{ for atomic } A$$
$$|A \rightarrow B| = (!|A|) \rightarrow |B|$$

[17]There seems to be a certain symmetry here: whereas the embedding of strong in weak involves marking the antecedents of implications with structurally facilitating operators, embedding of weak in strong involves marking the consequents with structurally inhibiting operators.

extraction of NPs is semi-acceptable from certain subject relative clauses, the objects of some verbs (e.g. 'destroy'), the complements of possessive NPs, some adverbials, sentential noun complements, subjects, and the complements of factive verbs. Furthermore, a domain which is not an island to NP extraction may nethertheless be an island to PP extraction:

(B2) a. the man who I went to London without speaking to
b. *the man to whom I went to London without speaking

We consider here how the existing apparatus can be sensitised to deal with such facts. The relevant concept is resource bounded structural potential. Thus one of the branches of linear logic deals with structural modalities that license a limited amount of structural manipulation (e.g. *n* contractions for some finite *n*); see e.g. Girard, Scedrov and Scott (1990). The approach here is to define fillers that can penetrate bracketings to bounded degrees of nesting. We shall actually just consider penetration of one or zero levels of bracketing. In this respect we echo Chomsky's notion of barriers, with more than one barrier creating an absolute island.

In terms of Moortgat;s $\uparrow$, there are now operators $\uparrow_n$ for $n = 0, 1, \ldots$ capable of penetrating up to *n* bracketings.

(B3) For finite *n*, if A and B are category formulas,
$B{\uparrow_n}A$ is a category formula;
$T(B{\uparrow_n}A) = t_A \to t_B$;
$D(B{\uparrow_n}A) = \{\langle s, m\rangle \mid$ there exists a point nested not deeper than *n* in s such that where $s\rfloor P\lfloor$ represents interpolation of P at this point, $\forall\langle s', m'\rangle \in D(A), \langle s\rfloor s'\lfloor, m(m')\rangle \in D(B)\}$;

Suppose now that without has category $([\,]^{-1}(VP\backslash VP))/VP$, i.e. that it projects a level of bracketing. If the filler *who* has category $(CN\backslash CN)/(S{\uparrow_1}N)$ it can still penetrate this single bracketing, as in (A1a); but if 'to whom' only has type $(CN\backslash CN)/(S{\uparrow_0}PP)$, the adverbial is an island to the extraction, as in (B2b). Absolute islands, on the other hand, would be created be double bracketings; e.g. *and* would be $(S\backslash([\,]^{-1}([\,]^{-1}S)))/S$.

Note finally that by using operators $[\,]_a$, $[\,]^{-1}_a$, $]_b$, $[\,]^{-1}_b$, for a variety of labels *a*, *b*, *labelled* bracketings are generated. In this way any combination of hierarchical continuous domains can be described. Constraints with respect to these would be approachable in the style of the previous paragraph, with operators with bounded structural potential with respect to each of the bracket 'colours': e.g. 'I have resources to pass through up to two *a* bracketings and one *b* bracketing'.

References

Bach, E. and G. Horn (1976), 'Remarks on "Conditions on Transformations"', *Linguistic Inquiry* 7, 265-299.

Barry, G., M. Hepple, N, Leslie and G. Morrill (1991), 'Proof Figures and Structural Operators for Categorial Grammar'. *Proceedings of the meeting of the European Chapter of the Association for Computational Linguistics*, Berlin, Germany.

van Benthem, J. (1989a), 'Categorial Grammar and Type Theory', *Journal of Philosophical Logic* 19, 115-168.

van Benthem, J. (1989b), 'Categorial Grammar Meets Unification', to appear in J. Wedekind (ed.).

van Benthem, J. (1991), *Language in Action: Categories, Lambdas and Dynamic Logic*, North Holland, Amsterdam.

Bresnan, J. (1972), Theory of complementation in English Syntax. PhD thesis, MIT.

Buszkowski, W. (1986) 'Completeness Results for Lambek Syntactic Calculus', *Zeitschrift für mathematische Logik und Grundlagen der Mathematik* 32, 13-28.

Chomsky, N. (1973), 'Coditions on transformations', in S. Anderson and P. Kiparsky (eds.), *A Festschrift for Morris Halle*, Holt, Rinehart and Winston, New York, pp. 232-286.

Chomsky, N. and H. Lasnik (1977), 'Filters and Control', *Linguistic Inquiry* 8, 425-504.

Došen, K. (1990a), 'Modal logic as metalogic', SNS report, Tübingen.

Došen, K. (1990b), 'Modal translations in substructural logics'. Report 10-90, Universität Konstanz.

Gazdar, G. (1991), 'Coordinate structures and unbounded dependencies', *Linguistic Inquiry*.

Gazdar, G., E. Klein, G. Pullum, I. Sag (1985), *Generalised Phrase Structure Grammar*, Blackwell, Oxford.

Girard, J-Y. (1987) 'Linear Logic', *Theoretical Computer Science* 50, 1-102.

Girard, J-Y, A. Scedrov and P. Scott (1990), 'Bounded Linear Logic', to appear in S. Boss and P. Scott (eds.) *Feasible Mathematics*

Hepple, M. (1990), *The Grammar and Processing of Order and Dependency: A Categorial Approach*, PhD thesis, Centre for Cognitive Science, Edinburgh.

Kayne, R. (1984), *Connectedness and Binary Branching*, Foris, Dordrecht.

Lambek, J. (1958), 'The Mathematics of Sentence Structure', *Americam Mathematical Monthly* 65, pp. 154-169. Reprinted in W. Buszkowski, W. Marciszewski and J. van Benthem 1988 (eds.) *Categorial Grammar*, Benjamins, Amsterdam.

Lambek, J. (1961), 'On the Calculus of Syntactic Types', in R. Jakobson (ed.) *Structure of Language and its Mathematical Aspects*, Providence, pp. 166-178.

Moortgat, M. (1988a), 'Mixed Composition and Discontinuous Dependencies', in R. Oehrle, E. Bach and D. Wheeler (eds.), *Categorial Grammars and Natural Language Structures*, Kluwer Academic Publishers, Dordrecht, pp. 319-348.

Moortgat, M. (1988b), *Categorial Investigations: Logical and Linguistic Aspects of the Lambek Calculus*, Foris, Dordrecht.

Moortgat, M. (1990), 'The quantification calculus: questions of axiomatisation', in H. Hendriks and M. Moortgat 1990, *Theory of Flexible Interpretation*. Esprit DYANA Report R1.2.A.

Moortgat, M. (1991), 'Discontinuous Type-Constructors', ms. OTS, Rijksuniversiteit Utrecht.

Moortgat, M. (1992), 'Labelled Deductive Systems for categorial theorem proving', Working Paper, OTS, Rijksuniversiteit Utrecht; also in P. Dekker and M. Stokhof (eds.) *Proceedings of the Eighth Amsterdam Colloquium*.

Morrill, G. (1992), *Type-Logical Grammar*, Report de Recerca, Departament de Llenguatges i Sistemes Informàtics, Universitat Politècnica de Catalunya and Working Paper, OTS, Rijksuniversiteit Utrecht. Research monograph to appear in the series *Studies in Language, Logic and Information*, Kluwer Academic Publishers, Dordrecht.

Morrill (1989) 'Intensionality, Boundedness, and Modal Logic', Edinburgh University Centre for Cognitive Science Research Paper No. 32.

Morrill, G. (1990a) 'Grammar and Logical Types', in G. Barry and G. Morrill (eds.) *Studies in Categorial Grammar*, Edinburgh Working Papers in Cognitive Science, Volume 5, pp. 127--148. Also in *Proceedings of the Seventh Amsterdam Colloquium*, Institute for Language, Logic and Information, Amsterdam, Pt. 2, pp. 429--450.

Morrill (1990b), 'Intensionality and Boundedness', *Linguistics and Philosophy* 13, pp. 699--726.

Morrill (1990c), 'Categorial logic for binding and coordination phenomena', Esprit DYANA Report R1.2.D.

Morrill, G., N. Leslie, M. Hepple and G. Barry (1990) 'Categorial Deductions and Structural Operations', in G. Barry and G. Morrill (eds.) *Studies in Categorial Grammar*, Edinburgh Working Papers in Cognitive Science, Volume 5, pp. 1-21.

Pollard, C. (1988), 'Categorial Grammar and Phrase Structure Grammar: An Excursion on the Syntax-Semantics Frontier', in R. Oehrle, E. Bach and D. Wheeler (eds.), *Categorial Grammars and Natural Language Structures*, Kluwer Academic Publishers, Dordrecht, pp. 391-415.

Pollard, C. and I. Sag (1987), *An Information-Based Approach to Syntax and Semantics Volume 1*, CSLI Lecture Notes, Stanford.

Pollard, C. and I. Sag (1992), *An Information-Based Approach to Syntax and Semantics Volume 2*, to appear, University of Chicago Press, Chicago.

Ross, J. (1967), *Constraints on variables in syntax*. PhD thesis, MIT; Indiana University Linguistics Club.

Steedman, M. (1985), 'Dependency and Coordination in the Grammar of Dutch and English', *Language* 61, pp. 523-568.

Steedman, M. (1987), 'Combinatory Grammars and Parasitic Gaps', *Natural Language and Linguistic Theory* 5, pp. 403-439.

Zielonka, W. (1981), 'Axiomatisability of Ajdukiewicz-Lambek Calculus by Means of Cancellation Schemes', *Zeitschrift für mathematische Logik und Grundlagen der Mathematik* 27, 215-224.

Venema, Y. (1992), 'Meeting Modalities?', ms., Department of Philosophy, Rijksuniversiteit Utrecht.