A General Framework for Hybrid Substructural Categorial Logics

Mark Hepple
Institute for Research in Cognitive Science
University of Pennsylvania
3401 Walnut Street, Suite 400C
Philadelphia, PA 19104-6228
215-898-0330
hepple@linc.cis.upenn.edu

Draft: August 31, 1993.

1 Introduction

For the class of systems known as Categorial Grammars, grammar formalisms consist of logics. Lexical assignment is of complex formulas or types, which by their structure may encode various syntactic information such as subcategorisation and word order requirements. Syntactic derivation is via deduction over lexical formulas. Alternative categorial systems differ in the logic they employ to provide a notion of derivability.

There exist a wide range of different logics, which may be classified with respect to their limitations (if any) on the use of the resources that are available to serve deduction, and their consequent sensitivity to the specific structured nature of those resources. Comparison of logics in terms of their resource sensitivity gives rise to the so-called 'substructural hierarchy of logics'. Various systems on this hierarchy have been employed within Categorial Grammar (e.g. associative Lambek calculus, Lambek 1958; non-associative Lambek calculus, 1962). Indeed, some systems situated at previously unoccupied locations on this hierarchy have been proposed for specifically linguistic purposes (e.g. Moortgat & Morrill, 1991), and, no doubt, further ones remain to be developed.

[0] The author is pleased to acknowledge the support of an I.R.C.S. Postdoctoral Fellowship award.

It has become clear, however, that access to more than one substructural level is required not only cross-linguistically, for producing the grammars of very different languages, but also for specifying the grammar of any one language. The most plausible current model for how this may be done involves firstly selecting a specific resource logic as the 'basic' logic, providing the predominant level at which the grammar to be stated will operate. This choice sets the *default* characteristics of resource sensitivity. Then, special modal operators, termed *structural modalities*, may be used to allow controlled access to the resource sensitivity of higher substructural levels, by restrictedly undermining sensitivity to some specific aspect of resource structure. A number of theoretical and computational problems arise for the use of such operators. Furthermore, the complexity of syntactic analyses where extensive use of structural modalities is required tends to encourage the selection of the highest workable level possible for the default level of the logic. The discarding of resource sensitivity that such 'high selection' involves may not always be in the long term interests of developing adequate linguistic accounts.

It is my intention in this paper to propose an alternative general model of *hybrid substructural systems*, which should eliminate the need for structural modalities, and avoid their associated problems. The new model exploits natural relations between different substructural levels in terms of the relative informativeness of their characterisations. Under this model, the range of substructural levels form a single unified descriptive system, which may be used for different languages, although very different languages will tend to exploit different modes of description made available by the overall system. Such a unified approach should both facilitate producing grammars for individual languages, and provide a better basis for cross-linguistic generalisation.

OUTLINE: I will begin by introducing the topic of resource sensitivity in the context of some categorial logics, and illustrate how structural modalities may be used at a substructural level to give controlled access to the resource sensitivity of stronger logics. Discussion of systems involving structural modalities is used to arrive at a view of how different substructural levels might be interrelated in hybrid logics where different substructural levels coexist without structural rules. Then, a general approach for formulating hybrid logics from a sequent perspective is presented. Issues of term assignment and the treatment of word order are then discussed. An algebraic semantics for the hybrid approach is provided, and used in formulating a *labelled* natural deduction formulation for hybrid logics. Finally, the general approach is illustrated by discussion of some linguistic applications.

2 Some sample substructural systems

The topic of sensitivity to resource structure is most easily introduced in relation to sequent formulations of logics. Restricting our attention to the intuitionistic realm, a sequent is an object of the form $\Gamma \Rightarrow A$, indicating that the formula A may be deduced from the structured configuration of antecedent formulas Γ.

Let us consider a logic where Γ of any sequent $\Gamma \Rightarrow A$ is a binarily bracketted sequence of types, where the set of types is freely generated from a set of basic types using the three binary connectives $\{\circ, \leftarrow, \rightarrow\}$, and which has the inference rules in (1) and (2). ($\Gamma[\Delta]$ represents a configuration having Δ as a subconfiguration, and $\Gamma[\Delta']$ represents the result of replacing Δ with Δ' in $\Gamma[\Delta]$.)

(1)

$$A \Rightarrow A \quad (\mathrm{id}) \qquad\qquad \frac{\Delta \Rightarrow B \qquad \Gamma[B] \Rightarrow A}{\Gamma[\Delta] \Rightarrow A}\,[\mathrm{cut}]$$

(2)

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

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

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

The axiom and Cut rule in (1) express the reflexivity and transitivity of derivability. From the operational rules in (2) it should be clear that $\rightarrow$ and $\leftarrow$ are directional versions of implicational connectives. The 'product' connective $\circ$ is a form of conjunction, corresponding to 'matter-like addition' of substructures.

If no further rules are added, the above logic is a version of what is known as the non-associative Lambek calculus (Lambek, 1962). This logic is the weakest familiar system in the substructural landscape,[1] where deduction is sensitive to the linear order and bracketting of assumptions, each of which must be used precisely once in the course of a deduction. However, it is possible to undermine sensitivity to aspects of resource structure by inclusion of *structural rules* such as those in (3), which act to modify the structure of the antecedent configuration. (The double line of the Association rule indicates that it may be used in either direction.)

(3)

$$\begin{array}{cccc} \text{Association} & \text{Permutation} & \text{Weakening} & \text{Contraction} \\[4pt] \dfrac{\Gamma[(B, (C, D))] \Rightarrow A}{\overline{\Gamma[((B, C), D)] \Rightarrow A}}\,[\mathrm{A}] & \dfrac{\Gamma[B, C] \Rightarrow A}{\Gamma[C, B] \Rightarrow A}\,[\mathrm{P}] & \dfrac{\Gamma[\,] \Rightarrow A}{\Gamma[B] \Rightarrow A}\,[\mathrm{W}] & \dfrac{\Gamma[B, B] \Rightarrow A}{\Gamma[B] \Rightarrow A}\,[\mathrm{C}] \end{array}$$

The rules of Association and Permutation undermine sensitivity to the bracketting and linear order of assumptions, respectively. The Weakening rule allows resources to be 'wasted' (i.e.

[1] However, weaker systems are possible, and have been proposed, e.g. Moortgat & Morrill (1991).

not used at all in a deduction), whilst the contraction rule allows multiple use of resources.[2]

Various logics may be defined via the structural rules that they include. For example, adding the Association rule to the earlier non-associative logic undermines sensitivity to bracketting, and gives a version of the *associative* Lambek calculus (Lambek, 1962). If Permutation is also added, sensitivity to linear order is undermined, and we have a system sometimes known as LP ('Lambek+Permutation') or the Lambek-van Benthem calculus (van Benthem, 1983; 1991), also corresponding to a fragment of linear logic (Girard, 1987).[3] Adding also Weakening gives a resource discipline appropriate for Relevance logic (again, we have only a fragment thereof). Including all four structural rules gives a resource discipline appropriate to intuitionistic logic. The very free resource discipline of this latter system (as well as classical logic) has been seen, historically, as the most natural, and so systems having less freedom are called *sub*-structural. Comparison of such systems in terms of their resource discipline gives rise to the so-called 'substructural hierarchy' of logics.[4] For the remainder of this paper, we shall focus on systems with linear resource usage (i.e. where Contraction and Weakening are excluded, so that each resource must be used precisely once.)

Note that in systems with Permutation, the distinction between $\leftarrow$ and $\rightarrow$ breaks down, i.e. we have the interderivability $\mathrm{Y}{\rightarrow}\mathrm{X}{\Leftrightarrow}\mathrm{X}{\leftarrow}\mathrm{Y}$, and more generally, substituting $\mathrm{X}{\leftarrow}\mathrm{Y}$ for any subformula $\mathrm{Y}{\rightarrow}\mathrm{X}$, or vice versa, in any sequent preserves theoremhood. Hence, $\leftarrow$ and $\rightarrow$ are not normally even distinguished in systems with Permutation. Crucial cases identifying systems that include Association are the possibility of 'composing' implications (e.g. $\mathrm{X}{\leftarrow}\mathrm{Y},\ \mathrm{Y}{\leftarrow}\mathrm{Z} \Rightarrow \mathrm{X}{\leftarrow}\mathrm{Z}$), and the possiblity of transformations that change 'order of combination' (e.g. $(\mathrm{Y}{\rightarrow}\mathrm{X}){\leftarrow}\mathrm{Z} \Rightarrow \mathrm{Y}{\rightarrow}(\mathrm{X}{\leftarrow}\mathrm{Z})$).

It will be useful to distinguish notationally the cases of the three type constructors $\{\circ, \leftarrow, \rightarrow\}$ in systems have different structural rules. In general, for any product type constructor $\circ$ (here using $\circ$ as a placeholder), the corresponding left and right implications (or 'divisions') may be written $\overset{\circ}{\rightarrow}$ and $\overset{\circ}{\leftarrow}$. The four possible linear systems that arise by selection of rules from Association and Permutation will reappear throughout the paper, so it is convenient to have

[2] It is implicit here that the operation of Weakening should always be such as to induce an appropriate bracketting. In practice, inclusion of Weakening and/or Contraction only really makes sense in systems with Association and Permutation.

[3] Strictly, LP and linear logic differ. LP requires a non-empty antecedent whereas linear logic does not.

[4] See van Benthem (1991) and Moortgat (1992) for discussion of substructural hierarchy

distinguished notations for the connectives in these systems, shown in (4).

(4)	Association	Permutation	Connectives			
	−	−	$\odot$	$\obslash$	$\oslash$	(non-associative Lambek)
	−	+	$\ominus$	$\xrightarrow{\ominus}$	$\xleftarrow{\ominus}$	
	+	−	$\bullet$	$\backslash$	$/$	(associative Lambek)
	+	+	$\otimes$	$\multimap$	$\multimapinv$	(linear logic)

3 Structural modalities

Structural modalities are unary operators that allow for controlled involvement of structural rules, rules that are otherwise not freely available in a particular system.[5] For example, a version of particular structural rule might state that the rule can only apply if one, or perhaps all, of the formulas directly affected by the rule's use are marked with a given structural modality. Let us consider a linguistic case to illustrate this possibility.

The associative Lambek calculus has received considerable attention as a possible grammatical formalism. Lexical assignment of directional implicational formulas readily allows information of subcategorisation and word order to be encoded. The (usually implicit) availability of Association allows for considerable flexibility in derivation, flexibility which has provided a basis for accounts of various phenomena, including non-constituent coordination and extraction.

For example, a string such as *John spoke to* may be derived as s/np, effectively a 'sentence missing a NP on its right', so that a relative clause such as *whom John spoke to* may be derived by lexically assigning *whom* the type rel/(s/np). Such a treatment of extraction is too limited however. The Lambek implications can only represent an 'X missing Y' where Y is *left* or *right peripheral* in X, i.e. Y\X or X/Y. Thus, this approach does not extend to cases of extraction from non-peripheral position, as in e.g. *whom John gave _ a bagel.* Morrill *et al.* (1990) handle this problem using a *permutation* structural modality. Consider, for example, a unary operator $\triangle$ with the following inference rules (where $\triangle\Gamma$ indicates a configuration where all formulas are of the form $\triangle$X):

(5)
$$\begin{array}{c}\triangle\Gamma \Rightarrow \mathrm{A}\\ \hline \triangle\Gamma \Rightarrow \triangle\mathrm{A}\end{array}[\triangle\mathrm{R}] \qquad \begin{array}{c}\Gamma[\mathrm{B}] \Rightarrow \mathrm{A}\\ \hline \Gamma[\triangle\mathrm{B}] \Rightarrow \mathrm{A}\end{array}[\triangle\mathrm{L}] \qquad \begin{array}{c}\Gamma[(\triangle\mathrm{B},\ \mathrm{C})] \Rightarrow \mathrm{A}\\ \hline\hline \Gamma[(\mathrm{C},\ \triangle\mathrm{B})] \Rightarrow \mathrm{A}\end{array}[\triangle\mathrm{P}]$$

The left and right rules for $\triangle$ are the same as for necessity in S4. The restricted permutation

[5] The original structural modalities are linear logic's 'exponentials' (Girard, 1987). For categorial work on linguistic uses of structural modalities, and on their proof theory and semantics, see Morrill *et al.* (1990), Barry *et al.* (1991), Moortgat (1992), Venema (1993a, 1993b), Versmissen (1992, 1993), Hollenberg (1991).

rule [$\triangle$P] allows that any formula of the form $\triangle$X may permute freely, i.e. undermining linear order for such an assumption. The left rule [$\triangle$L] allows that such a marking may be freely discarded. A formula such as s/($\triangle$np) corresponds to a sentence missing a NP *at some position*, and so a type assignment rel/(s/($\triangle$np)) for *whom* allows for extraction from both peripheral and non-peripheral location.

Other structural modalities may be used to give controlled reintroduction of other structural rules. A single modality may even be used to readmit more than one structural rule. Linear logic's *exponentials* ! and ? give controlled reintroduction of Contraction and Weakening. In a non-associative system, an associativity modality could be used to reintroduce Associativity.

Structural modalities allow that stronger logics may be embedded within weaker ones, via *embedding translations*. For example, the translation (6) embeds a fragment of linear logic within associative Lambek calculus, so that $\Gamma \Rightarrow A$ is a theorem of the former *iff* $\triangle|\Gamma| \Rightarrow |A|$ is a theorem of the latter (see Došen, 1990, for discussion). Another example is that intuitionistic logic can be embedded within (intuistionistic) linear logic, via an embedding translation using the exponentials !, ?. Of course, what such an embedding shows is that when the appropriate structural modalities are added to a weaker logic, the resulting combined system is anything but 'weaker' than the related 'stronger' logic.

(6)
$$\begin{array}{lll} |A| & := & A \qquad\qquad (A \text{ atomic}) \\ |(A \otimes B)| & := & ((\triangle|A|) \bullet (\triangle|B|)) \\ |(B \multimap A)| & := & ((\triangle|B|) \backslash |A|) \\ |(A \multimapinv B)| & := & (|A| \,/\, (\triangle|B|)) \end{array}$$

4 Relations between substructural levels

Consider again the case of Lambek calculus and linear logic and the embedding translation (6). The successful embedding shows that linear logic may be 'represented' within the system of Lambek calculus *plus* permutation modality ('LC$\triangle$'). Of course, it is also true that pure Lambek calculus may also be 'represented' within LC$\triangle$, trivially so indeed. In a sense then, LC$\triangle$ is a system where we can observe coexistence of linear logic and Lambek calculus, or rather perhaps coexistence of 'images' of these systems (in one case, an image obtained by a trivial mapping). Furthermore, the involvement of the permutation modality $\triangle$ allows us to observe relations between the two levels.

Consider, for example, a linear logic formula such as $X \otimes Y$ and its translation under (6) $(\triangle X) \bullet (\triangle Y)$ (or strictly its translation assuming X,Y are atomic — the difference doesn't matter for the immediate point). The translated formula exhibits the following interderivability: $(\triangle X) \bullet (\triangle Y) \Leftrightarrow (\triangle Y) \bullet (\triangle X)$, akin to $X \otimes Y \Leftrightarrow Y \otimes X$ for the original formula. Furthermore, the $\triangle$s may be 'dropped', i.e. $(\triangle X) \bullet (\triangle Y) \Rightarrow X \bullet Y$, and $(\triangle X) \bullet (\triangle Y) \Rightarrow Y \bullet X$. In this context then, a formula such as $X \otimes Y$ may be viewed as telling us that X,Y may be related to each other in either order. The transformation to a term $X\bullet Y$ or $Y\bullet X$ may be viewed as merely the step of selecting one of the permitted orders. Hence, there appears to be a natural relation between $X\otimes Y$ and $X\bullet Y$, as if $X\otimes Y \Rightarrow X\bullet Y$ were a theorem of some mixed logic.

Consider next a linear implicational such as $X\circ\!\!-Y$, whose translation might be $X/(\triangle Y)$. Here it is notable that $LC\triangle$ allows the transition $X/Y \Rightarrow X/(\triangle Y)$ (as well as $X/Y \Rightarrow X\backslash(\triangle Y)$). This transformation may again be viewed as a step of selecting a particular one of a range of options allowed by a formula. Thus, X/Y merely requires an argument Y to its right. One particular case of a Y that can 'appear to the right' is a Y that can appear anywhere, i.e. $\triangle Y$. In this case then, there appears a natural relation between $X\circ\!\!-Y$ and X/Y, as if $X/Y \Rightarrow X\circ\!\!-Y$ were a theorem.

The above discussion raises the interesting possibility of having a logic where substructurally different connectives coexist, and where transformations illustrating 'natural relations' between the levels, such as those discussed above, are indeed *theorems*. In particular, we might expect that for systems with products $*$ and $\circ$, where the former system has greater freedom for resource usage, the following transformations should be allowed:

$$X\circ Y \Rightarrow X*Y$$
$$X\overset{*}{\leftarrow}Y \Rightarrow X\overset{\circ}{\leftarrow}Y$$

Such an approach is outlined in the next section.

5 A hybrid substructural system

Let us consider a system in which connectives that are subject to different structural rules coexist.[6] Again, Γ of any $\Gamma \Rightarrow A$ is a binarily bracketted sequence, but now there is a different bracket pair corresponding to each system specific product operator, e.g. for the linear subsystem, with multiplicative $\otimes$, there is a bracket pair $(.,.)^{\otimes}$. In the following statement of the rules, the symbol $\circ$ is used in place of any of the product constructors. Note that the operational rules for the connectives of the different systems are the same. However, the different systems may differ regarding the structural rules that apply for them (see the side conditions on the structural rules in (9)). For present purposes we consider only systems arising by choices from the structural rules Association and Permutation, although the approach readily generalises.

(7) $$A \Rightarrow A \quad \text{(id)} \qquad\qquad \frac{\Delta \Rightarrow B \quad \Gamma[B] \Rightarrow A}{\Gamma[\Delta] \Rightarrow A}\text{[cut]}$$

[6] Oehrle & Zhang (1989) and Morrill (1990) propose systems in which there is coexistence of different substructural levels — associative and non-associative systems in both cases. However, there is no interrelation between the different levels, in contrast to the hybrid system to be described here.

(8)

$$\frac{(\mathrm{B},\ \Gamma)^{\circ} \Rightarrow \mathrm{A}}{\Gamma \Rightarrow \mathrm{B}\overset{\circ}{\rightarrow}\mathrm{A}}[\overset{\circ}{\rightarrow}\mathrm{R}] \qquad \frac{\Delta \Rightarrow \mathrm{C} \quad \Gamma[\mathrm{B}] \Rightarrow \mathrm{A}}{\Gamma[(\Delta,\ \mathrm{C}\overset{\circ}{\rightarrow}\mathrm{B})^{\circ}] \Rightarrow \mathrm{A}}[\overset{\circ}{\rightarrow}\mathrm{L}]$$

$$\frac{(\Gamma,\ \mathrm{B})^{\circ} \Rightarrow \mathrm{A}}{\Gamma \Rightarrow \mathrm{A}\overset{\circ}{\leftarrow}\mathrm{B}}[\overset{\circ}{\leftarrow}\mathrm{R}] \qquad \frac{\Delta \Rightarrow \mathrm{C} \quad \Gamma[\mathrm{B}] \Rightarrow \mathrm{A}}{\Gamma[(\mathrm{B}\overset{\circ}{\leftarrow}\mathrm{C},\ \Delta)^{\circ}] \Rightarrow \mathrm{A}}[\overset{\circ}{\leftarrow}\mathrm{L}]$$

$$\frac{\Gamma \Rightarrow \mathrm{A} \quad \Delta \Rightarrow \mathrm{B}}{(\Gamma,\ \Delta)^{\circ} \Rightarrow \mathrm{A}\circ\mathrm{B}}[\circ\mathrm{R}] \qquad \frac{\Gamma[(\mathrm{B},\ \mathrm{C})^{\circ}] \Rightarrow \mathrm{A}}{\Gamma[\mathrm{B}\circ\mathrm{C}] \Rightarrow \mathrm{A}}[\circ\mathrm{L}]$$

(9)

$$\begin{array}{c}\Gamma[(\mathrm{B},\ (\mathrm{C},\ \mathrm{D})^{\circ})^{\circ}] \Rightarrow \mathrm{A} \\ \hline\hline \Gamma[((\mathrm{B},\ \mathrm{C})^{\circ},\ \mathrm{D})^{\circ}] \Rightarrow \mathrm{A}\end{array}[\mathrm{A}] \qquad \Big\{\circ \ \in \ \{\bullet, \otimes\}$$

$$\frac{\Gamma[(\mathrm{B},\ \mathrm{C})^{\circ}] \Rightarrow \mathrm{A}}{\Gamma[(\mathrm{C},\ \mathrm{B})^{\circ}] \Rightarrow \mathrm{A}}[\mathrm{P}] \qquad \Big\{\circ \ \in \ \{\ominus, \otimes\}$$

(10)

$$\frac{\Gamma[(\mathrm{B},\ \mathrm{C})^{\circ'}] \Rightarrow \mathrm{A}}{\Gamma[(\mathrm{B},\ \mathrm{C})^{\circ}] \Rightarrow \mathrm{A}}[<_S] \qquad \Big\{\ \circ' \ <_S \ \circ$$

With only the rules (8-9), we would have a system where different substructural levels merely coexist, with no interrelation. Such interrelation is effected by the rule (10). Note that the rule's side condition employs a relation $<_S$, for relating different levels in term of 'degree of structural freedom'. In particular, $\circ' <_S \circ$ just in case $\circ'$ and $\circ$ are corresponding product operators where the latter's system exhibits greater freedom for resource usage than the former's.[7] Thus, the pairs $\langle\circ, \circ'\rangle$ such that $\circ, \circ' \in \{\odot, \ominus, \bullet, \otimes\}$ are:

$$\{\langle\odot, \ominus\rangle\,,\ \langle\odot, \bullet\rangle\,,\ \langle\odot, \otimes\rangle\,,\ \langle\ominus, \otimes\rangle\,,\ \langle\bullet, \otimes\rangle\}$$

The rule $[<_S]$ allows a bracket pair of one system to be replaced in the conclusion sequent by the bracket of another system, just in case the latter's system exhibits greater freedom of resource usage.

Let us examine some consequences of linking substructural levels in this way. In the previous section, I suggested that $\mathrm{A}\otimes\mathrm{B} \Rightarrow \mathrm{A}\bullet\mathrm{B}$ would be an appropriate theorem in a hybrid system.

[7] The phrasing 'corresponding product operators' is unnecessary for the example systems $\{\odot, \ominus, \bullet, \otimes\}$ we have addressed so far, but is important with regard to 'dependency systems' (Moortgat & Morrill, 1991), to be discussed later in the paper. In dependency systems, a level typically has two complementary product operators, one left-biased and one right-biased. In a hybrid framework, we may have more than one such biased level, and interrelation of products under $<_S$ should preverve bias, e.g. relating the left-biased connective of one level to the left-biased connective of the other level, etc. Note that for our non-biased example systems $(\{\odot, \ominus, \bullet, \otimes\})$, we find $\circ' <_S \circ$ *iff* the structural rules of $\circ'$ are a subset of those for $\circ$. Again, the situation is not quite so simple when we consider dependency systems.

This transition is derivable in the present approach, as shown in (11).

(11) $$\dfrac{\dfrac{\dfrac{B \Rightarrow B \qquad A \Rightarrow A}{(A,\ B)^{\bullet} \Rightarrow A{\bullet}B}\,[{\bullet}\mathrm{R}]}{(A,\ B)^{\otimes} \Rightarrow A{\bullet}B}\,[<_S]}{A{\otimes}B \Rightarrow A{\bullet}B}\,[{\otimes}\mathrm{L}]$$

(12) $$\dfrac{\dfrac{\dfrac{A \Rightarrow A \qquad B \Rightarrow B}{(A/B,\ B)^{\bullet} \Rightarrow A}\,[/\mathrm{L}]}{(A/B,\ B)^{\otimes} \Rightarrow A}\,[<_S]}{A/B \Rightarrow A{\multimapinv}B}\,[{\multimapinv}\mathrm{R}]$$

Note that the converse transition is *not* derivable, i.e. since the converse substitution of brackets under $[<_S]$ to that in (11) is not allowed. The transformation $A/B \Rightarrow A{\multimapinv}B$, also discussed earlier, can be derived as in (12). More generally, transformations similar to (11) and (12) may be derived for the connectives of any two appropriately related subsystems, e.g. $A{\odot}B \Rightarrow A{\ominus}B$ and $A{\oslash}B \Rightarrow A/B$.

Let us consider some other theorems of this hybrid system. Note that although the composition (13a) is not allowed within a non-associative level, the same types may be composed, provided the result is in an associative level, as in (13b), proven in (14). (The inference step marked $[<_S{*}]$ corresponds to multiple uses of $[<_S]$.) Likewise, two counterdirectional arguments cannot be reordered purely within a non-associative system, as in (13c), but can if the result type is in an associative system, e.g. (13d), proven in (15).

(13) a. * $(A{\oslash}B,\ B{\oslash}C)^{\odot} \Rightarrow A{\oslash}C$

b. $(A{\oslash}B,\ B{\oslash}C)^{\bullet} \Rightarrow A/C$

c. * $(B{\obslash}A){\oslash}C \Rightarrow B{\obslash}(A{\oslash}C)$

d. $(B{\obslash}A){\oslash}C \Rightarrow B\backslash(A/C)$

(14) $$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{C \Rightarrow C \qquad B \Rightarrow B}{(B{\oslash}C,\ C)^{\odot} \Rightarrow B}\,[{\oslash}\mathrm{L}] \qquad A \Rightarrow A}{(A{\oslash}B,\ (B{\oslash}C,\ C)^{\odot})^{\odot} \Rightarrow A}\,[{\oslash}\mathrm{L}]}{(A{\oslash}B,\ (B{\oslash}C,\ C)^{\bullet})^{\bullet} \Rightarrow A}\,[<_S{*}]}{((A{\oslash}B,\ B{\oslash}C)^{\bullet},\ C)^{\bullet} \Rightarrow A}\,[\mathrm{A}]}{(A{\oslash}B,\ B{\oslash}C)^{\bullet} \Rightarrow A/C}\,[/\mathrm{R}]$$

(15) $$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{C \Rightarrow C \qquad \dfrac{B \Rightarrow B \qquad A \Rightarrow A}{(B,\ B{\obslash}A)^{\odot} \Rightarrow A}\,[{\obslash}\mathrm{L}]}{(B,\ ((B{\obslash}A){\oslash}C,\ C)^{\odot})^{\odot} \Rightarrow A}\,[{\oslash}\mathrm{L}]}{(B,\ ((B{\obslash}A){\oslash}C,\ C)^{\bullet})^{\bullet} \Rightarrow A}\,[<_S{*}]}{((B,\ (B{\obslash}A){\oslash}C)^{\bullet},\ C)^{\bullet} \Rightarrow A}\,[\mathrm{A}]}{(B,\ (B{\obslash}A){\oslash}C)^{\bullet} \Rightarrow A/C}\,[/\mathrm{R}]}{(B{\obslash}A){\oslash}C \Rightarrow B\backslash(A/C)}\,[\backslash\mathrm{R}]$$

Although associative Lambek calculus allows composition, directionally mixed compositions such as (16a) are not permitted. However, such types may be composed provided that the result is in the permutative linear subsystem, as in (16b), which has proof (17).

(16) a. * $(A/B,\ C\backslash B)^{\bullet} \Rightarrow C\backslash A$

b. $(A/B,\ C\backslash B)^{\otimes} \Rightarrow A\circ\!\!-C$

c. * $(A/B,\ C\backslash B)^{\bullet} \Rightarrow A\circ\!\!-C$

d. $(C\backslash B,\ A/B)^{\otimes} \Rightarrow A\circ\!\!-C$

(17)
$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{C \Rightarrow C \qquad B \Rightarrow B}{(C,\ C\backslash B)^{\bullet} \Rightarrow B}\,[\backslash L] \qquad A \Rightarrow A}{(A/B,\ (C,\ C\backslash B)^{\bullet})^{\bullet} \Rightarrow A}\,[/L]}{(A/B,\ (C,\ C\backslash B)^{\otimes})^{\otimes} \Rightarrow A}\,[<_{S}*]}{(A/B,\ (C\backslash B,\ C)^{\otimes})^{\otimes} \Rightarrow A}\,[P]}{(A/B,\ C\backslash B)^{\otimes},\ C)^{\otimes} \Rightarrow A}\,[A]}{(A/B,\ C\backslash B)^{\otimes} \Rightarrow A\circ\!\!-C}\,[\circ\!\!-R]$$

This example is suggestive that linear implication in a hybrid system might allow treatment of non-peripheral extraction, e.g. with relative pronouns bearing type rel/(s∘–np), eliminating the need for a permutative structural modality. However, the same example also brings up a potential problem. Note that the two implicationals in the antecedent of (16b) are configured with the linear brackets $(.,.)^{\otimes}$. This is necessarily so: the similar sequent in (16c), where the associative Lambek brackets appear, is not a theorem. Obviously then, the sequent (16d), where the two implicationals appear 'out of order', is also a theorem. If word order were to be determined from the linear order of formulas in a configuration, then the approach would be unable to distinguish (e.g.) a relative clause from some of its ungrammatical permutations. The method for determining the word order consequences of proofs depends on augmenting proofs with a system of term assignment, introduced in the next section.

6 Term assignment for hybrid substructural systems

I will next describe a system of term assignment for the logic introduced in the previous section, by which the formulas in a proof are associated with lambda terms in accordance with the well known Curry–Howard interpretation of proofs (Howard, 1969). Antecedent formulas are associated with variables. The terms associated with succedent formulas record the *natural deduction* structure of the dominating subproof.

(18) $A : v \Rightarrow A : v$ (id) $\qquad \dfrac{\Delta \Rightarrow B : b \qquad \Gamma[B : v] \Rightarrow A : a}{\Gamma[\Delta] \Rightarrow A : a[b/v]}\,[\text{cut}]$

(19)

$$\frac{(\mathrm{B}:v,\ \Gamma)^{\circ} \Rightarrow \mathrm{A}:a}{\Gamma \Rightarrow \mathrm{B}\overset{\circ}{\rightarrow}\mathrm{A}:[\underset{\circ}{\rightarrow}]v.a}[\overset{\circ}{\rightarrow}\mathrm{R}] \qquad \frac{\Delta \Rightarrow \mathrm{C}:c \qquad \Gamma[\mathrm{B}:v] \Rightarrow \mathrm{A}:a}{\Gamma[(\Delta,\ \mathrm{C}\overset{\circ}{\rightarrow}\mathrm{B}:w)^{\circ}] \Rightarrow \mathrm{A}:a[(c\underset{\circ}{\rightarrow}w)/v]}[\overset{\circ}{\rightarrow}\mathrm{L}]$$

$$\frac{(\Gamma,\ \mathrm{B}:v)^{\circ} \Rightarrow \mathrm{A}:a}{\Gamma \Rightarrow \mathrm{A}\overset{\circ}{\leftarrow}\mathrm{B}:[\underset{\circ}{\leftarrow}]v.a}[\overset{\circ}{\leftarrow}\mathrm{R}] \qquad \frac{\Delta \Rightarrow \mathrm{C}:c \qquad \Gamma[\mathrm{B}:v] \Rightarrow \mathrm{A}:a}{\Gamma[(\mathrm{B}\overset{\circ}{\leftarrow}\mathrm{C}:w,\ \Delta)^{\circ}] \Rightarrow \mathrm{A}:a[(w\underset{\circ}{\leftarrow}c)/v]}[\overset{\circ}{\leftarrow}\mathrm{L}]$$

$$\frac{\Gamma \Rightarrow \mathrm{A}:a \qquad \Delta \Rightarrow \mathrm{B}:b}{(\Gamma,\ \Delta)^{\circ} \Rightarrow \mathrm{A}{\circ}\mathrm{B}:\langle a,b\rangle^{\circ}}[\circ\mathrm{R}] \qquad \frac{\Gamma[(\mathrm{B}:v,\ \mathrm{C}:w)^{\circ}] \Rightarrow \mathrm{A}:a}{\Gamma[\mathrm{B}{\circ}\mathrm{C}:x] \Rightarrow \mathrm{A}:[x/v{\circ}w].a}[\circ\mathrm{L}]$$

(20)

$$\begin{array}{c}\Gamma[(\mathrm{B}:b,\ (\mathrm{C}:c,\ \mathrm{D}:d)^{\circ})^{\circ}] \Rightarrow \mathrm{A}:a \\ \hline\hline \Gamma[((\mathrm{B}:b,\ \mathrm{C}:c)^{\circ},\ \mathrm{D}:d)^{\circ}] \Rightarrow \mathrm{A}:a\end{array}[\mathrm{A}] \qquad \Big\{\ \circ\ \in\ \{\bullet,\otimes\}$$

$$\frac{\Gamma[(\mathrm{B}:b,\ \mathrm{C}:c)^{\circ}] \Rightarrow \mathrm{A}:a}{\Gamma[(\mathrm{C}:c,\ \mathrm{B}:b)^{\circ}] \Rightarrow \mathrm{A}:a}[\mathrm{P}] \qquad \Big\{\ \circ\ \in\ \{\ominus,\otimes\}$$

(21)

$$\frac{\Gamma[(\mathrm{B}:b,\ \mathrm{C}:c)^{\circ'}] \Rightarrow \mathrm{A}:a}{\Gamma[(\mathrm{B}:b,\ \mathrm{C}:c)^{\circ}] \Rightarrow \mathrm{A}:a}[<_S] \qquad \Big\{\ \circ'\ <_S\ \circ$$

For implicational connectives, left inferences (corresponding to natural deduction elimination inferences) are interpreted via functional application, whilst right inferences (c.f. natural deduction introduction inferences) are interpreted via functional abstraction. Note, however, that a different abstraction and application operator is required for each distinct implicational connective, so that terms fully record the proof structure.[8] For an implication $\overset{\circ}{\leftarrow}$, for example, the corresponding application operator is notated as $\underset{\circ}{\leftarrow}$, e.g. $a\underset{\circ}{\leftarrow}b$ is a applied to b in the relevant mode. Note carefully that in $a\underset{\circ}{\rightarrow}b$ the term b is applied to a, not *vice versa.* The corresponding abstraction operator is notated as $[\underset{\circ}{\leftarrow}]$, e.g. $[\underset{\circ}{\leftarrow}]v.a$ represents abstraction over v in a. Left inferences for product operators are interpreted via system specific pairing, e.g. $\langle a,b\rangle^{\circ}$. Note the labelling of the corresponding right rule. A term such as e.g. $[b/v{\circ}w].a$ implicitly represents the substitution of b for $v{+}w$ in a.[9] Observe that the structural and $[<_S]$ rules have no term assignment effect.

[8] Buszkowski (1987) specifies a term system for implicational associative Lambek calculus, where the typing of functional expressions encodes directional distinctions, and where two directional modes of application and abstraction are distinguished. A straightforwardly identifiable subclass of the possible expressions of this system correspond to correct deductions of implicational associative Lambek calculus. Wansing (1990) discusses 'formulas-as-types' term systems for a range of substructural logics.

[9] This labelling operator is essentially just a compact notation for an operator used with the linear multiplicative $\otimes$ by Benton *et al.* (1992) in their system of term assignment for linear logic. Their operator is **let** _ **be** _$\otimes$_ **in** _, giving terms such as e.g. **let** b **be** $v{\otimes}w$ **in** a, parallel to $[b/v{\otimes}w].a$ for us. It is important to note that the operator serves to bind the variables v and w in the subterm a. Thus, $\mathrm{FV}([b/v{\circ}w].a) = (\mathrm{FV}(a) - \{v,w\}) \cup \mathrm{FV}(b)$ (where FV returns the set of variables free in a term).

7 Word order, semantics and term labelling

Consider again the type combinations (16b,d), repeated here as (22a,b), which illustrate the inadequacy of determining word order consequences of combinations from the order of formulas in configurations. For example, infering by this method that either order is possible for the two antecedent formulas in (22a,b), we appear to have lost the word order import of the directional connective of A/B, indicating that this functor requires its complement to appear to its right.

(22) a. $(\mathrm{A}/\mathrm{B},\ \mathrm{C}\backslash\mathrm{B})^{\otimes} \Rightarrow \mathrm{A}\mathbin{\circ\!-}\mathrm{C}$

c. $(\mathrm{C}\backslash\mathrm{B},\ \mathrm{A}/\mathrm{B})^{\otimes} \Rightarrow \mathrm{A}\mathbin{\circ\!-}\mathrm{C}$

(23) $[\overset{\leftarrow}{\otimes}]z.\, x \overset{\leftarrow}{\bullet} (z \overset{\rightarrow}{\bullet} y)$

Note that (22a) would receive proof term (23) under proof (17), and (22b) would receive the *same* proof term under a proof differing from (17) only by a semantically non-potent use of [P]. This common proof term for the two related theorems implicitly encodes the linear order import of the A/B functor's connective. Thus, since $\bullet$ is a non-permutative operator, an applicative term of the form $a \overset{\leftarrow}{\bullet} b$ indicates that the 'orderable elements' of a precede those of b. For the subterm $x \overset{\leftarrow}{\bullet} (z \overset{\rightarrow}{\bullet} y)$, this suggests that z precedes y, and that x precedes z and y. The abstraction over z in the complete term (23) would seem to discount z as an 'orderable element', leaving just the result that x precedes y, i.e. that the A/B functor precedes the C\B functor, as we would expect.

Since $\otimes$ is permutative, we expect less ordering information to be available from a term such as $a \overset{\leftarrow}{\otimes} b$, but this does not mean that no consequences follow for the relative ordering of the 'orderable elements' of a, b. For example, although the position of x relative to y, z in $x \overset{\leftarrow}{\otimes} (y \overset{\leftarrow}{\odot} z)$ is not completely fixed, the non-associativity of $\odot$ implies an 'integrity' for y, z in $(y \overset{\leftarrow}{\odot} z)$ such that we would not expect $y \prec x \prec z$ to be a possible order. These observations suggest that the method for determining the linear order consequences of proofs must be sensitive to the specific modes of structuring and their properties.

The first step in extracting word order information from proof terms involves deriving a second term called a *yield term*. For each syntactic product operator, we require a corresponding 'yield' operator, which will be notated identically. The set of possible yield terms Y is derived by closing the set $(\mathtt{VAR} \cup \mathtt{ATOM})$ under the yield operators ($\mathtt{VAR}$ the set of variables, $\mathtt{ATOM}$ the set of atoms).[10] The yield term is generated by a procedure $\mathcal{F}$, which is defined by induction on the structure of proof terms as in (24) (where $\circ$ is used in place of any conjunctive operator, and $[\overset{\leftrightarrow}{\circ}]$ and $\overset{\leftrightarrow}{\circ}$ are used to represent abstraction and application where

[10] Note that atoms are included here to allow that word order may be derived *after* lexical substitution.

directionality doesn't matter).

(24)
$$\begin{array}{llll}
\Omega(a) & = & a, & \text{where } \ a \in \ (\mathtt{VAR} \cup \mathtt{ATOM}) \\
\Omega(a \overset{\leftrightarrow}{\circ} b) & = & p \circ q, & \text{where } \ \Omega(a) = p,\ \Omega(b) = q \\
\Omega(\langle a, b \rangle^{\circ}) & = & p \circ q, & \text{where } \ \Omega(a) = p,\ \Omega(b) = q \\
\Omega([\overset{\leftrightarrow}{\circ}]v.a) & = & p[q], & \text{where } \ \Omega(a) = p[v{\circ'}q] \ \ \mathit{or} \ \ \Omega(a) = p[q{\circ'}v] \\
\Omega([b/v{\circ}w].a) & = & p[q], & \text{where } \ \Omega(a) = p[v{\circ'}w],\ \Omega(b) = q
\end{array}$$

The cases for applicative and pair terms merely put together the yield terms for the subexpressions. The case for abstraction deletes the variable from within the yield term of the subexpression. The final case, for product introduction labellings, is the most problematic. It is not always true that $\mathcal{F}(a) = p[v{\circ'}w]$ for the subterm a, in which case the procedure fails, and the word order consequences of the term cannot be projected. I will return to this problem shortly.

A yield term returns the orderable elements of a proof term structured in accordance with their original manner of combination. E.g. $x \overset{\leftarrow}{\otimes} (y \overset{\leftarrow}{\odot} z)$ gives a yield term $x{\otimes}(y{\odot}z)$, and $[\overset{\leftarrow}{\otimes}]z.\, x \overset{\leftarrow}{\bullet} (z \overset{\rightarrow}{\bullet} y)$ gives $x{\bullet}y$. Let $\leadsto$ represent restructuring of yield terms in ways appropriate to the different operators. E.g. since the syntactic product $\otimes$ is subject to [P], 'commutation' restructurings are allowed for the yield term operator $\otimes$, i.e. for yield terms a, b, $a \leadsto b$ if $a = p[q \otimes r]$ and $b = p[r \otimes q]$. Likewise for other operators and structural rules. Possible linear orders allowed by a proof term can be simply 'read off' the variants of its yield term under restructuring, e.g. the proof term $x \overset{\leftarrow}{\otimes} (y \overset{\leftarrow}{\odot} z)$ gives orders $x{\cdot}y{\cdot}z$ and $y{\cdot}z{\cdot}x$, since $\mathcal{F}(x \overset{\leftarrow}{\otimes} (y \overset{\leftarrow}{\odot} z))=$ $x{\otimes}(y{\odot}z)$, and $x{\otimes}(y{\odot}z) \leadsto (y{\odot}z){\otimes}x$.

Consider next the theorem (25) and its two alternative proofs (26) and (27), which assign the readings in (28a,b), respectively.

(25) $\quad (\mathrm{X/Y,\ Y})^{\otimes} \Rightarrow \mathrm{X}$

(26)
$$\dfrac{\dfrac{\dfrac{\dfrac{\mathrm{Y} \Rightarrow \mathrm{Y} \qquad \mathrm{X} \Rightarrow \mathrm{X}}{(\mathrm{X/Y,\ Y})^{\bullet} \Rightarrow \mathrm{X}}\,[/\mathrm{L}]}{(\mathrm{X/Y,\ Y})^{\otimes} \Rightarrow \mathrm{X}}\,[{<}_S]}{\mathrm{X/Y} \Rightarrow \mathrm{X}{\circ\!\!-}\mathrm{Y}}\,[{\circ\!\!-}\mathrm{R}] \qquad \dfrac{\mathrm{Y} \Rightarrow \mathrm{Y} \qquad \mathrm{X} \Rightarrow \mathrm{X}}{(\mathrm{X}{\circ\!\!-}\mathrm{Y},\ \mathrm{Y})^{\otimes} \Rightarrow \mathrm{X}}\,[{\circ\!\!-}\mathrm{L}]}{(\mathrm{X/Y,\ Y})^{\otimes} \Rightarrow \mathrm{X}}\,[\mathrm{cut}]$$

(27)
$$\dfrac{\dfrac{\mathrm{Y} \Rightarrow \mathrm{Y} \qquad \mathrm{X} \Rightarrow \mathrm{X}}{(\mathrm{X/Y,\ Y})^{\bullet} \Rightarrow \mathrm{X}}\,[/\mathrm{L}]}{(\mathrm{X/Y,\ Y})^{\otimes} \Rightarrow \mathrm{X}}\,[{<}_S]$$

(28) a. $\quad ([\overset{\leftarrow}{\otimes}]v.\, x \overset{\leftarrow}{\bullet} v) \overset{\leftarrow}{\otimes} y$

b. $\quad x \overset{\leftarrow}{\bullet} y$

Despite the directional connective of (26), (28a) gives a yield term $x{\otimes}y$ leaving x and y unordered (i.e. allows both orders $x{\cdot}y$ and $y{\cdot}x$). This example suggests that order should be determined from the *normal form* of proof terms. Thus, (28a) normalises to (28b), from which order can be appropriately determined ($x{\cdot}y$). The contraction rules for normalising terms of this systems are given in (29).[11]

(29)
$$\begin{array}{lcl}
([\overset{\leftarrow}{\circ}]v.a)\overset{\leftarrow}{\circ}b & \overset{\beta}{\rightsquigarrow} & a[b/v] \\
b\overset{\rightarrow}{\circ}([\overset{\rightarrow}{\circ}]v.a) & \overset{\beta}{\rightsquigarrow} & a[b/v] \\
[\langle b,c\rangle^{\circ}/v{\circ}w].a & \overset{\beta}{\rightsquigarrow} & a[b/v,c/w] \\
([b/v{\circ}'w].a)\overset{\leftarrow}{\circ}c & \overset{c}{\rightsquigarrow} & [b/v{\circ}'w].(a\overset{\leftarrow}{\circ}c) \\
c\overset{\rightarrow}{\circ}([b/v{\circ}'w].a) & \overset{c}{\rightsquigarrow} & [b/v{\circ}'w].(c\overset{\rightarrow}{\circ}a)
\end{array}$$

A function Ω, which returns the set of possible orderings allowed by a proof term may be defined as in (30) (where $\beta(a)$ returns the normal form of a, and $\mathtt{STRING}(a)$ a string of the atomic elements in the yield term a, in the order that they appear therein).

(30)
$$\Omega(a) = \{\, \mathtt{STRING}(p) \mid \mathcal{F}(\beta(a)) \rightsquigarrow p \,\}$$

We noted above that $\mathcal{F}$ fails to produce a yield term for some proof terms. For example, the proof term $(x\overset{\leftarrow}{\odot}y)\overset{\leftarrow}{\odot}z$ has yield term $(x{\odot}y){\odot}z$, which contains no subterm of the form $y{\circ}'z$. Hence, the proof term $[w/y{\bullet}z].(x\overset{\leftarrow}{\odot}y)\overset{\leftarrow}{\odot}z$ (corresponding to a proof of $((\mathrm{X}{\not\circ}\mathrm{Z}){\not\circ}\mathrm{Y},\ \mathrm{Y}{\bullet}\mathrm{Z})^{\bullet} \Rightarrow \mathrm{X}$) has no yield term. I will take the position that linear order determinations should be of *word order*, assessed only after substitution of lexical strings. The assumption is that lexical string terms may be complex expressions of the term labelling system, and that lexical assignments will specify string terms that provide 'appropriate resources' to allow a proper word order determination to be made. For example, a lexical assignment bearing a product type would have a pair string term. Substituting this for (e.g.) w in $[w/y{\bullet}z].(x\overset{\leftarrow}{\odot}y)\overset{\leftarrow}{\odot}z$ would give a term that simplified with normalisation to a result that did allow a full word order determination.

Allowing complex string terms constructed using the operators of the term labelling system in lexical assignments raises various possibilities concerning lexical encoding of partial derivations and string term encoding of word order relevant information. A simple (unrealistic) example of this possibility is that we might assign a word $\widehat{\mathrm{w}}$ a type such as $\mathrm{s}{\circ}\!{-}\mathrm{np}$, whose connective does not specify an order relative to its argument, but provide a string term such

[11] The correctness of the two commutative conversion rules (notated $\overset{c}{\rightsquigarrow}$) requires that $v, w \notin \mathtt{FV}(c)$, i.e. to avoid changes of meaning due to accidental binding of v, w occurrences in c. Given the linear usage characteristic of the systems under consideration, all proof terms are such that no problem arises in relation to this requirement.

as $[\overset{\leftarrow}{\otimes}]v.\,\widehat{\mathrm{w}}\overset{\leftarrow}{\bullet}v$, which *does* encode information of the relative order of $\widehat{\mathrm{w}}$ and its argument, which could emerge when word order was determined via normalisation and $\mathcal{F}$. Note that this lexical string term lexically encodes a derivation corresponding to the transformation: $\mathrm{s/np} \Rightarrow \mathrm{s}{\circ}\!\!-\mathrm{np}$.

Proof terms are also useful for determining the natural language semantic consequences of type combinations. However, the fine-grained distinctions between different forms of functional abstraction and application etc. seems somewhat inappropriate for linguistic semantics. I assume a procedure $\overset{\alpha}{\leadsto}$ that transforms a proof term to a a simpler form, with only a single form of abstraction (λ) and application notated by juxtaposition (with f functor and a argument in any fa). E.g.: $[\overset{\leftarrow}{\otimes}]z.\,x\overset{\leftarrow}{\bullet}(z\overset{\rightarrow}{\bullet}y) \quad \overset{\alpha}{\leadsto} \quad \lambda z.x(yz)$

8 Algebraic semantics for categorial logics

We can define semantics for various categorial logics which include a 'product' conjuction operator $\circ$ and dual division operators $\overset{\circ}{\rightarrow}, \overset{\circ}{\leftarrow}$ in terms of algebras $(\mathcal{L},\circ)$, where the algebra's binary operator $\circ$ is the semantic counterpart of the syntactic $\circ$.[12] An intepretation function $\llbracket\rrbracket$ assigns some subset of $\mathcal{L}$ to each type, satisfying the following conditions for non-primitive types and antecedent configurations:

(31)
$$\begin{aligned}
\llbracket \mathrm{X}\circ\mathrm{Y} \rrbracket &= \{x\circ y \in \mathcal{L} \mid x \in \llbracket \mathrm{X} \rrbracket \ \wedge \ y \in \llbracket \mathrm{Y} \rrbracket\} \\
\llbracket \mathrm{X}\overset{\circ}{\leftarrow}\mathrm{Y} \rrbracket &= \{x \in \mathcal{L} \mid \forall y \in \llbracket \mathrm{Y} \rrbracket .\ x\circ y \in \llbracket \mathrm{X} \rrbracket\} \\
\llbracket \mathrm{Y}\overset{\circ}{\rightarrow}\mathrm{X} \rrbracket &= \{x \in \mathcal{L} \mid \forall y \in \llbracket \mathrm{Y} \rrbracket .\ y\circ x \in \llbracket \mathrm{X} \rrbracket\} \\
\llbracket (\Gamma,\, \Delta)^{\circ} \rrbracket &= \{x\circ y \in \mathcal{L} \mid x \in \llbracket \Gamma \rrbracket \ \wedge \ y \in \llbracket \Delta \rrbracket\}
\end{aligned}$$

A sequent $\Gamma \Rightarrow \mathrm{A}$ *holds* in a model $((\mathcal{L},\circ), \llbracket\rrbracket)$, if $\llbracket\Gamma\rrbracket \subseteq \llbracket\mathrm{A}\rrbracket$. A sequent $\Gamma \Rightarrow \mathrm{A}$ is *valid* if it is true in all models. For systems defined in this way, the following laws hold:

$$\mathrm{A} \Rightarrow \mathrm{C}\overset{\circ}{\leftarrow}\mathrm{B} \quad \textit{iff} \quad \mathrm{A}\circ\mathrm{B} \Rightarrow \mathrm{C} \quad \textit{iff} \quad \mathrm{A} \Rightarrow \mathrm{B}\overset{\circ}{\rightarrow}\mathrm{C}$$

The particular categorial logic for which a semantics of the above pattern is appropriate depends on the further properties of the semantic operator $\circ$. For example, if $\circ$ is associative (i.e. we assume the equivalence axiom $x\circ(y\circ z) = (x\circ y)\circ z$) then we have a semantics for

[12] In general, I will notate such corresponding syntactic and semantic operators identically, though obviously the operators of the two domains are distinct. It should always be clear from context whether an object is in the syntactic or interpretive domain. Note in particular that lower case italics are used for semantic objects (e.g. $x\otimes y$), whereas syntactic types are written in in upper case roman (schematic types) or lower case roman (actual types).

the associative Lambek calculus. With no such additional equivalences, we have a system appropriate for the non-associative Lambek calculus.[13]

The properties of a semantics that is appropriate for a categorial logic gives some extent of a perspective on the relation between that logic and natural language. The algebraic objects in interpretations in some sense constitute abstract representations or descriptions of 'linguistic objects'. Given the character of $\circ$ in the semantics for associative Lambek calculus, we might loosely say that this system 'views' natural language as a system of strings. Similarly we might say that *non*-associative Lambek calculus views natural language as if it consisted of binary branching tree-like structures.[14]

9 Types as categories of partial descriptions

Let us consider how the above approach to algebraic semantics for categorial logics may be modified to be applicable for a hybrid substructural system. We have seen how the sets of objects categorised by types under the above interpretation scheme may be seen as abstract descriptions of linguistic objects. The crucial change for the hybrid approach is that the interpretive algebra must provide a range of operators for constructing 'descriptions', where these different modes of description are comparable. Thus, it is possible to have two objects in the algebra which may be possible descriptions of the same linguistic object, where one description may be 'less informative' than the other. Types are then viewed as categorising *partial* descriptions of linguistic objects, only categorising linguistic objects themselves indirectly.

For example, consider an interpretive algebra which includes at least the three operators $\{\odot,\ \bullet,\ \otimes\}$, for which we assume the following equivalence axioms:

$$x\bullet(y\bullet z) = (x\bullet y)\bullet z \qquad x\otimes(y\otimes z) = (x\otimes y)\otimes z \qquad (x\otimes y) = (y\otimes x)$$

Consider a possible object of this algebra $j\odot(s\odot m)$, which might loosely be viewed as providing a structural description of a particular sentence (perhaps *John saw Mary*). The object $j\bullet(s\bullet m)$ provides a less informative description of the same sentence, i.e. it indicates the words that are present, and their linear order, but not a specific (linguistic) bracketting (i.e. since $j\bullet(s\bullet m) = (j\bullet s)\bullet m$). Likewise, the object $j\otimes(s\otimes m)$ provides an even less informative description, which does not encode the (linguistic) order for the words that appear (since $\otimes$ is commutative).

This example suggests that different semantic operators may be compared in terms of the

[13] See Buszkowski (1986) for associative Lambek calculs, and Kandulski (1988) for the non-associative case. See Moortgat & Morrill (1991) for broader discussion of applying the general groupoid semantic approach to a range of categorial logics.

[14] Note quite trees, however, since the above model does not impose a condition of *acyclicity*, c.f. Venema, 1993/draft.

degree of informativeness of the descriptions whose construction they allow, e.g. for any x, y, the description $x{\otimes}y$ is less informative than the description $x{\bullet}y$. Assume an ordering $<_i$ of 'degree of informativeness' over semantic operators.[15] We can define a relation $\trianglerighteq_i$ over obects of the algebra ('descriptions') in terms of $<_i$ as in (32), which is such that $x \trianglerighteq_i y$ just in case the description x is *at least as* informative as y, e.g. $x \odot y \trianglerighteq_i x \odot y$, $x \odot y \trianglerighteq_i x{\bullet}y$, $x \odot y \trianglerighteq_i y \otimes x$. (This definition uses a convention such that $p[q]$ is an expression containing a subexpression q, and $p[r]$ is the result of replacing that subexpression by r.)

(32) $$\begin{array}{l} \textit{if} \quad x = w[u \circ' v] \;\;\&\;\; y = w[u \circ v] \;\;\&\;\; \circ <_i \circ' \;\; \textit{then} \;\; x \trianglerighteq_i y \\ \textit{if} \quad x = y \;\; \textit{then} \;\; x \trianglerighteq_i y \\ \textit{if} \quad x \trianglerighteq_i y \;\;\&\;\; y \trianglerighteq_i z \;\; \textit{then} \;\; x \trianglerighteq_i z \end{array}$$

If we assume that types categorise partial descriptions, it seems reasonable to expect that interpretations should satisfy a criterion of *coherence*, stated in (33). The criterion requires that if a description x is in an interpretation, then so also should be all related less informative descriptions (i.e. all descriptions y such that $x \trianglerighteq_i y$). (The idea behind the naming of the criterion is that if a certain description of a possible linguistic object is in an interpretation, it would be *incoherent* for the interpretation to deny any weaker description of the same object. I will sometimes loosely say that a type is coherent, meaning that its interpretation is coherent.) The interpretive scheme to be specified is such that all types are coherent in this sense.

(33) Coherence (of interpretation S):

$$\forall x, y.\ (x \in S \;\wedge\; x \trianglerighteq_i y) \rightarrow y \in S$$

A model is a triple $\mathcal{M} = ((\mathcal{L}, \mu), <_i, [\![\,]\!])$, where μ is a set of semantic operators and $(\mathcal{L}, \mu)$ an algebra defined on those operators, where $<_i$ is a partial ordering (of 'degree of informativeness') over μ, and $[\![\,]\!]$ is an interpretation function which assigns to each primitive type a *coherent* subset of $\mathcal{L}$ and satisfies the following conditions for non-primitive types and antecedent configurations:

(34) $$\begin{array}{ll} [\![\mathrm{X}{\circ}\mathrm{Y}]\!] & = \{z \in \mathcal{L} \mid \exists x, y.\ x \in [\![\mathrm{X}]\!] \;\wedge\; y \in [\![\mathrm{Y}]\!] \;\wedge\; x{\circ}y \trianglerighteq_i z\} \\ [\![\mathrm{X}\overset{\circ}{\leftarrow}\mathrm{Y}]\!] & = \{x \in \mathcal{L} \mid \forall y \in [\![\mathrm{Y}]\!].\ x{\circ}y \in [\![\mathrm{X}]\!]\} \\ [\![\mathrm{Y}\overset{\circ}{\rightarrow}\mathrm{X}]\!] & = \{x \in \mathcal{L} \mid \forall y \in [\![\mathrm{Y}]\!].\ y{\circ}x \in [\![\mathrm{X}]\!]\} \\ [\![(\Gamma, \Delta)^{\circ}]\!] & = \{z \in \mathcal{L} \mid \exists x, y.\ x \in [\![\Gamma]\!] \;\wedge\; y \in [\![\Delta]\!] \;\wedge\; x{\circ}y \trianglerighteq_i z\} \end{array}$$

[15] Note that for our (non-biased) example systems (i.e. for any $\circ, \circ' \in \{\odot, \ominus, \bullet, \otimes\}$), $\circ <_i \circ'$ *iff* the equivalence axioms of $\circ'$ are a subset of those of $\circ$ (c.f. the discussion of footnote 7). Again, the situation is not so simple when we consider dependency systems.

As before a sequent $\Gamma \Rightarrow \mathrm{A}$ holds in a model if $[\![\Gamma]\!] \subseteq [\![\mathrm{A}]\!]$, and is valid if it is true in all models. The conditions in (34) for product types and antecedent configurations directly ensure that their interpretations are coherent (assuming coherence for subtypes and subconfigurations assembled). The conditions for implicationals do not need any special manoeuvring to ensure coherence, coherence follows automatically. Thus, the coherence of any type also ensures coherence with respect to 'fragments' of descriptions in the type, i.e. if $x[y] \in [\![\mathrm{T}]\!]$, then for all z such that $y \trianglerighteq_i z$, $x[z] \in [\![\mathrm{T}]\!]$. All objects in (e.g.) $[\![\mathrm{X} \overset{\circ}{\leftarrow} \mathrm{Y}]\!]$ are fragments of objects in $[\![\mathrm{X}]\!]$, and so $[\![\mathrm{X} \overset{\circ}{\leftarrow} \mathrm{Y}]\!]$ is coherent provided that $[\![\mathrm{X}]\!]$ is. Since primitive types are coherent by stipulation, and the $[\![]\!]$ conditions for complex types/configurations preserve coherence, it follows that all interpretations are coherent.[16] Note that in the case where μ is singleton, the $\trianglerighteq_i$ relation collapses to equality, and this interpretation scheme becomes the same as that described in the previous section for single level categorial logics. Note that 'degree of informativeness' for algebraic operators (c.f. $<_i$) and 'degree of structural freedom' for syntactic operators (c.f. $<_S$) are intimately related. In particular, for any syntactic $\circ, \circ'$, we have $\circ' <_S \circ$ *if and only if* $\circ <_i \circ'$ for the corresponding semantic operators.

Crucial issues arise as to the soundness and completeness of the proof system with respect to this semantics. A proof system is sound with respect to a semantics if all derivable sequents are valid. It is complete if all valid sequents are derivable. I will briefly sketch a soundness proof. The problem of completeness is left for further research.

9.1 Soundness

We require the following lemma:

(35) **Lemma 1.**
For all configurations Δ, Π : *if* $[\![\Delta]\!] \subseteq [\![\Pi]\!]$ *then* $[\![\Gamma[\Delta]]\!] \subseteq [\![\Gamma[\Pi]]\!]$

Proof (sketch): By definition of $[\![]\!]$, if $[\![\Delta]\!] \subseteq [\![\Pi]\!]$, then for any configuration Γ, $[\![(\Delta,\Gamma)^\circ]\!] \subseteq [\![(\Pi,\Gamma)^\circ]\!]$ and $[\![(\Gamma,\Delta)^\circ]\!] \subseteq [\![(\Gamma,\Pi)^\circ]\!]$. Lemma follows by simple induction.

(36) **Theorem.**
If a sequent is derivable, then it is valid
(i.e. if $\Gamma \Rightarrow \mathrm{A}$ is derivable, then $[\![\Gamma]\!] \subseteq [\![\mathrm{A}]\!]$ in all models).

Proof (sketch): By induction on length of derivation.
Cases for final inference of a derivation:

- $\mathrm{X} \Rightarrow \mathrm{X}$ Immediate.

[16] As we have just seen, coherence is directly imposed in this approach. It is possible that such coherence, or at least its effects, could be achieved more simply, by conditions which link together the behaviour of the different semantic operations, in the manner of Moortgat (1993/??). This possibility remains to be explored.

- $\dfrac{\Delta \Rightarrow \mathrm{B} \quad \Gamma[\mathrm{B}] \Rightarrow \mathrm{A}}{\Gamma[\Delta] \Rightarrow \mathrm{A}}[\mathrm{cut}]$ By induction hypothesis (IH), and Lemma 1.
- $\dfrac{(\Gamma, \mathrm{B})^{\circ} \Rightarrow \mathrm{A}}{\Gamma \Rightarrow \mathrm{A}\overset{\circ}{\leftarrow}\mathrm{B}}[\overset{\circ}{\leftarrow}\mathrm{R}]$ Assume $t \in \llbracket\Gamma\rrbracket$. By IH, $(t \circ b) \in \llbracket\mathrm{A}\rrbracket$ for any $b \in \llbracket\mathrm{B}\rrbracket$. Hence, $t \in \llbracket\mathrm{A}\overset{\circ}{\leftarrow}\mathrm{B}\rrbracket$.
- $\dfrac{\Delta \Rightarrow \mathrm{C} \quad \Gamma[\mathrm{B}] \Rightarrow \mathrm{A}}{\Gamma[(\mathrm{B}\overset{\circ}{\leftarrow}\mathrm{C}, \Delta)^{\circ}] \Rightarrow \mathrm{A}}[\overset{\circ}{\leftarrow}\mathrm{L}]$ $\llbracket(\mathrm{B}\overset{\circ}{\leftarrow}\mathrm{C}, \mathrm{C})^{\circ}\rrbracket \subseteq \llbracket\mathrm{B}\rrbracket$, by definition of $\llbracket\rrbracket$. Hence, by IH and Lemma 1, $\llbracket(\mathrm{B}\overset{\circ}{\leftarrow}\mathrm{C}, \Delta)^{\circ}\rrbracket \subseteq \llbracket\mathrm{B}\rrbracket$, and $\llbracket\Gamma[(\mathrm{B}\overset{\circ}{\leftarrow}\mathrm{C}, \Delta)^{\circ}]\rrbracket \subseteq \llbracket\Gamma[\mathrm{B}]\rrbracket \subseteq \llbracket\mathrm{A}\rrbracket$
- $\dfrac{\Gamma \Rightarrow \mathrm{A} \quad \Delta \Rightarrow \mathrm{B}}{(\Gamma, \Delta)^{\circ} \Rightarrow \mathrm{A}{\circ}\mathrm{B}}[{\circ}\mathrm{R}]$ $\llbracket(\mathrm{A}, \mathrm{B})^{\circ}\rrbracket = \llbracket\mathrm{A}{\circ}\mathrm{B}\rrbracket$, by definition of $\llbracket\rrbracket$. Hence, by IH and Lemma 1, $\llbracket(\Gamma, \Delta)^{\circ}\rrbracket \subseteq \llbracket\mathrm{A}{\circ}\mathrm{B}\rrbracket$
- $\dfrac{\Gamma[(\mathrm{B}, \mathrm{C})^{\circ}] \Rightarrow \mathrm{A}}{\Gamma[\mathrm{B}{\circ}\mathrm{C}] \Rightarrow \mathrm{A}}[{\circ}\mathrm{L}]$ $\llbracket(\mathrm{A}, \mathrm{B})^{\circ}\rrbracket = \llbracket\mathrm{A}{\circ}\mathrm{B}\rrbracket$, by definition of $\llbracket\rrbracket$. Hence, by IH and Lemma 1, $\llbracket\Gamma[\mathrm{B}{\circ}\mathrm{C}]\rrbracket \subseteq \llbracket\Gamma[(\mathrm{B}, \mathrm{C})^{\circ}]\rrbracket \subseteq \llbracket\mathrm{A}\rrbracket$
- $\begin{array}{c}\Gamma[(\mathrm{B}, (\mathrm{C}, \mathrm{D})^{\circ})^{\circ}] \Rightarrow \mathrm{A} \\ \hline\hline \Gamma[((\mathrm{B}, \mathrm{C})^{\circ}, \mathrm{D})^{\circ}] \Rightarrow \mathrm{A}\end{array}[\mathrm{A}] \;\Big\{{\circ}{\in}\{\bullet, \otimes\}$ For any syntactic ${\circ}{\in}\{\bullet, \otimes\}$, $\llbracket((\mathrm{B},\mathrm{C})^{\circ},\mathrm{D})^{\circ}\rrbracket = \llbracket(\mathrm{B},(\mathrm{C},\mathrm{D})^{\circ})^{\circ}\rrbracket$, by associativity of corresponding semantic $\circ$. Hence, by IH and Lemma 1, $\llbracket\Gamma[((\mathrm{B},\mathrm{C})^{\circ},\mathrm{D})^{\circ}]\rrbracket \subseteq \llbracket\Gamma[(\mathrm{B},(\mathrm{C},\mathrm{D})^{\circ})^{\circ}]\rrbracket \subseteq \llbracket\mathrm{A}\rrbracket$.
- $\dfrac{\Gamma[(\mathrm{B}, \mathrm{C})^{\circ}] \Rightarrow \mathrm{A}}{\Gamma[(\mathrm{C}, \mathrm{B})^{\circ}] \Rightarrow \mathrm{A}}[\mathrm{P}] \;\Big\{{\circ}{\in}\{\ominus, \otimes\}$ For any syntactic ${\circ}{\in}\{\ominus, \otimes\}$, $\llbracket(\mathrm{B}, \mathrm{C})^{\circ}\rrbracket = \llbracket(\mathrm{C}, \mathrm{B})^{\circ}\rrbracket$, by commutativity of corresponding semantic $\circ$. Hence, by IH and Lemma 1, $\llbracket(\mathrm{C}, \mathrm{B})^{\circ}\rrbracket \subseteq \llbracket(\mathrm{B}, \mathrm{C})^{\circ}\rrbracket \subseteq \llbracket\mathrm{A}\rrbracket$.
- $\dfrac{\Gamma[(\mathrm{B}, \mathrm{C})^{\circ'}] \Rightarrow \mathrm{A}}{\Gamma[(\mathrm{B}, \mathrm{C})^{\circ}] \Rightarrow \mathrm{A}}[\subseteq] \;\Big\{\circ' <_S \circ$ Syntactic $\circ' <_S \circ$ implies semantic $\circ <_i \circ'$. Hence, by definition of $\llbracket\rrbracket$, $\llbracket(\mathrm{B}, \mathrm{C})^{\circ}\rrbracket \subseteq \llbracket(\mathrm{B}, \mathrm{C})^{\circ'}\rrbracket$. Hence, by Lemma 1 and IH, $\llbracket\Gamma[(\mathrm{B}, \mathrm{C})^{\circ}]\rrbracket \subseteq \llbracket\Gamma[(\mathrm{B}, \mathrm{C})^{\circ'}]\rrbracket \subseteq \llbracket\mathrm{A}\rrbracket$.

10 Natural deduction and for hybrid substructural systems

Although sequent formulations have various proof-theoretic advantages, it is useful for linguistic purposes to be able to present proofs in the more readable format of natural deduction. I shall develop such a natural deduction formulation in this section. The natural deduction proofs of this approach are isomorphic in structure to their associated proof terms.[17] Thus, merely stating the basic form of the natural deduction rules presents little problem, since these map trivially to the rules of formation for terms of the enriched lambda system. The difficulty lies in appropriately restricting the use of the rules to only allow correct inferencing. This is achieved by adopting an approach which falls within the general framework of *labelled deductive systems* (LDS: Gabbay, 1991).

In labelled deduction, every formula is associated with a label, which records information of the use of resources in proving that formula. Inference rules indicate how labels are

[17] Given this isomorphism, *proof normalisation* for natural deduction proofs maps trivially to normalisation of terms (for which the requisite contraction rules are given above). Proof normalisation for natural deduction is addressed in detail in Prawitz (1965). For the case of natural deduction linear logic see Benton *et al.* (1992), and for Lambek calculus, see Morrill *et al.* (1990), Barry *et al.* (1991).

propagated. Of course, this characterisation applies to any logic with a standard system of term assignment. Crucially in LDS, however, inference rules may have conditions on their use, which refer to labels, and which serve to ensure correct inferencing. Such labelling may be used for the purpose of "bringing semantics into syntax", as a basis for ensuring that the proof system behaves in a way appropriate to its intended semantics.

For the natural deduction formulation of the hybrid substructural sytem, the labelling discipline required is a standard style system of term assignment, i.e. formulas in proofs are associated with terms of the (enriched) lambda system, where assumptions are labelled with variables, and propagation of labels is in accordance with the Curry–Howard interpretation of proofs (Howard, 1969). This labelling system is used to 'import' the system of algebraic semantics presented in the previous section, via a mapping from label terms to objects which (in some sense) correspond to elements in the interpretive domain. Let us begin by stating the natural deduction rules, with term labelling.[18]

(37)

$$\frac{\mathrm{A}\overset{\circ}{\leftarrow}\mathrm{B} : a \qquad \mathrm{B} : b}{\mathrm{A} : a \overset{\leftarrow}{\circ} b}\,\overset{\circ}{\leftarrow}\mathrm{E} \qquad\qquad \frac{\begin{matrix}[\mathrm{B} : v]\\ \mathrm{A} : a\end{matrix}}{\mathrm{A}\overset{\circ}{\leftarrow}\mathrm{B} : [\overset{\leftarrow}{\circ}]v.a}\,\overset{\circ}{\leftarrow}\mathrm{I}$$

$$\frac{\mathrm{B} : b \qquad \mathrm{B}\overset{\circ}{\rightarrow}\mathrm{A} : a}{\mathrm{A} : b \overset{\rightarrow}{\circ} a}\,\overset{\circ}{\rightarrow}\mathrm{E} \qquad\qquad \frac{\begin{matrix}[\mathrm{B} : v]\\ \mathrm{A} : a\end{matrix}}{\mathrm{B}\overset{\circ}{\rightarrow}\mathrm{A} : [\overset{\rightarrow}{\circ}]v.a}\,\overset{\circ}{\rightarrow}\mathrm{I}$$

$$\frac{\begin{matrix}[\mathrm{B} : v]\ [\mathrm{C} : w] & \\ \mathrm{A} : a & \qquad \mathrm{B}\circ\mathrm{C} : b\end{matrix}}{\mathrm{A} : [b/v\circ w].a}\,\circ\mathrm{E} \qquad\qquad \frac{\mathrm{A} : a \qquad \mathrm{B} : b}{\mathrm{A}\circ\mathrm{B} : \langle a, b\rangle^{\circ}}\,\circ\mathrm{I}$$

Note that there are no rules corresponding to the sequent structural rules. All structure sensitivity is handled in terms of constraints on rule use.

Let us consider how the inference rules must be constrained. In this paper, we have confined our attention to systems with linear resource usage, i.e. where each assumption must be used precisely once. We can ensure firstly that no assumption is used *more* than once by excluding proofs where two or more assumptions are labelled with the same variable. This may be achieved either by simple stipulation, or indirectly by conditions on inferencing. In particular, we may require for all rules that combine two subproofs (i.e. all rules except the implicational introduction rules) that their proof terms have disjoint free variable sets, i.e.

[18] Morrill *et al.* (1990) and Barry *et al.* (1991) give natural deduction formulations of the associative Lambek calculus. See Troelstra (1993) and Benton *et al.* (1992) for natural deduction formulations of linear logic. The [$\circ$E] rule is of a form akin to multiplicative elimination rules discussed by Schroeder-Heister (1984), which ensure that *both* projections of the multiplicative must be used.

$\mathtt{FV}(b) \cap \mathtt{FV}(c) = \emptyset$ for inferences of the form: $\dfrac{\mathrm{B}:b \quad \mathrm{C}:c}{\mathrm{A}:a}$

I will adopt this second approach, but leave the condition as implicit. We can ensure that each resource is used *at least* once by requiring that rules which discharge assumptions do discharge at least one assumption,[19] or, equivalently, that the associated label abstractions bind non-vacuously. This requirement need not be directly imposed, since it follows from later conditions. As is standard with categorial logics, we require that deductions depend on at least one assumption. This condition can be enforced by requiring that the proof term that results with any inference must contain at least one free variable, i.e. for any proof term a, $\mathtt{FV}(a) \neq \emptyset$.

All further conditioning of the inference rules is directed toward importing the system of algebraic semantics discussed in the previous section. The method involves mapping from proof terms to objects that I will call *markers*, which in some sense denote elements of the model. We require an operator in the marker domain for each operator of the interpretive algebra (i.e. each $\circ \in \mu$), which will be notated identically. The set of possible marker terms M is derived by closing the set of variables VAR under the marker operators. A relation $\unrhd_i$ over marker terms operates in a precisely analogous manner to the $\unrhd_i$ relation over elements of the algebra (bringing with it the relation $<_i$ and the equivalence axioms of the algebra). Each proof term maps to a *set* of marker terms, under a procedure Σ, which is defined as in (38) by induction on the structure of terms.[20]

(38)
$$\begin{array}{lcl}
\Sigma(v) & = & \{\, v \,\}, \quad v \in \mathtt{VAR} \\
\Sigma(a \overset{\leftrightarrow}{\circ} b) & = & \{\, z \in M \mid \exists x, y.\ x \in \Sigma(a) \ \wedge \ y \in \Sigma(b) \ \wedge \ x{\circ}y \unrhd_i z \,\} \\
\Sigma(\langle a, b \rangle^{\circ}) & = & \{\, z \in M \mid \exists x, y.\ x \in \Sigma(a) \ \wedge \ y \in \Sigma(b) \ \wedge \ x{\circ}y \unrhd_i z \,\} \\
\Sigma([\overset{\leftarrow}{\circ}]v.a) & = & \{\, z \in M \mid z{\circ}v \in \Sigma(a) \,\} \\
\Sigma([\overset{\rightarrow}{\circ}]v.a) & = & \{\, z \in M \mid v{\circ}z \in \Sigma(a) \,\} \\
\Sigma([b/v{\circ}w].a) & = & \{\, z \in M \mid \exists x, y.\ x \in \Sigma(a) \ \wedge \ y \in \Sigma(b) \ \wedge \ x{=}u[v{\circ}w] \ \wedge \ u[y] \unrhd_i z \,\}
\end{array}$$

Under the Curry–Howard 'formulas-as-types' interpretation of proofs, a proof of A from B_1,..,B_n may be seen as a method or function for constructing an element in (the interpretation of) A from elements in the B_1,..,B_n. A proof term is essentially just a representation

[19] Compare, for example, to intuitionistic logic where (e.g.) use of [→I] may be accompanied by discharge of any number of assumptions, including zero.

[20] A similar method is used in Hepple (1993/draft), although there terms map to a *single* marker whose form is used in determining the applicability of inference rules. Both of these 'marker' systems may be compared to Buszkowski (19**), where a mapping is specified from directionally typed lambda terms (for implicational associative Lambek calculus) to the interpretive model (of strings).

of such a proof / method in a different notation. For us, the elements in the $B_1,..,B_n$ map to a *set* of elements in A. The procedure Σ serves to make explicit the schematic form of those elements in A so constructed. The crucial point for our purposes is that proof terms which do not correspond to correct deductions *fail* as methods of construction, so that for any term a not corresponding to a correct deduction, $\Sigma(a) = \emptyset$. Hence, we can ensure that deduction is appropriate for the semantics by requiring that the proof term a that results with any inference must be such that $\Sigma(a) \neq \emptyset$.

For example, consider the following two proofs, of which the second is incorrect. As the reader can easily verify, applying Σ to the proof term of (39) gives a non-empty set of markers, whereas the marker set for the proof term of (40) is empty.

(39)
$$\dfrac{\mathrm{X}\bullet\mathrm{Y}:w \qquad \dfrac{[\mathrm{X}:x] \qquad [\mathrm{Y}:y]}{\mathrm{X}\odot\mathrm{Y}:\langle x,y\rangle^{\odot}}\odot\mathrm{I}}{\mathrm{X}\odot\mathrm{Y}:[w/x\bullet y].\langle x,y\rangle^{\odot}}\bullet\mathrm{E}$$

(40)
$$\dfrac{\mathrm{X}\bullet\mathrm{Y}:w \qquad \dfrac{[\mathrm{X}:x] \qquad [\mathrm{Y}:y]}{\mathrm{X}\otimes\mathrm{Y}:\langle x,y\rangle^{\otimes}}\otimes\mathrm{I}}{\mathrm{X}\otimes\mathrm{Y}:[w/x\bullet y].\langle x,y\rangle^{\otimes}}\bullet\mathrm{E}$$

Thus, for the subterm $\langle x,y\rangle^{\odot}$, Σ gives $\{x\odot y, x\ominus y, x\bullet y, x\otimes y\}$. Crucially, this contains the element $x\bullet y$, and so Σ on $[w/x\bullet y].\langle x,y\rangle^{\odot}$ gives $\{w\}$. For the second proof, Σ on $\langle x,y\rangle^{\otimes}$ gives only $\{x\otimes y\}$, and so $\Sigma([w/x\bullet y].\langle x,y\rangle^{\otimes})$ is empty. Note that it is only uses of inferences that discharge assumptions (i.e. $[\overset{\circ}{\leftarrow}\mathrm{I}]$, $[\overset{\circ}{\rightarrow}\mathrm{I}]$, $[\circ\mathrm{E}]$) that are liable to produce incorrect proofs. The reason why should be clear from the definition of Σ. The clauses for these inference rules require that there are markers for the subterms which have certain forms, and if there are none such, an empty marker set results. The Σ clauses for the other cases merely put subterm markers together.

As another example, consider the following two proofs. Again the first proof is correct whilst the second is not. The crucial observation is that the marker set $\{x\bullet y, x\otimes y\}$ of the subterm $x \overset{\leftarrow}{\bullet} y$ contains an element $x\otimes y$ so that the $[\multimapinv\mathrm{I}]$ inference of (41) goes through, but contains no element $x\odot y$ and so the $[\not\circ\mathrm{I}]$ inference of (42) is incorrect.

(41)
$$\dfrac{\dfrac{\mathrm{X}/\mathrm{Y}:x \qquad [\mathrm{Y}:y]}{\mathrm{X}:x \overset{\leftarrow}{\bullet} y}/\mathrm{E}}{\mathrm{X}\multimapinv\mathrm{Y}:[\overset{\leftarrow}{\otimes}]y.\,x \overset{\leftarrow}{\bullet} y}\multimapinv\mathrm{I}$$

(42)
$$\dfrac{\dfrac{\mathrm{X}/\mathrm{Y}:x \qquad [\mathrm{Y}:y]}{\mathrm{X}:x \overset{\leftarrow}{\bullet} y}/\mathrm{E}}{\mathrm{X}\not\circ\mathrm{Y}:[\overset{\leftarrow}{\odot}]y.\,x \overset{\leftarrow}{\bullet} y}\not\circ\mathrm{I}$$

Since this natural deduction formulation is intended to be convenient for presenting linguistic derivations, it is undesirable that users should have to compute (potentially large) marker sets for labels at each stage of derivation. In practice, the natural method to employ is to construct only a *single* 'exemplar' marker term alongside each subderivation, which is arrived at in a *goal directed* fashion, and constructed from the marker terms that have been built for the immediate subterms. The marker term so generated at each stage should be the 'strongest description' possible, which is 'diluted' (by steps replacing a marker operator with a ‘less informative’ marker operator) only so far as is necessary to guarantee that the intended next inference goes through. Such dilution is only required to allow for inferences $[\overset{\circ}{\leftarrow}\mathrm{I}]$, $[\overset{\circ}{\rightarrow}\mathrm{I}]$ or $[\circ\mathrm{E}]$ (for reasons that should be clear from the above discussion). For example, consider the following derivation, where an exemplar marker for each proof term is shown to the right of $\Sigma\!\mapsto$ (except for assumptions, whose marker is simply their labelling variable). To allow for the $[\circ\!-\mathrm{I}]$ inference, the exemplar marker of the preceding term must be diluted, so that it can be matched to the pattern $\alpha\otimes y$. However, the term is only diluted to the minimum extent required — the subterm $w{\bullet}x$ is not diluted to $w\otimes x$, for example, as this is unnecessary to producing an overall term of the form $\alpha\otimes y$.

(43)
$$\dfrac{\dfrac{\dfrac{\dfrac{((\mathrm{W}/\mathrm{Z})/\mathrm{Y})/\mathrm{X}:w \qquad \mathrm{X}:x}{(\mathrm{W}/\mathrm{Z})/\mathrm{Y}:w \overset{\leftarrow}{\bullet} x}{/\mathrm{E}} \quad [\mathrm{Y}:y]}{\mathrm{W}/\mathrm{Z}:(w \overset{\leftarrow}{\bullet} x) \overset{\leftarrow}{\bullet} y}{/\mathrm{E}} \quad \mathrm{Z}:z}{\mathrm{W}:((w \overset{\leftarrow}{\bullet} x) \overset{\leftarrow}{\bullet} y) \overset{\leftarrow}{\bullet} z}{/\mathrm{E}}}{\mathrm{W}\circ\!-\mathrm{Y}:[\overset{\leftarrow}{\otimes}]y.((w \overset{\leftarrow}{\bullet} x) \overset{\leftarrow}{\bullet} y) \overset{\leftarrow}{\bullet} z}{\circ\!-\mathrm{I}}$$

$$\begin{array}{l} \overset{\Sigma}{\mapsto}\ w{\bullet}x \\ \overset{\Sigma}{\mapsto}\ (w{\bullet}x){\bullet}y \\ \overset{\Sigma}{\mapsto}\ ((w{\bullet}x){\bullet}y){\bullet}z \quad \trianglerighteq_i \quad ((w{\bullet}x)\otimes z)\otimes y \\ \overset{\Sigma}{\mapsto}\ (w{\bullet}x)\otimes z \end{array}$$

For purposes of presentation, stating exemplar markers in this way might lead to cluttered proofs, so it may be preferable to state only such marker elements as are needed to demonstrate that the next inference goes through. For example, (43) might be presented showing only the marker for the proof term preceding the $[\circ\!-\mathrm{I}]$ step.

Observe that the configuration of assumptions that we observe for the sequent formulation is absent under this natural deduction approach, so that the combination allowed by a proof might most naturally be written as $\Gamma \Rightarrow \mathrm{A}:a$, where Γ is a *set* of assumptions. We may ask what is the relation between such unconfigured sequents and their configured counterparts. The answer is straightforward: each marker in the marker set $\Sigma(a)$ for an unconfigured sequent $\Gamma \Rightarrow \mathrm{A}:a$ provides a ‘pattern’ for how we may configure the assumptions Γ. The procedure τ, defined in (44), is such that $\tau(\Gamma,m)$ returns a configuration, when Γ is a set of assumptions and m is a marker term defined on the labelling variables of Γ. If an unconfigured sequent $\Gamma \Rightarrow \mathrm{A}:a$ is provable under the natural deduction formulation, then for all $m{\in}\Sigma(a)$, $\tau(\Gamma,m) \Rightarrow \mathrm{A}$ is a theorem of the sequent system.

(44)
$$\begin{array}{lcl} \tau(\Gamma,v) & = & \mathrm{A}:v,\ \ \mathit{where}\ \ (\mathrm{A}:v)\in\Gamma \\ \tau(\Gamma,a{\circ}b) & = & (\tau(\Gamma,a),\tau(\Gamma,b))^{\circ} \end{array}$$

11 A labelled sequent formulation

The labelled natural deduction formulation of the previous section is in turn suggestive of how we might construct an alternative sequent formulation of the hybrid substructural approach.

This is a set-based sequent system, i.e. where sequents are objects of the form $\Gamma \Rightarrow \mathrm{A}$ such that Γ is a set of assumptions. The role of configurations in ensuring appropriate structure sensitivity is eliminated, being replaced by label conditions.[21]

The sequent rules are stated below. Linear use of resources follows from the way the rules are stated, particularly the use of $\uplus$ (indicating the union of two sets that are required to be disjoint).[22] To ensure correct inferencing, we need only require that with each inference, the conclusion's proof term a is such that $\mathtt{FV}(a) \neq \emptyset$ and $\Sigma(a) \neq \emptyset$.

(45)

$$\{\mathrm{A}:v\} \Rightarrow \mathrm{A}:v \quad (\mathrm{id}) \qquad\qquad \frac{\Delta \Rightarrow \mathrm{B}:b \quad \{\mathrm{B}:v\}\uplus\Gamma \Rightarrow \mathrm{A}:a}{\Delta\uplus\Gamma \Rightarrow \mathrm{A}:a[b/v]}[\mathrm{cut}]$$

$$\frac{\Gamma\uplus\{\mathrm{B}:v\} \Rightarrow \mathrm{A}:a}{\Gamma \Rightarrow \mathrm{A}\overset{\circ}{\leftarrow}\mathrm{B}:[\overset{\leftarrow}{\circ}]v.a}[\overset{\circ}{\leftarrow}\mathrm{R}] \qquad\qquad \frac{\Delta \Rightarrow \mathrm{C}:c \quad \Gamma\uplus\{\mathrm{B}:v\} \Rightarrow \mathrm{A}:a}{\Gamma\uplus\Delta\uplus\{\mathrm{B}\overset{\circ}{\leftarrow}\mathrm{C}:w\} \Rightarrow \mathrm{A}:a[(w\overset{\leftarrow}{\circ}c)/v]}[\overset{\circ}{\leftarrow}\mathrm{L}]$$

$$\frac{\Gamma\uplus\{\mathrm{B}:v\} \Rightarrow \mathrm{A}:a}{\Gamma \Rightarrow \mathrm{B}\overset{\circ}{\rightarrow}\mathrm{C}:[\overset{\rightarrow}{\circ}]v.a}[\overset{\circ}{\rightarrow}\mathrm{R}] \qquad\qquad \frac{\Delta \Rightarrow \mathrm{C}:c \quad \Gamma\uplus\{\mathrm{B}:v\} \Rightarrow \mathrm{A}:a}{\Gamma\uplus\Delta\uplus\{\mathrm{C}\overset{\circ}{\rightarrow}\mathrm{B}:w\} \Rightarrow \mathrm{A}:a[(c\overset{\rightarrow}{\circ}w)/v]}[\overset{\circ}{\rightarrow}\mathrm{L}]$$

$$\frac{\Gamma \Rightarrow \mathrm{A}:a \quad \Delta \Rightarrow \mathrm{B}:b}{\Gamma\uplus\Delta \Rightarrow \mathrm{A}\circ\mathrm{B}:\langle a,b\rangle^{\circ}}[\circ\mathrm{R}] \qquad\qquad \frac{\Gamma\uplus\{\mathrm{B}:v,\mathrm{C}:w\} \Rightarrow \mathrm{A}:a}{\Gamma\uplus\{\mathrm{B}\circ\mathrm{C}:x\} \Rightarrow \mathrm{A}:[x/v\circ w].a}[\circ\mathrm{L}]$$

This labelled sequent approach may have computational advantages over the earlier sequent formulation. To determine if some collection of types may be combined to give a certain result, we do not need to try deriving all the possible configurations of those types. Furthermore, the $[<_S]$ rule and the Associativity and Permutation structural rules are not required, thereby avoiding much spurious effort that these rules allow in proof search. In practice, however, whether theorem proving with this formulation is more efficient will depend on the cost of performing $\Sigma(a) \neq \emptyset$ checks. The marker sets of even relatively small label terms may contain large numbers of elements, and generating these elements may reintroduce much of the rebracketting and permuting that we would hope to have lost by eliminating the structural rules. However, it is my suspicion that for many marker systems (such as e.g. that above based on operators $\{\odot,\ominus,\bullet,\otimes\}$), all marker sets M will contain a non-empty subset of *maximal*

[21] A further possibility for formulating a labelled deduction version of hybrid substructural systems is in terms of *proof nets*, originally proposed in relation to linear logic, and developed in relation to associative Lambek calculus and LP by Roorda (1991). In a proof search context, it is useful to employ an algorithm that yields objects termed *proof structures*, of which only a subset meet the inductive definition for proof nets. Roorda provides a method by which the proof structures that are proof nets may be identified by conditions on proof terms. It is promising that Roorda's approach could be adapted to give a LDS proof net method for hybrid substructural systems, using the enriched term labelling system and appropriate marker conditions.

[22] Alternatively, the rules could have been stated in a way that would in general allow contraction and weakening (implicitly), but with label conditions on the specific rules ensuring linear usage.

elements — elements m, such that $m \trianglerighteq_i e$ for all $e \in \mathrm{M}$ — of which any one can serve as a representative for M. It is possible that the task of verifying a non-empty marker set for some label might be made more efficient by reformulating it in terms of computing single maximal marker elements for terms and subterms. This topic requires further research.

12 Linguistic applications

The central aim of this paper is to introduce a *general* framework for hybrid substructural categorial logics, rather than to put forward any one hybrid system. No claim of particular merit is made for the specific hybrid system that has been used for illustration. In the remainder of the paper, I will seek to give an idea of how a hybrid substructural approach might look in action, by presenting some possible linguistic accounts which depend on use of a hybrid framework.

12.1 Extraction: simple case

As noted earlier, linear implicationals allow for a possible treatment for extraction, including the problematic non-peripheral extraction case. A "sentence missing NP" may be derived as a constituent of type so–np (or equivalently np–os). The following proof, for example, is suitable for deriving the relative clause *which Mary gave to Bill.* Lexical string substitution proceeds as in (47a), where Ω gives a single order for the lexical atoms (indicated to the right of $\Omega\!\mapsto$, with quotation marks indicating an atom string). Lexical semantic substitution proceeds as in (47b), and simplifies under $\overset{\alpha}{\rightsquigarrow}$ (i.e. by deleting distinctions of directionality and level) to the result shown.

(46) (which) (mary) (gave) (to) (bill)

$$\dfrac{\mathrm{rel}/(\mathrm{s}{\circ\!-}\mathrm{np}) : v \qquad \dfrac{\dfrac{\mathrm{np} : w \qquad \dfrac{\dfrac{((\mathrm{np}\backslash\mathrm{s})/\mathrm{pp})/\mathrm{np} : x \quad [\mathrm{np} : u]}{(\mathrm{np}\backslash\mathrm{s})/\mathrm{pp} : x \overleftarrow{\bullet} u}\,/\mathrm{E} \qquad \dfrac{\mathrm{pp}/\mathrm{np} : y \quad \mathrm{np} : z}{\mathrm{pp} : y \overleftarrow{\bullet} z}\,/\mathrm{E}}{\mathrm{np}\backslash\mathrm{s} : (x \overleftarrow{\bullet} u) \overleftarrow{\bullet} (y \overleftarrow{\bullet} z)}\,/\mathrm{E}}{\mathrm{s} : w \overrightarrow{\bullet} ((x \overleftarrow{\bullet} u) \overleftarrow{\bullet} (y \overleftarrow{\bullet} z))}\,\backslash\mathrm{E}}{\mathrm{s}{\circ\!-}\mathrm{np} : [\overleftarrow{\otimes}]u.w \overrightarrow{\bullet} ((x \overleftarrow{\bullet} u) \overleftarrow{\bullet} (y \overleftarrow{\bullet} z))}\,{\circ\!-}\mathrm{I}}{\mathrm{rel} : v \overleftarrow{\bullet} ([\overleftarrow{\otimes}]u.w \overrightarrow{\bullet} ((x \overleftarrow{\bullet} u) \overleftarrow{\bullet} (y \overleftarrow{\bullet} z)))}\,/\mathrm{E}$$

$$\overset{\Sigma}{\mapsto} ((w \bullet x) \otimes (y \bullet z)) \otimes u$$

(47) a. Lexical string substitution:
$\mathit{which} \overset{\leftarrow}{\bullet} ([\overset{\leftarrow}{\otimes}]u.\mathit{mary} \overset{\rightarrow}{\bullet} ((\mathit{gave} \overset{\leftarrow}{\bullet} u) \overset{\leftarrow}{\bullet} (\mathit{to} \overset{\leftarrow}{\bullet} \mathit{bill})))$
$\Omega\!\mapsto$ *“which mary gave to bill”*

b. Lexical semantic substitution:
$\mathrm{which}' \overset{\leftarrow}{\bullet} ([\overset{\leftarrow}{\otimes}]u.\mathrm{mary}' \overset{\rightarrow}{\bullet} ((\mathrm{gave}' \overset{\leftarrow}{\bullet} u) \overset{\leftarrow}{\bullet} (\mathrm{to}' \overset{\leftarrow}{\bullet} \mathrm{bill}')))$
$\overset{\alpha}{\leadsto}$ $\mathrm{which}'\,(\lambda u.\mathrm{gave}'\ u\ (\mathrm{to}'\ \mathrm{bill}')\ \mathrm{mary}')$

12.2 Wrapping

We noted earlier that the possibility of assigning complex string terms to lexical items could allow for lexical encoding of partial derivations.[23] Consider a possible type ((np\s)/prt)/np for the particle verb *calls* (as in e.g. *John calls Mary up*), which explicitly subcategorises for a particle (prt). The derivation (48) might be used to combine the verb with its particle, but without the object NP.

(48)
$$\dfrac{\dfrac{\dfrac{\overset{\text{(calls)}}{((\mathrm{np}\backslash\mathrm{s})/\mathrm{prt})/\mathrm{np} : x} \quad [\mathrm{np} : y]}{(\mathrm{np}\backslash\mathrm{s})/\mathrm{prt} : x \underset{y}{\leftarrow}}/\mathrm{E} \quad \overset{\text{(up)}}{\mathrm{prt} : z}}{\mathrm{np}\backslash\mathrm{s} : (x \overset{\leftarrow}{\bullet} y) \overset{\leftarrow}{\bullet} z}/\mathrm{E}}{(\mathrm{np}\backslash\mathrm{s})\circ\!\!-\mathrm{np} : [\overset{\leftarrow}{\otimes}]y.\,(x \overset{\leftarrow}{\bullet} y) \overset{\leftarrow}{\bullet} z}\circ\!\!-\mathrm{I} \quad \overset{\Sigma}{\mapsto}\ (x\bullet y)\bullet z\ \trianglerighteq\ (x\otimes z)\otimes y$$

Lexical string substitution gives the string term: $[\overset{\leftarrow}{\otimes}]y.\,(\mathit{calls} \overset{\leftarrow}{\bullet} y) \overset{\leftarrow}{\bullet} \mathit{up}$. We could then treat the verb+particle *calls up* as a single discontinuous lexical functor with the lexical entry (49), whose string term is the complex term just derived, effectively giving a lexical encoding of the derivation (48). Note that in (49), however, the verb+particle complex has been assigned an *atomic* semantic term, illustrating that the approach allows for a mismatching status of string and semantic assignments.

(49) $<(\mathrm{np}\backslash\mathrm{s})\circ\!\!-\mathrm{np},\ \ [\overset{\leftarrow}{\otimes}]v.\,(\mathit{calls} \overset{\leftarrow}{\bullet} v) \overset{\leftarrow}{\bullet} \mathit{up},\ \ \mathrm{calls\text{-}up}'>$

Using the lexical entry (49) (plus obvious other lexical entries) *John calls Mary up* can be derived using the proof (50). Lexical substitution proceeds as in (51).

[23] This characteristic invites comparisons to some other formalisms, in particular *lexicalised tree adjoining grammar* (LTAG) (see Joshi *et al*, 1991, for discussion of this formalism), where the basic lexical and derivational units are partial phrase structure trees, each of which is ‘anchored’ to one or more words.

(50)

$$\begin{array}{c}
\dfrac{\begin{array}{c}(\text{mary})\\ \text{np}:x\end{array} \qquad \dfrac{\begin{array}{c}(\text{calls-up})\\ (\text{np}\backslash\text{s})\circ\!\!-\text{np}:y\end{array} \qquad \begin{array}{c}(\text{john})\\ \text{np}:z\end{array}}{\text{np}\backslash\text{s}:y\overset{\leftarrow}{\otimes}z}\,\circ\!\!-\text{E}}{\text{s}:x\overset{\rightarrow}{\bullet}(y\overset{\leftarrow}{\otimes}z)}\,\backslash\text{E}
\end{array}$$

(51) a. Lexical string substitution:

$\mathit{mary}\overset{\rightarrow}{\bullet}(([\overset{\leftarrow}{\otimes}]v.(\mathit{calls}\overset{\leftarrow}{\bullet}v)\overset{\leftarrow}{\bullet}\mathit{up})\overset{\leftarrow}{\otimes}\mathit{john})$

$\overset{\beta}{\leadsto}\ \mathit{mary}\overset{\rightarrow}{\bullet}((\mathit{calls}\overset{\leftarrow}{\bullet}\mathit{john})\overset{\leftarrow}{\bullet}\mathit{up})\ \overset{\Omega}{\mapsto}$ *"mary calls john up"*

b. Lexical semantic substitution:

$\text{mary}'\overset{\rightarrow}{\bullet}(\text{calls-up}'\overset{\leftarrow}{\otimes}\text{john}')\ \overset{\alpha}{\leadsto}\ \text{calls-up}'\ \text{john}'\ \text{mary}'$

The use of a complex lexical string term in this example has allowed us to simulate the non-concatenative string operation known as *wrap*, used most notably in work in Montague Grammar (e.g. Bach, 1981). A wrapping functor is one whose string is 'wrapped around' that of its argument under combination.

12.3 Quantification

Moortgat (1991), developing earlier proposals in Moorgat (1990), proposes a generalised quantifier type constructor q, which (informally) is such that a type q(X,Y,Z) is one that binds a position typed X within an expression typed Y, yielding a result typed Z. For example, a sententially scoped NP quantifier might have type q(np,s,s), i.e. so that a NP position is bound within a sentence to give sentence result.[24]

Moortgat (1991) discusses the quantor type constructor in relation to binary connectives known as 'extraction' (↑) and 'infixation' (↓), first proposed in Moortgat (1988). Informally, X↑Y is a function from Y to X, corresponding to an 'X missing a Y at some position'. Likewise informally, X↓Y is a function from Y to X, which yields an X by *infixing* into the string of the argument Y. Moortgat points out that a quantor type q(X,Y,Z) *would* be definable as Z↓(Y↑X) *if* infixation and extraction could be linked so that infixation was to the position of the 'missing X' of Y↑X. However, no such linkage is possible within the Moortgat (1991) formalism. Morrill & Solias (1993) and Hepple (1993/draft) propose formalisms in which versions of ↑ and ↓ can be defined which do allow such linkage of infixation and extraction, so that Moortgat's suggestion for quantification can be implemented.

For the present approach, we can recreate the insights of these proposals, but without needing to use distinguished extraction and infixation connectives. Instead, a quantifier may

[24] Such an account can be contrasted with the Lambek framework account of Emms (1990), which does not require any additional type forming operators, relying instead on the use of type *polymorphism*.

be treated as $\mathrm{s}\multimap(\mathrm{s}\multimap\mathrm{np})$, receiving the word order position of the missing NP in $\mathrm{s}\multimap\mathrm{np}$ due to the complex lexical string term assigned. This lexical string term is that which would arise under the 'type-raising' transformation: $\mathrm{np} \Rightarrow \mathrm{s}\multimap(\mathrm{s}\multimap\mathrm{np})$. Given (52) and other obvious lexical entries, *John gave someone money* can be derived under the proof (53), with lexical substitution as in (54). For examples with more than one quantifier, alternative quantifier scopings are derivable.

(52) $<\mathrm{s}\multimap(\mathrm{s}\multimap\mathrm{np}),\ [\overset{\leftarrow}{\otimes}]u.\,u\overset{\leftarrow}{\otimes}\,\mathit{someone},\ \mathrm{someone}'>$

(53)

$$\dfrac{\overset{\text{(someone)}}{\mathrm{s}\multimap(\mathrm{s}\multimap\mathrm{np}):q} \qquad \dfrac{\dfrac{\overset{\text{(john)}}{\mathrm{np}:x} \qquad \dfrac{\dfrac{\overset{\text{(gave)}}{((\mathrm{np}\backslash\mathrm{s})/\mathrm{np})/\mathrm{np}:y} \quad [\mathrm{np}:v]}{(\mathrm{np}\backslash\mathrm{s})/\mathrm{np}:y\overset{\leftarrow}{\bullet}v}\,/\mathrm{E} \qquad \overset{\text{(money)}}{\mathrm{np}:z}}{\mathrm{np}\backslash\mathrm{s}:(y\overset{\leftarrow}{\bullet}v)\overset{\leftarrow}{\bullet}z}\,/\mathrm{E}}{\mathrm{s}:x\overset{\rightarrow}{\bullet}((y\overset{\leftarrow}{\bullet}v)\overset{\leftarrow}{\bullet}z)}\,\backslash\mathrm{E}}{\mathrm{s}\multimap\mathrm{np}:[\overset{\leftarrow}{\otimes}]v.x\overset{\rightarrow}{\bullet}((y\overset{\leftarrow}{\bullet}v)\overset{\leftarrow}{\bullet}z)}\,\multimap\mathrm{I}}{\mathrm{s}:q\overset{\leftarrow}{\otimes}([\overset{\leftarrow}{\otimes}]v.x\overset{\rightarrow}{\bullet}((y\overset{\leftarrow}{\bullet}v)\overset{\leftarrow}{\bullet}z))}\,\multimap\mathrm{E}$$

(54) a. Lexical string subsitution:

$([\overset{\leftarrow}{\otimes}]u.\,u\overset{\leftarrow}{\otimes}\,\mathit{someone})\overset{\leftarrow}{\otimes}([\overset{\leftarrow}{\otimes}]v.\mathit{john}\overset{\rightarrow}{\bullet}((\mathit{gave}\overset{\leftarrow}{\bullet}v)\overset{\leftarrow}{\bullet}\mathit{money}))$

$\overset{\beta}{\rightsquigarrow}\ \mathit{john}\overset{\rightarrow}{\bullet}((\mathit{gave}\overset{\leftarrow}{\bullet}\mathit{someone})\overset{\leftarrow}{\bullet}\mathit{money})$

$\overset{\Omega}{\mapsto}$ *"john gave someone money"*

b. Lexical semantic subsitution:

$\mathrm{someone}'\overset{\leftarrow}{\otimes}([\overset{\leftarrow}{\otimes}]v.\mathrm{john}'\overset{\rightarrow}{\bullet}((\mathrm{gave}'\overset{\leftarrow}{\bullet}v)\overset{\leftarrow}{\bullet}\mathrm{money}'))$

$\overset{\alpha}{\rightsquigarrow}\ \mathrm{someone}'(\lambda v.\mathrm{gave}'\ v\ \mathrm{money}'\ \mathrm{john}')$

There is a sense in which this view of quantifiers seems very natural. Quantifiers behave distributionally very much like simple NPs, and so might be expected to have a basic string term compatible with a simple np type. Semantically, however, quantifiers are of a higher type, and so the string component must be raised for compatibility in constructing a well-formed lexical assignment. The specific transformation $\mathrm{np} \Rightarrow \mathrm{s}\multimap(\mathrm{s}\multimap\mathrm{np})$ achieves this compatibility without imposing additional word order constraints.

12.4 Extraction: pied piping

Morrill (1991) employs Moortgat's quantor operator in an account of pied piping. For pied piping of (e.g.) PPs, a relative pronoun is assigned type: $q(\mathrm{np}, \mathrm{pp}, \mathrm{rel}/(\mathrm{s}\uparrow\mathrm{pp}))$
This type allows the relative pronoun to infix to a NP position within a PP, giving a functor

rel/(s↑pp), i.e. which prefixes to a 'sentence missing PP' to give a relative clause. Hence, for example, *to whom* may be type rel/(s↑pp), and so *to whom John spoke* is a relative clause. The lexical semantics of *whom* 'canonicalises' the meaning, so that *to whom John spoke* is equivalent to *whom John spoke to.* In both Morrill & Solias (1993) and Hepple (1993/draft), this account is reconstructed using versions of ↑ and ↓ in place of q, as for quantification above.

Morrill's proposal can similarly be reconstructed under the present approach. Lexical entries for non-pied piping and PP pied piping *whom* are given in (55a,b), respectively. Assuming obvious additional lexical entries, *to whom John spoke* can be derived using the proof (56), with lexical substitution as in (57).

(55) a. $<\text{rel}/(\text{s}\circ\!\!-\text{np}),\ \textit{whom},\ \text{whom}'>$

b. $<(\text{rel}/(\text{s}\circ\!\!-\text{pp}))\circ\!\!-(\text{pp}\circ\!\!-\text{np}),\ [\overleftarrow{\otimes}]u.u\overleftarrow{\otimes}\,\textit{whom},\ \lambda q\lambda p.\text{whom}'(\lambda r.p(qr))>$

(56)

$$\dfrac{\dfrac{\begin{array}{c}\text{(whom)}\\ (\text{rel}/(\text{s}\circ\!\!-\text{pp}))\circ\!\!-(\text{pp}\circ\!\!-\text{np}):q\end{array} \qquad \dfrac{\dfrac{\begin{array}{c}\text{(to)}\\ \text{pp/np}:v \qquad [\text{np}:w]\end{array}}{\text{pp}:v\overleftarrow{\bullet}w}\,/\text{E}}{\text{pp}\circ\!\!-\text{np}:[\overleftarrow{\otimes}]w.v\overleftarrow{\bullet}w}\,\circ\!\!-\text{I}}{\text{rel}/(\text{s}\circ\!\!-\text{pp}):q\overleftarrow{\otimes}([\overleftarrow{\otimes}]w.v\overleftarrow{\bullet}w)}\,\circ\!\!-\text{E} \qquad \dfrac{\dfrac{\begin{array}{c}\text{(john)}\\ \text{np}:x\end{array} \qquad \dfrac{\begin{array}{c}\text{spoke}\\ (\text{np}\backslash\text{s})/\text{pp}:y \qquad [\text{pp}:z]\end{array}}{\text{np}\backslash\text{s}:y\overleftarrow{\bullet}z}\,/\text{E}}{\text{s}:x\overrightarrow{\bullet}(y\overleftarrow{\bullet}z)}\,\backslash\text{E}}{\text{s}\circ\!\!-\text{np}:[\overleftarrow{\otimes}]z.x\overrightarrow{\bullet}(y\overleftarrow{\bullet}z)}\,\circ\!\!-\text{I}}{\text{rel}:(q\overleftarrow{\otimes}([\overleftarrow{\otimes}]w.v\overleftarrow{\bullet}w))\overleftarrow{\otimes}([\overleftarrow{\otimes}]z.x\overrightarrow{\bullet}(y\overleftarrow{\bullet}z))}\,/\text{E}$$

(57) a. Lexical string subsitution:

$([\overleftarrow{\otimes}]u.u\overleftarrow{\otimes}\,\textit{whom})\overleftarrow{\otimes}([\overleftarrow{\otimes}]w.\textit{to}\overleftarrow{\bullet}w))\overleftarrow{\bullet}([\overleftarrow{\otimes}]z.\textit{john}\overrightarrow{\bullet}(\textit{spoke}\overleftarrow{\bullet}z))$

$\overset{\beta}{\rightsquigarrow}\ (\textit{to}\overleftarrow{\bullet}\,\textit{whom})\overleftarrow{\bullet}([\overleftarrow{\otimes}]z.\textit{john}\overrightarrow{\bullet}(\textit{spoke}\overleftarrow{\bullet}z))$

$\overset{\Omega}{\mapsto}$ *"to whom john spoke"*

b. Lexical semantic subsitution:

$((\lambda q\lambda p.\text{whom}'(\lambda r.p(qr)))\overleftarrow{\otimes}([\overleftarrow{\otimes}]w.\textit{to}\overleftarrow{\bullet}w))\overleftarrow{\bullet}([\overleftarrow{\otimes}]z.\textit{john}\overrightarrow{\bullet}(\textit{spoke}\overleftarrow{\bullet}z))$

$\overset{\beta}{\rightsquigarrow}\overset{\alpha}{\rightsquigarrow}\ \text{whom}'(\lambda r.\text{spoke}'(\text{to}'r)\,\text{john}')$

12.5 Dependency systems within a hybrid framework

Although the non-associative Lambek calculus is the weakest of the substructural logics as this landscape is typically conceived, there are possibilities for introducing further dimensions of resource structure and structure sensitivity. Moortgat & Morrill (1991) make a linguistically motivated proposal to allow binary trees that are *biased* or *headed*, where for each mother node, one daughter is distinguished as being the *head* or *primary* element, and the other as the *non-head*, *dependent* or *secondary* element. Such biases allow construction of trees which encode *dependency structure*. For the sequent approach, there might be right and left headed brackets $(.,.)^{\lhd}$ and $(.,.)^{\rhd}$ for building configurations, such that Γ is the primary subconfiguration in both $(\Delta,\Gamma)^{\lhd}$ and $(\Gamma,\Delta)^{\rhd}$, and corresponding groups of syntactic operators such as e.g. $\{\, \lhd,\, \overset{\lhd}{\leftarrow},\, \overset{\lhd}{\rightarrow} \}$ and $\{\, \rhd,\, \overset{\rhd}{\leftarrow},\, \overset{\rhd}{\rightarrow} \}$. From a semantic perspective, there would be similarly biased algebraic operators $\lhd$ $\rhd$, giving objects such as e.g. $j \lhd (s \rhd m)$.

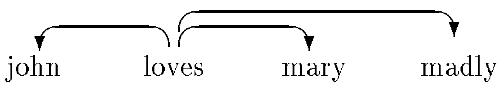
Moortgat & Morrill illustrate there proposals by addressing the use of metrical trees for handling stress phenomena within metrical phonology. In this context, they discuss how different equivalence axioms over the objects of the algebra will serve to maintain or undermine certain possible distinctions. For example, let the *most prominent* element of an object be that reached by recursing down through the object always following primary branches, and the *least prominent* element be that reached by following secondary branches, e.g. x is most prominent in $(x \rhd y) \rhd z$ and least prominent in $(x \lhd y) \lhd z$. The equivalence axiom (58a) preserve only most prominence, i.e. if α and α' are equivalent under (58a), then α and α' must have the same most prominent element, but may not have the same least prominent element. The axiom in (58b), however, preserves *both* most and least prominence. Thus, the selection of equivalence axioms over the objects of the algebra is significant to the distinctions that are represented.

(58) a. $(x \rhd y) \rhd z \quad = \quad x \rhd (y \lhd z)$

b. $(x \rhd y) \rhd z \quad = \quad x \rhd (y \rhd z)$

Within a hybrid substructural framework, the possibity exists of having a range of coexisting dependendency levels which differ (speaking in terms of the interpretive algebra, for the moment) in the equivalence axioms that apply and hence the distinctions that are represented. Each dependency level might be related to other dependency levels, as well as to non-dependency levels, where bias is not represented. In the case where one dependency level is related to another, we can expect that relations between operators should preserve bias. For example, given a level with (semantic) operators $\{\overset{a}{\rhd},\ \overset{a}{\lhd}\}$ and a 'less informative' level $\{\overset{b}{\rhd},\ \overset{b}{\lhd}\}$, we expect that $\overset{b}{\rhd} <_i \overset{a}{\rhd}$ and $\overset{b}{\lhd} <_i \overset{a}{\lhd}$, but *not* $\overset{b}{\rhd} <_i \overset{a}{\lhd}$ or $\overset{b}{\lhd} <_i \overset{a}{\rhd}$. When a dependency level is related to a non-dependency level, however, bias will not be preserved, e.g. we might have $\otimes <_i \overset{a}{\rhd}$ *and* $\otimes <_i \overset{a}{\lhd}$.

In follows subsections, I will address the use of various interrelated dependency systems within a hybrid substructural approach. The notion of head that will concern me is syntactic / grammatical, c.f. syntactic grammars where heads are distinguished, e.g. dependency grammars. Note that in dependency grammar, a sentence such as e.g. *John loves Mary madly* receives a flat structure with *loves* as head:



In the present approach, the use of binary operators requires that a hierarchical structure be imposed, e.g. as in: $(\mathit{john} \triangleleft ((\mathit{loves} \triangleright \mathit{mary}) \triangleright \mathit{madly}))$
For immediate purposes, it is useful to employ a notion of 'head' more akin to that used in dependency grammar. This notion may be characterised recursively, i.e. the atomic element reached by recursing down through a structure following only primary branches (c.f. 'most prominent element' above). This notion will be referred to simply as 'head' (or sometimes 'recursive head'), and contrasted with the notion 'immediate head' which is h in any expression $h \triangleright x$ or $x \triangleleft h$. We may refer to subexpression x as 'a projection of h' if h is the (recursive) head of x. The dependents of a head are those substructures which are 'immediate dependents' of its projections, i.e. d is a dependent of an atomic h in s, just in case $s \equiv t[g \triangleright d]$ or $s \equiv t[d \triangleleft g]$, where h is head of g.

12.6 Lexical subcategorisation and grammatical relations

Within the categorial framework known as Montague Grammar, it has been suggested that grammatical relations (GRs), i.e. notions such as *subject* and *direct object*, might be defined in terms of subcategorisation and combination (see especially Dowty, 1982). For example, a *subject* is defined to be that complement which combines with an intransitive verb (or VP) to give a sentence, whilst a *direct object* is defined as that complement which combines with a transitive verb type to give an intransitive verb. This view of GRs is problematic for use with systems such as associative Lambek calculus for various reasons.[25] One problem is due to the flexibility of derivation under the associative Lambek calculus, which is such that for a constituent of type s/np or np\s, the sought argument NP does not need to be the subject of the constituent's main verb.

One possible solution to this problem is that lexical subcategorisation might be specified in terms of the non-associative connectives $\circ\!\!\!/$, $\circ\!\!\!\backslash$, i.e. since then the np argument of any constituent s$\circ\!\!\!/$np or np$\circ\!\!\!\backslash$s must be the maximal verb's subject. Selecting this level for specifying lexical subcategorisation would not lead to a general loss of flexibility in a hybrid framework, since any non-associative implication may be freely rewritten as a corresponding associative implication (e.g. X$\circ\!\!\!/$Y $\Rightarrow$ X/Y).[26]

For various reasons, it seems likely that 'lexical subcategorisation' should also encode dependency information. (We shall see some possible uses for this information shortly.) Let us

[25] A considerable problem for using argument order to encode grammatical hierarchy involves appropriately handling word order when indirect objects and other more oblique complements are taken into consideration. This problem is addressed in Montague Grammar by use of 'wrapping' string operations. I will ignore this problem here. See Hepple (1990) for discussion.

[26] Hepple (1990) provides connectives $\circ\!\!\!\backslash$, $\circ\!\!\!/$ for specifying lexical subcategorisation (which are *not* the same as the $\circ\!\!\!\backslash$, $\circ\!\!\!/$ we have considered in this paper), which are such that e.g. X$\circ\!\!\!/$Y $\Rightarrow$ X/Y is derivable, but X/Y $\Rightarrow$ X$\circ\!\!\!/$Y is not. This asymmetry is achieved by the (unsatisfactory) method of providing the lexical subcategorisation connectives with an incomplete logic, with only elimination rules.

assume a level of non-associative dependency structure, having syntactic product operators $\rhd$ and $\lhd$. The associated implicationals might be used for specifying lexical subcategorisation. For example, a transitive verb might be: $(\mathrm{np}\overset{\lhd}{\rightarrow}\mathrm{s})\overset{\rhd}{\leftarrow}\mathrm{np}$. A VP-level adverbial might have type: $(\mathrm{np}\overset{\lhd}{\rightarrow}\mathrm{s})\overset{\rhd}{\rightarrow}(\mathrm{np}\overset{\lhd}{\rightarrow}\mathrm{s})$, where, crucially, the principal connective $\overset{\rhd}{\rightarrow}$ serves to indicate that, although the adverbial is a functor over the VP, it is the VP which is head of the construction.[27] On the other hand, an auxiliary would be $(\mathrm{np}\overset{\lhd}{\rightarrow}\mathrm{s})\overset{\rhd}{\leftarrow}(\mathrm{np}\overset{\lhd}{\rightarrow}\mathrm{s})$, indicating its head status.

In the hybrid framework, this level would be linked to various other levels, some biased, some non-biased, and being subject to differing regimes of associativity and permutation. For example, it seems reasonable that $\odot <_i \rhd$ and $\odot <_i \lhd$, i.e. so that bias may be dropped, yielding the familiar non-associative level. Another related level (notated $\overset{p}{\rhd}$ / $\overset{p}{\lhd}$) might involve maintaining bias, but adding a bias-preserving form of permutation, e.g.: $x \overset{p}{\rhd} y = y \overset{p}{\lhd} x$

A possible use for this latter level is in characterising certain constituent notions without reference to language specific ordering information. For example, we could refer to VP *without* directional information as $\mathrm{s}\overset{\rhd p}{\leftarrow}\mathrm{np}$ (or equivalently $\mathrm{np}\overset{\lhd p}{\rightarrow}\mathrm{s}$). (Note that $\mathrm{s}\overset{\rhd}{\leftarrow}\mathrm{np} \Rightarrow \mathrm{s}\overset{\rhd p}{\leftarrow}\mathrm{np}$ and $\mathrm{np}\overset{\lhd}{\rightarrow}\mathrm{s} \Rightarrow \mathrm{s}\overset{\rhd p}{\leftarrow}\mathrm{np}$.) We might take the view that a head which requires a VP complement should not make reference to (language specific) directional information encoded by its VP complement, merely the fact that its complement is a VP. This could be avoided by assigning such the head a type of the form $\mathrm{X}\overset{\rhd}{\leftarrow}(\mathrm{s}\overset{\rhd p}{\leftarrow}\mathrm{np})$.

In the remainder of the paper, I will address some possible uses for dependency systems which are less rigidly structured than the level for encoding lexical subcategorisation. Of course, to be linguistically useful, such levels must be 'accessible' from the level used in encoding lexical subcategorisation (i.e. for a level with operators $\overset{*}{\rhd}$ / $\overset{*}{\lhd}$, we expect $\overset{*}{\rhd} <_i \rhd$ and $\overset{*}{\lhd} <_i \lhd$).

12.7 Bounded head movement

Consider how we might handle bounded movement of heads, where the relevant sense of 'bounded' is that the head cannot move out of its own 'domain', i.e. the domain of head + complements/adjuncts that the head itself defines. Rephrasing this requirement, when a head h is moved to the peripery of some domain D, h must itself be head of D (in the 'recursive' sense discussed above). I will notate the system to be discussed via $\overset{*}{\rhd}$ / $\overset{*}{\lhd}$. Clearly, the equivalences defined over the relevant system of algebraic objects must be such as to preserve heads, i.e. if $\alpha = \alpha'$, then α and α' have the same head. Furthermore, to allow

[27] It should be clear that the *head/non-head* distinction is treated as a primitive, non-derived notion in the Moortgat & Morrill approach. This view may be contrasted with approaches such as that of Barry & Pickering (1990), where the notion of head is defined in terms of the notion of categorial functor. Problems arise for such approaches regarding the status of adjuncts.

abstraction over the head element, equivalence in the system must be such that if an atomic x is head of some α, then there exists some β such that $\alpha = \beta \overset{*}{\triangleleft} x$. The equivalence axioms in (59) are sufficient for this purpose. Bounded head movement may be handled by assigning a displaced head having canonical type H a further type of the form (e.g.): $\mathrm{A}/(\mathrm{B} \overset{\triangleleft *}{\leftharpoonup} \mathrm{H})$.

(59) a. $x \overset{*}{\triangleright} y \;= y \overset{*}{\triangleleft} x$

b. $x \overset{*}{\triangleright} (y \overset{*}{\triangleright} z) \;= (x \overset{*}{\triangleright} y) \overset{*}{\triangleright} z$

For example, consider the type combinations in (60) and the partial proof (61), which could be extended to give a proof of either (60a,b). (61) shows some marker cases for the proof term which are relevant to what are possible subsequent introduction steps. We can see from the 'maximal' marker term $x \triangleright (y \triangleright z)$ that the functor labelled x is 'head of the proof', whereas the functor labelled y is 'head' only of a subproof. The cases (i) and (ii) show that (60a) could be derived by discharging either functor assumption. Cases (iii) and (iv) show that to derive (60b), by a $[\overset{\triangleleft *}{\leftharpoonup}\mathrm{I}]$ inference, we may only discharge the assumption labelled x, i.e. the head of the overall proof.

(60) a. $\{\mathrm{X} \overset{\triangleright}{\leftharpoonup} \mathrm{X}, \mathrm{X}\} \Rightarrow \mathrm{X} \multimap (\mathrm{X} \overset{\triangleright}{\leftharpoonup} \mathrm{X})$

b. $\{\mathrm{X} \overset{\triangleright}{\leftharpoonup} \mathrm{X}, \mathrm{X}\} \Rightarrow \mathrm{X} \overset{\triangleleft *}{\leftharpoonup} (\mathrm{X} \overset{\triangleright}{\leftharpoonup} \mathrm{X})$

(61)
$$\frac{\mathrm{X} \overset{\triangleright}{\leftharpoonup} \mathrm{X} : x \qquad \dfrac{\mathrm{X} \overset{\triangleright}{\leftharpoonup} \mathrm{X} : y \qquad \mathrm{X} : z}{\mathrm{X} : y \overset{\leftharpoonup}{\triangleright} z}\overset{\triangleright}{\leftharpoonup}\mathrm{E}}{\mathrm{X} : x \overset{\leftharpoonup}{\triangleright} (y \overset{\leftharpoonup}{\triangleright} z)}\overset{\triangleright}{\leftharpoonup}\mathrm{E} \qquad \overset{\Sigma}{\mapsto} \;\; x \triangleright (y \triangleright z) \begin{cases} \text{(i)} & \trianglerighteq_i \;\; (y \otimes z) \otimes x \\ \text{(ii)} & \trianglerighteq_i \;\; (x \otimes z) \otimes y \\ \text{(iii)} & \trianglerighteq_i \;\; (y \triangleright z) \overset{*}{\triangleleft} x \\ \text{(iv)} & \neg (\trianglerighteq_i \;\; \alpha \overset{*}{\triangleleft} y) \end{cases}$$

One linguistic application for this notion of bounded head movement is the treatment of the Verb Second (V2) behaviour of languages such as Dutch and German, where systematic differences are observed between the word order of main and subordinate clauses. Various accounts of V2 in a range of different formalisms and frameworks have a common underlying informal basis: main clause word order is derived (in some sense) from subordinate clause word order by, firstly, fronting of the main clause finite verb, followed by optional movement of some constituent to the left of the fronted verb. Note that this verb movement is bounded — the finite verb may not come from any embedded clause, nor may any non-finite verb dominated by the finite verb take the latter's place in the fronting movement. A criticism that may be levelled against most V2 treatments is that the mechanism provided to allow fronting of the verb could as well allow unbounded movement (even if the boundedness requirement may be encoded), whereas the boundedness of V2 verb movement appears a very strong characteristic of the phenomenon cross-linguistically.

For the present approach, V2 verb movement might be handled using a lexical rule such as (62), generating an appropriate movement type for each finite verb.[28] Note that s and s_m in (62) are distinguished types for embedded and main clauses, respectively. The use of this mechanism for handling V2 verb movement explains not only why the movement is bounded, but also why it is only the 'maximal' verb in any cluster that may be fronted.

(62) $\mathrm{V} \;\Rightarrow\; \mathrm{s}_m/(\mathrm{s} \overset{\triangleleft *}{\leftarrow} \mathrm{V})$ (V a finite verb)

12.8 Local and non-local dependency relations

It is a crucial characteristic of the dependency algebra discussed in the previous section that for any object α, there is some β such that $\alpha = \beta \overset{*}{\triangleleft} x$ where x is atomic, and this is only possible where x is the head of α. It follows, at the level of the proof system, that a $[\overset{\triangleleft *}{\leftarrow}\mathrm{I}]$ (or $[\overset{\triangleright *}{\rightarrow}\mathrm{I}]$) inference may only discharge the assumption which is 'head of the proof'. Let us turn our attention to possibilities for abstracting over 'dependent' elements. At the level of the interpretive algebra, the crucial question is: for any α, what are the atomic x such that there exists some β such that $\alpha = \beta \overset{*}{\triangleright} x$ (or $\alpha = x \overset{*}{\triangleleft} \beta$) (I will again notate the system to be discussed via $\overset{*}{\triangleright} / \overset{*}{\triangleleft}$).

Assume again the bias-preserving equivalence (63a), so that the linear order of heads and dependents is not relevant to possibilities. Consider the systems arrived at by choice of 'associativity axioms' from the equivalences (63b,c).

(63) a. $x \overset{*}{\triangleright} y \;= y \overset{*}{\triangleleft} x$

b. $x \overset{*}{\triangleleft} (y \overset{*}{\triangleright} z) \;= (x \overset{*}{\triangleleft} y) \overset{*}{\triangleright} z$

c. $x \overset{*}{\triangleright} (y \overset{*}{\triangleright} z) \;= (x \overset{*}{\triangleright} y) \overset{*}{\triangleright} z$

Firstly, consider the system including only the axioms (63a,b). Note that various other equivalences follow automatically, e.g. $(x \overset{*}{\triangleright} y) \overset{*}{\triangleright} z \;= (x \overset{*}{\triangleright} z) \overset{*}{\triangleright} y$. Note that (63b) preserves both heads and dependents (in the sense defined earlier), i.e. on both sides of the equivalence, y is head, and x and z its dependents. Thus, for any α, α' such that $\alpha = \alpha'$ under (63a,b), α and α' have the same head h, and h has the 'same' dependents in both (or strictly, there is a one-to-one mapping between its dependents in the two cases, where the subexpressions so paired

[28] This treatment of V2 is closely related to that of Hepple (1990). That proprosal, however, relies on a general polymodal treatment of locality constraints for ensuring that the verb movement is appropriately bounded. Hepple (1990), following Jacobson (1987), gives a verb movement based 'pseudo-wrapping' treatment of English word order, to avoid problems for adopting the Montague Grammar treatment of grammatical relations (discussed earlier) within a concatenative CG framework. This view of English word order could also be reconstructed within the above framework, without need to stipulate or otherwise enforce the boundedness of verb movement.

are equivalent). Furthermore, free restructuring under (63a,b) will allow any dependent of the head of some α to 'move up' to topmost hierarchical position. For any α and atomic x, there is some β such that $\alpha = \beta\,{}^{*}_{\triangleright}\,x$ if and only if x is a dependent of the head of α.

This system would, for example, allow for a version of bounded movement of dependents. For any constituent derived of the form $\mathrm{B}\overset{\triangleright *}{\leftarrow}\mathrm{C}$, the argument C must correspond to a direct dependent of the constituent's head, and so assigning an element with canonical type C a movement type of the form (e.g.) $\mathrm{A}/(\mathrm{B} \overset{\triangleright *}{\leftarrow} \mathrm{C})$ would allow it to move to the periphery of the domain of its head but not beyond.

Equivalence under (63c) does preserve the head of any expression, but does *not* preserve dependents. It does, however, preserve 'least prominence' in the sense defined earlier (i.e. the least prominent element in an expression is the atom reached by recursing down always following secondary branches). In a system with just (63a,c), $\alpha = \beta\,{}^{*}_{\triangleright}\,x$ (for some β) if and only if (atomic) x is the least prominent element of α. When all three equivalences are included, much greater freedom results. In particular, for any α, $\alpha = \beta\,{}^{*}_{\triangleright}\,x$ (for some β) for any atomic x in α *except* its head.

One way that these observations might be summed up is as follows. The dependency system with equivalences (63a,b) is the maximally free system which preserves dependent status. With inclusion of (63c), this preservation is undermined so that any element in an object α (except its head) may be treated as if it were a top-level dependent. The 'division of labour' between the two equivalences (63b,c) may be characterised as follows: (63b) allows for 'local' restructuring of dependency relations (i.e. restructuring purely within the 'domain' of a single head), whereas (63c) allows an 'embedded' dependency to move up a level to be treated as if it were a dependent of a dominating head, and so introduces a 'non-local' aspect to dependency restructuring.

We may ask if the notion of locality that arises for the system (63a,b) is useful linguistically. For example, consider bounded extraction, and in particular, bounded movement of adverbial adjuncts. As is well known, extraction of such adjuncts from embedded clauses is in general ungrammatical (taking a simple view of the data, which is sufficient for immediate purposes):

(64) a. John [$_{\text{vp}}$ [$_{\text{vp}}$ opened the box] [$_{\text{adv}}$ with a crowbar]]

b. How$_{\text{i}}$ did John [$_{\text{vp}}$ [$_{\text{vp}}$ open the box] $\underline{\ \ }_i$]

c. *How$_{\text{i}}$ do you remember that John [$_{\text{vp}}$ [$_{\text{vp}}$ opened the box] $\underline{\ \ }_i$]
(* under intended reading)

It is readily seen, however, that the (63a,b) system is *too* restrictive for treating this phenomenon. An adverbial *can* move outside of the domain of its head, as in (65b), where the adverbial modifies the embedded VP, and hence is dependent on the embedded verb:

(65) a. John wants $[_{\mathrm{VP}}\ [_{\mathrm{VP}}$ to leave$]$ tomorrow$]$

b. When$_i$ does John want $[_{\mathrm{VP}}\ [_{\mathrm{VP}}$ to leave$]\ ___i\,]$

Another possible application is locality constraints on reflexivisation. Some recent proposals have explained the locality requirement for core cases of reflexivisation in terms of 'co-argumentation', i.e. reflexives and binder must be arguments of the same lexical predicate (Pollard & Sag, 1992; Reinhart & Reuland, 1991). In the present approach, co-argumentation might be reconstrued in terms of 'co-dependents'. A problem arises with examples such as *John$_i$ spoke to himself$_i$*, where the reflexive is embedded within a PP, suggesting need of the excluded 'non-local' equivalence (63c).

It appears then that the (63a,b) dependency system is too restrictive for most linguistic purposes. A possible alternative is that we might seek a system with some restricted involvement of the 'non-local' equivalence (63c), so that we could go beyond the limitations of the (63a,b) system only in limited circumstance and extent. I will explore two alternatives for a mixed local/non-local systems in the following subsections.

12.9 Distinguishing dependency relations

In the preceding subsections, I have treated the relationship between verb and complement and between verb and adjunct as being underlyingly the same, expressed at the 'lexical level' in the interpretive algebra via the same operators $\triangleright / \triangleleft$. However, various linguistic accounts have made explanatory use of the differential status of head-complement and head-adjunct relations. The possibility exists that we might use different operators to encode these different dependencies, e.g. a system $\overset{c}{\triangleright} / \overset{c}{\triangleleft} / \overset{a}{\triangleright} / \overset{a}{\triangleleft}$ where c indicates a head-complement relation and a indicates a head-adjunct relation. Then, for example, the dependency structure of *John loves Mary madly* might take the following form: $(john \overset{c}{\triangleleft} ((loves \overset{c}{\triangleright} mary) \overset{a}{\triangleright} madly))$ Further distinctions could be similarly encoded. For example, use has been made of the status of subjects as 'external argument', in contrast to the 'internal' status of the verb's other complements. This distinction could again be marked in the system of operators, so that our example sentence's structure might be: $(john \overset{e}{\triangleleft} ((loves \overset{c}{\triangleright} mary) \overset{a}{\triangleright} madly))$
I shall not attempt to enumerate all such distinctions that might be encoded.

The non-associative level at which such relations are lexically specified might be related to a range of other levels, subject to differing schemes of associativity and commutativity, etc. At some of these levels, the distinctions between differing kinds of head-dependent relations might be discarded, whilst at others, the distinctions might be maintained, and play a role in determining the possible derivable constituents. For example, it is possible that a non-commutative system where there was partial associativity conditioned by dependency distinctions might give rise to a level of 'flexible dependency structure' suitable for addressing phenomena of non-constituent coordination.[29]

I will here consider a possible role for fine-grained dependency distinctions in relation to island constraints. Extraction will be handled at a level of operators $\overset{\alpha*}{\triangleright}$ / $\overset{\alpha*}{\triangleleft}$ (α ranging over distinctions between dependencies), where the linkage of this level to the level of (distinguished) lexical relations preserves distinctions, i.e. so that $\overset{\alpha*}{\triangleright} <_i \overset{\alpha}{\triangleright}$ and $\overset{\alpha*}{\triangleleft} <_i \overset{\alpha}{\triangleleft}$ for each distinguished operator. The following equivalences are the same as those in (63), except that operators now bear dependency distinctions (α, β ranging over distinctions).

(66) a. $x \overset{\alpha*}{\triangleright} y \;= y \overset{\alpha*}{\triangleleft} x$

b. $x \overset{\alpha*}{\triangleleft} (y \overset{\beta*}{\triangleright} z) \;= (x \overset{\alpha*}{\triangleleft} y) \overset{\beta*}{\triangleright} z$

c. $x \overset{\alpha*}{\triangleright} (y \overset{\beta*}{\triangleright} z) \;= (x \overset{\alpha*}{\triangleright} y) \overset{\beta*}{\triangleright} z$

As should be clear from preceding discussion, admitting only (66a,b) would allow only (overly restrictive) head-bounded movement of dependents, whereas including also (66c) would freely allow (dependent) extraction of any element except the (maximal) head.

However, a middle ground may be achieved if we allow *restricted* involvement of (66c), which gives rise to the 'non-local' aspect of dependency restructuring. For example, if a certain case of (66c) is not allowed, (e.g.) if $x \overset{p*}{\triangleright} (y \overset{q*}{\triangleright} z) \neq (x \overset{p*}{\triangleright} y) \overset{q*}{\triangleright} z$, this has the effect that 'q-dependents' may not move up out of 'p-domains', so that 'p-domains' are islands to extraction of 'q-dependents'. For example, if we exclude equivalences $x \overset{a*}{\triangleright} (y \overset{\beta*}{\triangleright} z) \;= (x \overset{a*}{\triangleright} y) \overset{\beta*}{\triangleright} z$ for all distinctions β, the island status of adjuncts follows (recall that a is the dependency distinction for adjuncts). A restricted version of (66c) might take the form:

(67) $x \overset{\alpha*}{\triangleright} (y \overset{\beta*}{\triangleright} z) \;= (x \overset{\alpha*}{\triangleright} y) \overset{\beta*}{\triangleright} z \qquad$ where $\langle \alpha, \beta \rangle \in \{\ldots\}$

The set $\{\ldots\}$ in (67) might be used simply for stipulating observed possibilities. Alternatively, however, we might seek to encode more general claims about island phenomena in terms of requirements on the structure of the set, e.g. transitivity (viewing the set as an extensional characterisation of a relation).

12.10 Counting dependency chains and locality in binding

Let us turn our attention to locality constraints on binding. As noted earlier, some accounts have explained the locality requirement for core cases of reflexivisation in terms of 'co-argumentation', i.e. reflexive and binder must be arguments of the same lexical predicate

[29] C.f. the proposals of Barry & Pickering (1990, 1992/draft) concerning flexible dependency and the treatment of coordination.

(Pollard & Sag, 1992; Reinhart & Reuland, 1991). In the present approach, co-argumentation might be handled in terms of 'co-dependents'. As we have seen, a system with equivalence axioms (68) only allows abstraction over immediate dependents of a constituent's head.

(68) a. $x \overset{*}{\triangleright} y \;= y \overset{*}{\triangleleft} x$

b. $x \overset{*}{\triangleleft} (y \overset{*}{\triangleright} z) \;= (x \overset{*}{\triangleleft} y) \overset{*}{\triangleright} z$

Then, for example, a subject antecedent version of the reflexive *himself* might have the lexical type assignment (69) (where 'VP' is an abbreviation for the complex type).[30]

(69) $< \mathrm{VP} \xrightarrow{\triangleleft *} (\mathrm{VP} \xrightarrow{\triangleright *} \mathrm{np}),\;\; [\overrightarrow{\triangleright *}]u.\, u_{\,(\overleftarrow{\triangleright *})}\, \textit{himself},\;\; \lambda v.\, wvv >$

The reflexive is assigned a 'raised' type, and a likewise 'raised' string term such that the word order position that will result for the reflexive word corresponds to that of the NP argument of $\mathrm{VP}\xrightarrow{\triangleright *}\mathrm{np}$. The semantics of the reflexive effects co-binding of the NP argument sought by $\mathrm{VP}\xrightarrow{\triangleright *}\mathrm{np}$ with its next argument slot, i.e. with the subject position. Locality for reflexivisation follows because the NP argument of $\mathrm{VP}\xrightarrow{\triangleright *}\mathrm{np}$ must be a dependent of the head of the phrase, i.e. the verb, and hence a co-dependent of the verb's subject.

A problem arises with examples such as *John$_i$ spoke to himself$_i$*, where the reflexive is embedded within a PP, showing that simple 'co-dependency' is not sufficient. The embedded location of the NP appears to require involvement of the 'non-local' dependency restructuring equivalence: $x \overset{*}{\triangleright} (y \overset{*}{\triangleright} z) \;= (x \overset{*}{\triangleright} y) \overset{*}{\triangleright} z$, but as we have seen, free involvement of this axiom would undermine locality altogether. The Pollard & Sag (1991) approach allows the semantics of the PP to be 'identified' (via unification) with that of the embedded NP in such 'case-marking' preposition cases, giving a result semantically precisely as if the NP were a direct argument of the verb. One possible reconstrual of this solution is that the dependency relation between preposition and head is in some way treated as 'zero length'.

This observation is suggestive of an approach which exhibits sensitivity to the 'length' of chains of dependency (in terms of levels of embedding). A standard lexical dependency has length 1, whereas the lexical dependency between a case-marking preposition and its NP complement has length 0. Hence, the dependency chain between a verb taking a PP argument and the NP therein is, in sum, 1. Dependency chains with length >1 will not be allowed for reflexivisation. A system which allows full counting of the length of dependency chains, however, would appear too strong.[31] This might allow, for example, a version of reflevisation (or extraction) where binder and anaphor had to be precisely four levels of embedding apart;

[30] Adapting from the proposal of Szabolcsi (1987), where reflexives are functions over verbs, and have 'duplicator' semantics causing co-binding of two verb argument positions.

[31] C.f. Chomsky, 19**.

not less, not more. A restricted version of ‘counting’ might instead distinguish only the cases 0, 1 and >1.

Length of dependency must initially be recorded lexically. Assume we have operators $\overset{n}{\triangleright}$ / $\overset{n}{\triangleleft}$ for $0 \leq n \leq 1$. For example, a transitive verb might be $(\mathrm{np}\overset{\triangleleft 1}{\rightarrow}\mathrm{s})\overset{\triangleright 1}{\leftarrow}\mathrm{np}$, whilst a case-marking preposition might be $\mathrm{pp}\overset{\triangleright 0}{\leftarrow}\mathrm{np}$. Assume next a level with operators $\overset{*n}{\triangleright}$ / $\overset{*n}{\triangleleft}$ for $0 \leq n \leq 2$ (where 2 is used to signify all values >1). The linkage between the two levels preserves both bias and count. We require the equivalences in (70) (where n, m, r range over {0,1,2}), which are schematic for a greater number of fully instantiated equivalences (21 in fact). The bias-preserving permutation and local dependency restructuring equivalences (70a,b) do not change any count values. The non-local restructuring equivalence (70c), however, may give different values for operators on the two sides. Consider an object such as $x \overset{*1}{\triangleright} (y \overset{*1}{\triangleright} z)$. Here z is embedded to a depth of 2. Restructuring under (70c) yields the result $(x \overset{*1}{\triangleright} y) \overset{*2}{\triangleright} z$, where z appears as if it were a top-level dependent, but where its count records the fact that originates from a deeper level of embedding. The count for the ‘promoted’ dependent z is derived by adding together the counts for the previously chained dependencies. Note that if this sum is greater than 1, the result is taken as 2 (representing all values >1). The opposite direction of restructuring under (70c) is non-deterministic, e.g. $(x \overset{*2}{\triangleright} y) \overset{*2}{\triangleright} z$ may restructure to any of $x \overset{*2}{\triangleright} (y \overset{*2}{\triangleright} z)$, $x \overset{*2}{\triangleright} (y \overset{*1}{\triangleright} z)$ or $x \overset{*2}{\triangleright} (y \overset{*0}{\triangleright} z)$. However, all three alternatives correctly maintain the information that the depth of embedding of z is of an extent >1.

(70) a. $x \overset{*n}{\triangleright} y \ = y \overset{*n}{\triangleleft} x$

b. $x \overset{*n}{\triangleleft} (y \overset{*m}{\triangleright} z) \ = (x \overset{*n}{\triangleleft} y) \overset{*m}{\triangleright} z$

c. $x \overset{*n}{\triangleright} (y \overset{*m}{\triangleright} z) \ = (x \overset{*n}{\triangleright} y) \overset{*r}{\triangleright} z \qquad r = \textsc{max}(\{m{+}n,\, 2\})$

Under this approach, a subject antecedent version of *himself* might have the lexical type assignment (71). Here, the NP argument of $\mathrm{VP}\overset{\triangleright *1}{\rightarrow}\mathrm{np}$ must be a dependent at depth 1, and so may be either an argument of the verb or inside a PP argument of the verb.

(71) $< \mathrm{VP}\overset{\triangleleft *1}{\rightarrow} (\mathrm{VP} \overset{\triangleright *1}{\rightarrow} \mathrm{np}),\ [\overset{\rightarrow}{\triangleright *}]u.\, u_{(\overset{\leftarrow}{\triangleright *})}\ \textit{himself},\ \lambda v.\, wvv >$

This same mechanism may be used for handling certain aspects of ‘anti-locality’ conditions on binding. For example, the lexical entry (72) for the personal pronoun *him* allows for *bound* uses, with subject antecedent. This type requires that the bound argument position be embedded to a depth >1 in the structure, so allowing (e.g.) *Every boy$_i$ thinks Mary likes him$_i$* but not **Every boy$_i$ likes him$_i$*.

(72) $< \mathrm{VP}\overset{\triangleleft *2}{\rightarrow} (\mathrm{VP} \overset{\triangleright *2}{\rightarrow} \mathrm{np}),\ [\overset{\rightarrow}{\triangleright *}]u.\, u_{(\overset{\leftarrow}{\triangleright *})}\ \textit{him},\ \lambda v.\, wvv >$

References

[1] Bach, E. 1981, 'Discontinuous Constituents in Generalized Categorial Grammars.' *NELS*, **11**, pp1–12.

[2] Barry, G., Hepple, M., Leslie, N. and Morrill, G. 1991. 'Proof figures and structural operators for categorial grammar'. In *EACL–5*, Berlin.

[3] Barry, G. and Pickering, M. 1990. 'Dependency and Constituency in Categorial Grammar.' In Barry, G. and Morrill, G. (Eds), *Studies in Categorial Grammar.* Edinburgh Working Papers in Cognitive Science, Volume 5. Centre for Cognitive Science, University of Edinburgh.

[4] Barry, G. and Pickering, M. 1992/draft. 'Dependency, Flexible Constituency and Categorial Grammar.' ms. Centre for Cognitive Science, University of Edinburgh.

[5] van Benthem, J. 1983. 'The semantics of variety in Categorial Grammar.' Report 83-29, Department of Mathematics, Simon Fraser University. Also in W. Buszkowski, W. Marciszewski and J. van Benthem (Eds), *Categorial Grammar*, Vol. 25, Linguistic and Literary Studies in Eastern Europe, John Benjamins, Amsterdam/Philadelphia. 1988.

[6] van Benthem, J. 1991. *Language in Action: Categories, Lambdas and Dynamic Logic*, North-Holland, Amsterdam.

[7] Benton, N., Bierman, G., de Paiva, V. & Hyland, M. 1992. 'Term assignment for intuistionistic linear logic.' Technical Report, Cambridge University Computer Laboratory.

[8] Buszkowski, W. 1986. 'Completeness results for Lambek syntactic calculus.' *Zeitschrift für mathematischen Logik und Grundlagen der Mathematik*, **32**, pp13–28.

[9] Buszkowski, W. 1987. 'The Logic of Types.' In J. Srzednicki (Ed), *Initiatives in Logic*, Martinus Nijhoff Publishers, Dordrecht.

[10] Došen, K. 1990. 'Modal translations in substructural logics'. Report 10-90, Universität Konstanz.

[11] Dowty, D. 1982. 'Grammatical relations and Montague grammar.' In P. Jacobson and G.K. Pullum (Eds), *The Nature of Syntactic Representation*, D. Reidel, Dordrecht, 79–130.

[12] Emms, M. 1990. 'Polymorphic Quantifiers.' In Barry, G. and Morrill, G. (Eds), *Studies in Categorial Grammar.* Edinburgh Working Papers in Cognitive Science, Volume 5. Centre for Cognitive Science, University of Edinburgh.

[13] Gabbay, D. 1991. *Labelled deductive systems.* Draft 1991. (to appear: Oxford University Press).

[14] Girard, J-Y. 1987. 'Linear Logic'. In: *Theoretical Computer Science* **59**, pp1-102.

[15] Hepple, M. 1990. *The Grammar and Processing of Order and Dependency: A Categorial Approach.* Ph.D. dissertation, Centre for Cognitive Science, University of Edinburgh.

[16] Hepple, M. 1993/draft. 'Labelled Deduction and Discontinuous Constituency'. ms. I.R.C.S. University of Pennsylvania.

[17] Hollenberg, M. 1991. 'Proof nets and modalities'. In *Proceedings 2nd CLIN conference*, Amsterdam.

[18] Howard, W.A. 1969. The formulae-as-types notion of construction. In J.R. Hindley and J.P. Seldin (Eds), *To H.B. Curry, Essays on Combinatory Logic, Lambda Calculus and Formalism*, Academic Press, 1980.

[19] Jacobson, P. 1987. 'Phrase Structure, Grammatical Relations and Discontinuous Constituents.' In Huck, G.J. and Ojeda, A.E. (Eds), *Syntax and Semantics*, **20**: *Discontinuous Constituency*, Academic Press, New York. 27–69.

[20] Joshi, A.K., Vijay-Shanker, K. and Weir, D. 1991. 'The convergence of mildly context-sensitive formalisms'. In P. Sells, S. Shieber and T. Wasow (Eds.) *Foundational issues in Natural Language Processing.* MIT Press, Cambridge MA.

[21] Kandulski, M. 1988. 'The non-associative Lambek calculus'. In W. Buszkowski, W. Marciszewski and J. van Benthem (Eds), *Categorial Grammar*, Vol. 25, Linguistic and Literary Studies in Eastern Europe, John Benjamins, Amsterdam/Philadelphia. 1988.

[22] Lambek, J. 1958. The mathematics of sentence structure. *American Mathematical Monthly*, **65**, pp154–170.

[23] Lambek, J. 1961. 'On the calculus of syntactic types.' In R. Jakobson (Ed), *Structure of Language and its Mathematical Aspects*, Proceedings of the Symposia in Applied Mathematics XII, American Mathematical Society.

[24] Moortgat, M. 1988. *Categorial Investigations: Logical and Linguistic Aspects of the Lambek Calculus*, Foris, Dordrecht.

[25] Moortgat, M. 1990. 'The quantification calculus.' In Hendriks, H. and Moortgat, M. (Eds), *Theory of Flexible Interpretation.* Esprit DYANA Deliverable R1.2.A, Institute for Language, Logic and Information, University of Amsterdam.

[26] Moortgat, M. 1991. 'Generalized Quantification and Discontinuous type constructors'. To appear in W. Sijtsma & A. van Horck (Eds), *Proceedings Tilburg Symposium on Discontinuous Constituency.* De Gruyter, Berlin.

[27] Moortgat, M. 1992. 'Labelled Deductive Systems for categorial theorem proving.' In: P. Dekker & M. Stokhof (Eds), *Proceedings of Eigth Amsterdam Colloquium*, University of Amsterdam.

[28] Moortgat, M. & Morrill, G. 1991. 'Heads and Phrases: Type Calculus for Dependency and Constituency.' To appear: *Journal of Language, Logic and Information.*

[29] Morrill, G. 1990. 'Rules and Derivations: Binding Phenomena and Coordination in Categorial Logic.' Deliverable R1.2.D of DYANA Dynamic Interpretation of Natural Language, ESPRIT Basic Research Action BR3175.

[30] Morrill, G. 1992. Categorial Formalisation of Relativisation: Pied Piping, Islands and Extraction Sites. Research Report, Departament de Llenguatges i Sistemes Informàtics, Universitat Politècnica de Catalunya.

[31] Morrill, G., Leslie, N., Hepple, M. and Barry, G. 1990. 'Categorial deductions and structural operations.' In Barry, G. and Morrill, G. (Eds), *Studies in Categorial Grammar.* Edinburgh Working Papers in Cognitive Science, Volume 5. Centre for Cognitive Science, University of Edinburgh.

[32] Morrill, G. & Solias, M.T. 1993. 'Tuples, Discontinuity, and Gapping in Categorial Grammar.' In: *Proceedings of EACL–93*, Utrecht.

[33] Oehrle, R. & Zhang, S. 1989. 'Lambek calculus and Preposing of Embedded Subjects.' in: Chicago Linguistic Society 25.

[34] Prawitz, D. 1965. *Natural Deduction: a Proof Theoretical Study*, Almqvist and Wiksell, Uppsala.

[35] Pollard, C. & Sag, I. 1992. 'Anaphors in English and the Scope of Binding Theory'. *Linguistic Inquiry* **23**, 261-303.

[36] Reinhart, T. & Reuland, E. 1991. 'Anaphors and Logophors: an Argument Structure Perspective'. In: Koster, J. & Reuland, E. (eds), *Long Distance Anaphora*, CUP.

[37] Roorda, D. 1991. *Resource Logics: Proof Theoretical Investigations.* Ph.D. Dissertation, Amsterdam.

[38] Schroeder-Heister, P. 1984. 'A natural extension of natural deduction.' *The Journal of Symbolic Logic*, 49(4):1284–1300.

[39] Szabolcsi, A. 1987. 'Bound variables in syntax (are there any?)' In Groenendijk, J., Stokhof, M. and Veltman, F. (Eds), *Sixth Amsterdam Colloquium*, Institute for Language, Logic and Information, University of Amsterdam.

[40] Troelstra, A. 1992. *Lectures on Linear Logic.* CSLI.

[41] Venema, Y. 1993a. 'Meeting a modality? Restricted permutation for the Lambek calculus.' OTS Working Paper, Onderzoekinstituut voor Taal en Spraak, Universiteit Utrecht, Netherlands.

[42] Venema, Y. 1993a. 'Meeting Strength In Substructural Logics.' OTS Working Paper, Onderzoekinstituut voor Taal en Spraak, Universiteit Utrecht, Netherlands.

[43] Venema, Y. 1993/draft. 'Tree Models and (Labelled) Categorial Grammar'. ms. Universiteit Utrecht.

[44] Versmissen, K. 1992. 'Adding a universal modality to the Lambek calculus.' In: *OTS Yearbook 1992.* Onderzoekinstituut voor Taal en Spraak, Universiteit Utrecht, Netherlands.

[45] Versmissen, K. 1992. 'Categorial Grammar, modalities, and algebraic semantics.' In: *Proceedings of EACL-6*, Utrecht.

[46] Wansing, W. 1990. 'Formulas-as-types for a hierarchy of sublogics of Intuistionistic Propositional Logic.' ms. Institut für Philosophie, Freie Universität Berlin.