

Discontinuous Lambek Calculus

Glyn Morrill · Oriol Valentín · Mario Fadda

Received: date / Accepted: date

Abstract The search for a full treatment of wrapping in type logical grammar has been a task of long-standing. In this paper we present a calculus for discontinuity addressing this challenge, ω -**DL**. The calculus allows an unbounded number of points of discontinuity (hence the prefix ω -) and includes both deterministic and nondeterministic discontinuous connectives. We believe that it constitutes a general and natural extension of the Lambek calculus **L**. Like the Lambek calculus it has a sequent calculus which is a sequence logic without structural rules, and it enjoys such properties as Cut-elimination, the subformula property and decidability.

By n -**DL** we refer to ω -**DL** restricted to at most n points of discontinuity. **0-DL** is the original Lambek calculus **L**. Of particular interest is **1-DL** in which the unicity of the point of discontinuity means that the deterministic and nondeterministic discontinuous connectives coincide. We illustrate **1-DL** with linguistic applications to medial extraction, discontinuous idioms, parentheticals, gapping, VP ellipsis, reflexivization, quantification, pied-piping, appositive relativisation, comparative subdeletion, and cross-serial dependencies. We further illustrate deterministic **2-DL** with linguistic application to anaphora, and nondeterministic **2-DL** with linguistic application to particle shift and complement alternation. That is we address, so far as we are aware, all the constructions for which some version of discontinuity has been proposed in the type-logical literature.

Keywords categorial logic · Cut-elimination · discontinuity · Lambek calculus · Type Logical Grammar · wrapping.

Work partially funded by the DGICYT project TIN2005-08832-C03-03 (MOISES-BAR).

Glyn Morrill
Universitat Politècnica de Catalunya
E-mail: morrill@lsi.upc.edu

Oriol Valentín
Universitat Pompeu Fabra
E-mail: oriol.valentin@upf.edu

Mario Fadda
Universitat Politècnica de Catalunya
E-mail: mfadda@lsi.upc.edu

Contents

1	Introduction	2
2	Theory of Discontinuous Lambek Calculus	8
3	Applications of Discontinuous Lambek Calculus	17
4	Conclusion	39
5	Appendix: Cut-Elimination	40

1 Introduction

A critical issue in natural grammar is ‘discontinuity’: form-meaning mismatch. If natural languages were purely ‘continuous’, with dependent elements always adjacent, some variety of immediate constituent analysis would have sufficed. Within the framework of Type Logical Grammar (TLG), our concern here is to find a technically natural and empirically wide-ranging adaptation of the continuous model of Lambek categorial grammar to discontinuity.

Ojeda (2006)[?] identifies two early approaches to discontinuity as being ‘permutation’ (Chomsky 1955, p.405)[?]¹ and ‘wrapping’ (Yngver 1960, p.448)[?]. Permutation has been introduced into TLG via permutation structural modalities (Morrill et al. 1990; Morrill 1994, ch. 7; Moortgat 1999)[?][?][?] inspired by the ‘exponentials’ of linear logic (Girard 1987)[?]. Here we develop the wrapping approach to discontinuity in TLG.

In Section 1.1 we resume the (continuous) Lambek calculus **L** and in Section 1.2 we describe its type-logical semantics. **L** provides a model with features and properties which we aim to replicate in discontinuous Lambek calculus ω -**DL**. In Section 1.3 we survey the history of approaches to discontinuity in categorial grammar. In Section 2 we present the theory of the discontinuous Lambek calculus, and hypersequent calculus and labelled natural deduction for ω -**DL**. In Section 3 we present linguistic applications of discontinuous Lambek calculus and we conclude in Section 4. In the appendix Section 5 we prove Cut-elimination.

1.1 The (Continuous) Lambek Calculus **L**

We take as basic TLG **L**, by which we mean the system of Lambek (1958)[?] with type-logical semantics along the lines of van Benthem (1983)[?] and Lambek (1988)[?].

(1) **Definition** (*basic prosodic algebra*)

A *basic prosodic algebra* is an algebra $(L, +)$ of arity (2) which is a free semigroup. I.e. L is a set, and $+$ is a binary operation on L such that for all $s_1, s_2, s_3 \in L$,

$$s_1 + (s_2 + s_3) = (s_1 + s_2) + s_3 \text{ associativity;}$$

furthermore, up to associativity every element of L has a unique factorization into primes (freeness).²

¹ Chomsky (1965)[?] revoked permutation in favour of copying and deletion.

² Factors of an element s are elements s_1, \dots, s_n such that $s = s_1 + \dots + s_n$; a prime is an element which has no factorization other than into just itself.

We call + concatenation.

We use the term “prosodic” in the sense of both prose and suprasegmental phonology: naturally sentential grammar must relate textual word order (prose) and semantics, but we think that it is also here that speech intonational phrasing and contour (suprasegmental phonology) must be interfaced with its semantic reflexes; the term “syntax” we reserve for the mediation of prosodics and semantics by sentential grammar.

(2) **Definition** (*types of L*)

The set \mathcal{F} of *types* of \mathbf{L} is defined on the basis of a set \mathcal{P} of primitive or atomic basic types as follows:

$$\mathcal{F} ::= \mathcal{P} \mid \mathcal{F} \bullet \mathcal{F} \mid \mathcal{F} \setminus \mathcal{F} \mid \mathcal{F} / \mathcal{F}$$

The connective \bullet is called (continuous) ‘product’, \setminus is called ‘under’, and $/$ is called ‘over’.

(3) **Definition** (*prosodic interpretation of L*)

A *prosodic interpretation* of \mathbf{L} is a function $[[\cdot]]$ mapping each type $A \in \mathcal{F}$ into a subset of L such that:

$$\begin{aligned} [[A \setminus C]] &= \{s_2 \mid \forall s_1 \in [[A]], s_1 + s_2 \in [[C]]\} \\ [[C / B]] &= \{s_1 \mid \forall s_2 \in [[B]], s_1 + s_2 \in [[C]]\} \\ [[A \bullet B]] &= \{s_1 + s_2 \mid s_1 \in [[A]] \ \& \ s_2 \in [[B]]\} \end{aligned}$$

Such a formalized interpretation appears to have been first made explicit in Buskowski (1982)[?]. Observe that the (continuous) product \bullet inherits associativity from the basic prosodic algebra:

$$(4) \ A \bullet (B \bullet C) = (A \bullet B) \bullet C$$

Note also that $(\setminus, \bullet, /; \subseteq)$ constitutes a residuated triple, i.e.

$$(5) \ B \subseteq A \setminus C \quad \text{iff} \quad A \bullet B \subseteq C \quad \text{iff} \quad A \subseteq C / B$$

(6) **Definition** (*configurations and sequents of L*)

The set \mathcal{O} of *configurations* of \mathbf{L} is defined as follows:

$$\mathcal{O} ::= \mathcal{F} \mid \mathcal{F}, \mathcal{O}$$

The set Σ of *sequents* of \mathbf{L} is defined as follows:

$$\Sigma ::= \mathcal{O} \Rightarrow \mathcal{F}$$

\mathcal{O} is called the *antecedent* configuration and \mathcal{F} is called the *succedent* type.

(7) **Definition** (*prosodic interpretation of configurations and validity of sequents in L*)

We extend the interpretation of types to include configurations as follows:

$$[[A, \Gamma]] = \{s_1 + s_2 \mid s_1 \in [[A]] \ \& \ s_2 \in [[\Gamma]]\}$$

A sequent $\Gamma \Rightarrow A$ is *valid* iff $[[\Gamma]] \subseteq [[A]]$ in every interpretation.

(8) **Definition** (*sequent calculus for \mathbf{L}*)

The *sequent calculus* for \mathbf{L} is as follows, where $\Delta(\Gamma)$ indicates a configuration Δ with a distinguished subconfiguration Γ :

$$\begin{array}{c} \frac{}{A \Rightarrow A} \textit{id} \quad \frac{\Gamma \Rightarrow A \quad \Delta(A) \Rightarrow B}{\Delta(\Gamma) \Rightarrow B} \textit{Cut} \\ \\ \frac{\Gamma \Rightarrow A \quad \Delta(C) \Rightarrow D}{\Delta(\Gamma, A \setminus C) \Rightarrow D} \setminus L \quad \frac{A, \Gamma \Rightarrow C}{\Gamma \Rightarrow A \setminus C} \setminus R \\ \\ \frac{\Gamma \Rightarrow B \quad \Delta(C) \Rightarrow D}{\Delta(C/B, \Gamma) \Rightarrow D} /L \quad \frac{\Gamma, B \Rightarrow C}{\Gamma \Rightarrow C/B} /R \\ \\ \frac{\Delta(A, B) \Rightarrow D}{\Delta(A \bullet B) \Rightarrow D} \bullet L \quad \frac{\Gamma \Rightarrow A \quad \Delta \Rightarrow B}{\Gamma, \Delta \Rightarrow A \bullet B} \bullet R \end{array}$$

In \mathbf{L} , a *theorem* is a sequent which is derivable in this calculus.

Observe that in the sequent calculus for \mathbf{L} , for each connective there is a left (L) rule introducing it in the antecedent, and a right (R) rule introducing it in the succedent; the type in which this connective occurs is called *active*; the other types are called *side formulas*. These reflect respectively sufficient conditions for use, and necessary conditions for proof, of a type so-built. The sequent calculus fully modularizes the inferential properties of connectives: it deals with a single occurrence of a single connective at a time.

(9) **Proposition** (*soundness of \mathbf{L}*)

In \mathbf{L} , every theorem is valid.

Proof. Straightforward induction on the length of sequent proofs. \square

(10) **Theorem** (*Cut-elimination for \mathbf{L}*)

In \mathbf{L} , every theorem has a Cut-free sequent proof.

Proof. Lambek (1958)[?]. Where A is a type, let $d(A)$ be the number of separate occurrences of the connectives \setminus , \bullet , $/$ in A , and let $d(A_1, A_2, \dots, A_n) = d(A_1) + d(A_2) + \dots + d(A_n)$. The *degree* of an instance of Cut

$$\frac{\Gamma \Rightarrow A \quad \Delta_1, A, \Delta_2 \Rightarrow B}{\Delta_1, \Gamma, \Delta_2 \Rightarrow B}$$

is defined to be $d(\Gamma) + d(\Delta_1) + d(\Delta_2) + d(A) + d(B)$. It is shown that in any Cut the premises of which have been proved without Cut, the conclusion is either identical with one of the premises, or else the Cut can be replaced by one or two such Cuts of smaller degree. Therefore since no degree is negative, every theorem has a Cut-free proof. \square

(11) **Corollary** (*subformula property for \mathbf{L}*)

In \mathbf{L} , every theorem has a sequent proof containing only its subformulas.

Proof. Every rule except Cut has the property that all the types in the premises are either in the conclusion (side formulas) or are the immediate subtypes of the active formula, and Cut itself is eliminable. \square

(12) **Corollary** (*decidability of \mathbf{L}*)

In \mathbf{L} , it is decidable whether a sequent is a theorem.

Proof. By backward-chaining in the finite Cut-free sequent search space. \square

(13) **Theorem** (*completeness of \mathbf{L}*)

In \mathbf{L} , every valid sequent is a theorem.

Proof. By the sophisticated reasoning of Pentus (1993)[?], which goes via “quasi-models”. \square

1.2 Type-Logical Semantics

(14) **Definition** (*semantic types*)

The set \mathcal{T} of semantic types is defined on the basis of a set δ of primitive semantic types by:

$$\mathcal{T} ::= \delta \mid \mathcal{T} \& \mathcal{T} \mid \mathcal{T} \rightarrow \mathcal{T}$$

(15) **Definition** (*semantic frame*)

A *semantic frame* is a \mathcal{T} -indexed family of non-empty sets $\{D_\tau\}_{\tau \in \mathcal{T}}$ such that:

$$\begin{aligned} D_{\tau_1 \& \tau_2} &= D_{\tau_1} \times D_{\tau_2} && \text{cartesian product} \\ D_{\tau_1 \rightarrow \tau_2} &= D_{\tau_2}^{D_{\tau_1}} && \text{functional exponentiation} \end{aligned}$$

For example, we might select as basic types a type e of entities, $\delta(e)$ a nonempty set of individuals, and a basic type 2 of two truth values, $\delta(2) = \{\emptyset, \{\emptyset\}\}$.

(16) **Definition** (*semantic terms*)

The sets Φ_τ of *semantic terms* of type τ for each type τ are defined on the basis of a set C_τ of constants of type τ and an enumerably infinite set V_τ of variables of type τ for each type τ as follows:

$$\begin{aligned} \Phi_\tau &::= C_\tau \mid V_\tau \mid (\Phi_{\tau' \rightarrow \tau} \Phi_{\tau'}) \mid \pi_1 \Phi_{\tau \& \tau'} \mid \pi_2 \Phi_{\tau' \& \tau} \\ \Phi_{\tau \rightarrow \tau'} &::= \lambda V_\tau \Phi_{\tau'} \\ \Phi_{\tau \& \tau'} &::= (\Phi_\tau, \Phi_{\tau'}) \end{aligned}$$

(We allow ourselves to abbreviate $((\phi \ \psi) \ \chi)$ as $(\phi \ \psi \ \chi)$, etc.³) An occurrence of a variable x in a term is *free* iff it does not fall within any part of the term of the form λx ; otherwise it is *bound* (by the closest λx within the scope of which it falls). Each term $\phi \in \Phi_\tau$ receives a semantic value $[\phi]^g \in D_\tau$ with respect to a valuation f sending each constant in C_τ to an element in D_τ , and an assignment g sending each variable in V_τ to an element in D_τ , as follows:

³ Likewise, $((\phi, \psi), \chi)$ would be abbreviated (ϕ, ψ, χ) , etc.

(17)	$[c]^g = f(c)$ $[x]^g = g(c)$ $[(\phi \ \psi)]^g = [\phi]^g([\psi]^g)$ $[\pi_1 \phi]^g = \mathbf{fst}([\phi]^g)$ $[\pi_2 \phi]^g = \mathbf{snd}([\phi]^g)$ $[\lambda x_\tau \phi]^g = D_\tau \ni d \mapsto [\phi]^{(g - \{(x, g(x))\}) \cup \{(x, d)\}}$ $[(\phi, \psi)]^g = \langle [\phi]^g, [\psi]^g \rangle$	for $c \in C_\tau$ for $x \in V_\tau$ functional application first projection second projection functional abstraction ordered pair formation
------	--	---

The result $\phi\{\psi/x\}$ of substituting term ψ (of type τ) for variable x (of type τ) in a term ϕ is the result of replacing by ψ every free occurrence of x in ϕ . The application of the substitution is *free* iff no variable in ψ is bound in its new location. (Manipulations can be pathological if substitution is not free.) The following laws of lambda-conversion obtain:

$$(18) \quad \lambda y \phi = \lambda x (\phi\{x/y\}) \text{ if } x \text{ is not free in } \phi \text{ and } \phi\{x/y\} \text{ is free}$$

α -conversion

$$\begin{aligned} (\lambda x \phi \ \psi) &= \phi\{\psi/x\} && \text{if } \phi\{\psi/x\} \text{ is free} \\ \pi_1(\phi, \psi) &= \phi \\ \pi_2(\phi, \psi) &= \psi \end{aligned}$$

β -conversion

$$\begin{aligned} \lambda x (\phi \ x) &= \phi && \text{if } x \text{ is not free in } \phi \\ (\pi_1 \phi, \pi_2 \phi) &= \phi \end{aligned}$$

η -conversion

(19) **Definition** (*semantic type map for \mathbf{L}*)

The *semantic type map* for \mathbf{L} is a homomorphism T from syntactic types \mathcal{F} to semantic types \mathcal{T} such that:

$$\begin{aligned} T(A \bullet B) &= T(A) \& T(B) \\ T(A \setminus C) &= T(A) \rightarrow T(C) \\ T(C/B) &= T(B) \rightarrow T(C) \end{aligned}$$

Categorical semantics, Curry-Howard type-logical semantics, works because under such a type map categorical derivations are homomorphically sent to intuitionistic proofs, i.e. pure terms of the typed lambda calculus. These compose lexical semantics expressed as terms of higher-order logic into meanings in higher-order logic of projected expressions. Montague (1970)[?] observed that algebraically, compositionality is a homomorphism from syntax to semantics. TLG goes further in asserting that it is a homomorphism from syntactic *proofs* to semantic *proofs*.

1.3 Evolution of Discontinuity in Categorical Grammar

The idea of discontinuity operators for categorical grammar appears to originate in Bach (1981, 1984)[?][?]. Where $s = a_1 + \dots + a_n$ is the factorization of s into primes, let us define:

$$(20) \begin{aligned} FIRST(s) &=_{df} a_1 \\ RREST(s) &=_{df} a_2 + \dots + a_n \\ LAST(s) &=_{df} a_n \\ LREST(s) &=_{df} a_1 + \dots + a_{n-1} \end{aligned}$$

Bach (1984)[?] defined the operations *RWRAP* and *LWRAP*, and their converses *LINFIX* and *RINFIX* respectively, as follows:

$$(21) \begin{aligned} RWRAP(s_1, s_2) &= LINFIX(s_2, s_1) = FIRST(s_1) + s_2 + RREST(s_1) \\ LWRAP(s_1, s_2) &= RINFIX(s_2, s_1) = LREST(s_1) + s_2 + LAST(s_1) \end{aligned}$$

Bach had in mind such applications as a characterization of the object equi *persuade*/subject equi *promise* distinction in terms of alternative argument order, but here we would assume a coding of control properties in lexical semantics. He also proposed ‘long-distance’ functors in relation to Dutch word order, which we will address, but in terms of wrapping.

The first type-logical formulation of discontinuity, i.e. with an interpretation of types and with a sequent calculus, appeared in Moortgat (1988)[?]. Moortgat defined discontinuous types as follows (we modify his notation):

$$(22) \mathcal{F} ::= \mathcal{F} \downarrow_{\forall} \mathcal{F} \mid \mathcal{F} \uparrow_{\exists} \mathcal{F}$$

$$(23) \begin{aligned} [[A \downarrow_{\forall} C]] &= \{s \mid \forall s_1 + s_2 \in [[A]], s_1 + s + s_2 \in [[C]]\} \\ [[C \uparrow_{\exists} B]] &= \{s \mid \exists s_1, s_2, s = s_1 + s_2 \ \& \ \forall s' \in [[B]], s_1 + s' + s_2 \in [[C]]\} \end{aligned}$$

The following sequent rules were given:

$$(24) \frac{\Gamma \Rightarrow A \quad \Delta(C) \Rightarrow D}{\Delta(\Gamma(A \downarrow_{\forall} C)) \Rightarrow D} \downarrow_{\forall} L \quad \frac{\Gamma, B, \Delta \Rightarrow C}{\Gamma, \Delta \Rightarrow C \uparrow_{\exists} B} \uparrow_{\exists} R$$

Thus e.g. medial extraction, not otherwise derivable in the Lambek calculus, is obtained from a relative pronoun type $R/(S \uparrow_{\exists} N)$. And $S(\text{neg}) \downarrow_{\forall} S(\text{pos})$ would be the type of a freely floating negation particle, if there were really such an element. However, the other sequent rules cannot be formulated, so the logic is incomplete.⁴

Moortgat (1991)[?] ⁵ defined a three-place in-situ binder type-constructor Q for e.g. quantifier phrases, $Q(S, N, S)$, and subject-oriented reflexives, $Q(N \setminus S, N, N \setminus S)$. The left sequent rule is:

$$(25) \frac{\Gamma(A) \Rightarrow B \quad \Delta(C) \Rightarrow D}{\Delta(\Gamma(Q(B, A, C))) \Rightarrow D} QL$$

However the best that came be managed on the right is:

$$(26) \frac{\Gamma \Rightarrow A}{\Gamma \Rightarrow Q(B, A, B)} QR$$

⁴ We resolve this by decomposing Moortgat’s connectives into ones for which both rules of proof and use can be given, as follows: $A \downarrow_{\forall} C = \sim A \downarrow C$ and $C \uparrow_{\exists} B = \sim(C \uparrow B)$.

⁵ Moortgat (1991)[?] also proposed a substring product:

$$(i) [[A \odot_{\exists} B]] = \{s_1 + s_2 + s_3 \mid s_1 + s_3 \in [[A]] \ \& \ s_2 \in [[B]]\}$$

$$(ii) \frac{\Gamma_1, \Gamma_2 \Rightarrow A \quad \Delta \Rightarrow B}{\Gamma_1, \Delta, \Gamma_2 \Rightarrow A \odot_{\exists} B} \odot_{\exists} R$$

But again a left rule cannot be given. We resolve this by decomposing thus: $A \odot_{\exists} B = \sim A \odot B$.

This is insufficient to derive e.g. $Q(S, N, S) \Rightarrow Q(N \setminus S, N, N \setminus S)$ (that a quantifier phrase can occur in a verb phrase conjunct, H. Hendriks, p.c.) so the logic is incomplete again. Moortgat (1991)[?] indicated that $Q(B, A, C)$ might be decomposed into something like $(B \uparrow \exists A) \downarrow \vee C$, but he did not have a calculus ensuring that the two points of discontinuity would be one and the same, as is required in order to ensure, for example, that a quantifier phrase only binds the position it occupies.⁶

Versmissen (1991)[?] observed that we want in some way to mark points of discontinuity. Algebraic formulations, developed without knowledge of the head grammars of Pollard (1984)[?], were as follows:

- Solias (1992)[?]: prosodic algebra $(L, +, 0, \langle \cdot, \cdot \rangle)$ where $(L, +, 0)$ is a free monoid and $(L, \langle \cdot, \cdot \rangle)$ is a free groupoid. Wrap was a partial operation defined by $\langle s_1, s_3 \rangle W s_2 =_{df} s_1 + s_2 + s_3$.
- Morrill and Solias (1993)[?]: prosodic algebra $(L, +, 0, \langle \cdot, \cdot \rangle, 1, 2)$ where $(L, +, 0)$ is a monoid, $(L, \langle \cdot, \cdot \rangle)$ is a groupoid and $1 \langle s_1, s_2 \rangle = s_1, 2 \langle s_1, s_2 \rangle = s_2$ and $\langle 1s, 2s \rangle = s$. Wrap was a total operation defined by $s W s' =_{df} 1s + s' + 2s$.
- Morrill (1994, ch. 4; 1995)[?][?]: prosodic algebra $(L, +, 0, \langle \cdot, \cdot \rangle, W)$ where $(L, +, 0)$ is a monoid, $(L, \langle \cdot, \cdot \rangle)$ and (L, W) are groupoids and there is the structural rule of interaction $\langle s_1, s_3 \rangle W s_2 = s_1 + s_2 + s_3$. Wrap was a primitive total operation.

In Solias (1992)[?] wrapping was derived and partial. In Morrill and Solias (1993)[?] it was derived and total. In Morrill (1994, ch. 4; 1995)[?][?] it was primitive and total. But in all three cases, the representation of discontinuous expressions in an (unsorted) algebra introduced (infinitely many) prosodic terms in which points of discontinuity, because embedded, could necessarily never wrap, e.g. $s_1 + (s_2, s_3)$, so the prosodic ontology contained much useless junk.

Morrill and Merenciano (1996)[?] cleared this up admitting only n -tuples of strings sorted by their arity. But in the generalized case (i.e. with no upper bound on the number of points of discontinuity), both pairing and the empty tuple would be required for the construction of unboundedly long tuples. Here we reduce the machinery to a single operator of arity zero (i.e. a constant); cf. also Moortgat (1996)[?] for a constant operator, but here we use it to get the generalized case of discontinuity. The nullary operator is in general internal to concatenation, but whenever it is embedded it can be considered (by virtue of the associativity of concatenation) to be immediately embedded, and as such, always useful to undergo wrap.

2 Theory of Discontinuous Lambek Calculus

The key to our treatment of discontinuity is the notion of a “separator” (Morrill 2002)[?]:

(27) **Definition** (*graded prosodic algebra*)

A *graded prosodic algebra* is a free algebra $(L, +, 0, 1)$ of arity $(2, 0, 0)$ such that $(L, +, 0)$ is a monoid and 1 is a prime. I.e. L is a set, $0 \in L$ and $+$ is a binary operation on L such that for all $s_1, s_2, s_3, s \in L$,

$$\begin{aligned} s_1 + (s_2 + s_3) &= (s_1 + s_2) + s_3 && \text{associativity} \\ 0 + s &= s = s + 0 && \text{identity} \end{aligned}$$

⁶ We resolve this by realizing exactly the decomposition $Q(B, A, C) = (B \uparrow A) \downarrow C$.

The distinguished constant 1 is called a *separator*.

(28) **Definition** (*sorts*)

The *sorts* of discontinuous Lambek calculus are the naturals $0, 1, \dots$. The sort $\sigma(s)$ of an element s of a graded prosodic algebra $(L, +, 0, 1)$ is defined by the morphism of monoides σ to the additive monoid of naturals defined thus:

$$\begin{aligned}\sigma(1) &= 1 \\ \sigma(a) &= 0 && \text{for a prime } a \neq 1 \\ \sigma(s_1 + s_2) &= \sigma(s_1) + \sigma(s_2)\end{aligned}$$

I.e. the sort of a prosodic element is simply the number of separators it contains; we require the separator 1 to be a prime and the graded prosodic algebra to be free in order to ensure that this induction is well-defined. The fact that there is a homomorphism from a graded prosodic algebra to the additive monoid of naturals means that a graded prosodic algebra is an instance of what is known as a *graded algebra*, in particular a *graded monoid*.

(29) **Definition** (*sort domains*)

Where $(L, +, 0, 1)$ is a graded prosodic algebra, the *sort domains* L_i of sort i of generalized discontinuous Lambek calculus are defined as follows:

$$L_i = \{s \mid \sigma(s) = i\}, i \geq 0$$

(30) **Definition** (*discontinuous prosodic structure*)

The *discontinuous prosodic structure* defined by a graded prosodic algebra $(L, +, 0, 1)$ is the ω -sorted structure

$$(\{L_i\}_{i \in \mathcal{N}}, \iota_1, \iota_2, \{U_k\}_{k \in \mathcal{Z}^+}, +, \{W_k\}_{k \in \mathcal{Z}^+}; U, W)$$

where:

operation or relation	is such that
$\iota_1 : L_i \rightarrow L_{i+1}$	$\iota_1(s) = s+1$
$\iota_2 : L_i \rightarrow L_{i+1}$	$\iota_2(s) = 1+s$
$U_k : L_{i+1} \rightarrow L_i$	$U_k(s)$ is the result of replacing the k -th separator in s by 0
$+ : L_i \times L_j \rightarrow L_{i+j}$	as in the graded prosodic algebra
$W_k : L_{i+1} \times L_j \rightarrow L_{i+j}$	$W_k(s, t)$ is the result of replacing the k -th separator in s by t
$U : L_{i+1} \times L_i$	it is the smallest relation such that $U(s_1+1+s_2, s_1+s_2)$
$W : L_{i+1} \times L_j \times L_{i+j}$	it is the smallest relation such that $W(s_1+1+s_3, s_2, s_1+s_2+s_3)$

The types of discontinuous Lambek calculus are to be interpreted as subsets of L . The connectives and their prosodic interpretations are shown in Figures 1 and 2. Versions of the unary operators $\hat{}$ ('bridge') and $\check{}$ ('split') were introduced in Morrill and Merenciano (1996)[?]. Nondeterministic binary discontinuity connectives were given in Morrill (2002)[?] and Morrill, Fadda and Valentín (2007)[?]. Given the functionalities and relationalities of the operations and relations with respect to which the connectives are defined, there is a fixed pattern between the types and the sorts. Figures 3 and 4

$[[\triangleleft A]] = \{\iota_1(s) \mid s \in [[A]]\}$	left injection
$[[\triangleleft^{-1} B]] = \{s \mid \iota_1(s) \in [[B]]\}$	left projection
$[[\triangleright A]] = \{\iota_2(s) \mid s \in [[A]]\}$	right injection
$[[\triangleright^{-1} B]] = \{s \mid \iota_2(s) \in [[B]]\}$	right projection
$[[\overset{\wedge}{\sim} A]] = \{U_k(s) \mid s \in [[A]]\}$ deterministic bridge	$k > 0$
$[[\overset{\sim}{\sim} B]] = \{s \mid U_k(s) \in [[B]]\}$ deterministic split	$k > 0$
$[[\overset{\wedge}{\sim} A]] = \{s \mid \exists s_1 \in [[A], U(s_1, s)]\}$ nondeterministic bridge	
$[[\overset{\sim}{\sim} B]] = \{s_1 \mid \forall s, U(s_1, s) \Rightarrow s \in [[B]]\}$ nondeterministic split	

Fig. 1 Prosodic interpretation of ω -DL types, part I

gives both the grammar defining by mutual recursion the sets \mathcal{F}_i of types of sort i for each sort i on the basis of sets \mathcal{P}_i of primitive types of sort i for each sort i , and the homomorphic *prosodic sort map* S sending types to their sorts. The prosodic sort map is to prosodics what the semantic type map is to semantics: both homomorphisms mapping syntactic types to the datatypes of the respective components of their inhabiting signs in the dimensions of language: prosodic sort for form/signifier and semantic type for meaning/signified.⁷

Observe that (modulo sorting) $(\triangleleft^{-1}, \triangleleft)$, $(\triangleright^{-1}, \triangleright)$, $(\overset{\wedge}{\sim}, \overset{\sim}{\sim})$ and $(\overset{\wedge}{\sim}, \overset{\sim}{\sim})$ are residuated pairs with respect to \subseteq :

$$(31) \quad \begin{aligned} A \subseteq \triangleleft^{-1} B &\text{ iff } \triangleleft A \subseteq B \\ A \subseteq \triangleright^{-1} B &\text{ iff } \triangleright A \subseteq B \\ A \subseteq \overset{\wedge}{\sim} B &\text{ iff } \overset{\sim}{\sim} A \subseteq B \\ A \subseteq \overset{\sim}{\sim} B &\text{ iff } \overset{\wedge}{\sim} A \subseteq B \end{aligned}$$

Observe also that (modulo sorting) $(\setminus, \bullet, /; \subseteq)$, $(\downarrow_k, \odot_k, \uparrow_k; \subseteq)$ and $(\downarrow, \odot, \uparrow; \subseteq)$ are residuated triples:

$$(32) \quad \begin{aligned} B \subseteq A \setminus C &\text{ iff } A \bullet B \subseteq C \text{ iff } A \subseteq C / B \\ B \subseteq A \downarrow_k C &\text{ iff } A \odot_k B \subseteq C \text{ iff } A \subseteq C \uparrow_k B \\ B \subseteq A \downarrow C &\text{ iff } A \odot B \subseteq C \text{ iff } A \subseteq C \uparrow B \end{aligned}$$

⁷ If we had continuous and discontinuous product units $I = \{0\}$ and $J = \{1\}$ we could define away all the unary connectives by just the two nullary connectives as follows:

$$\begin{aligned} \triangleleft^{-1} A &= A / J & \triangleleft A &= A \bullet J & \triangleright^{-1} A &= J \setminus A & \triangleright A &= J \bullet A \\ \overset{\wedge}{\sim} A &= A \uparrow_k I & \overset{\sim}{\sim} A &= A \odot_k I & \overset{\wedge}{\sim} A &= A \uparrow I & \overset{\sim}{\sim} A &= A \odot I \end{aligned}$$

We do this as a means of obtaining the technical result of Cut-elimination in the Appendix, but there are four reasons why we do not go down this path in our official discontinuity calculus: 1) we would require a dummy semantics for the units, 2) lexical assignments to units can challenge decidability of recognition, 3) proof nets for units are problematic, and 4) there would appear to be problems of incompleteness.

$[[A \bullet B]] = \{s_1 + s_2 \mid s_1 \in [[A]] \ \& \ s_2 \in [[B]]\}$	(continuous) product
$[[A \setminus C]] = \{s_2 \mid \forall s_1 \in [[A]], s_1 + s_2 \in [[C]]\}$	under
$[[C / B]] = \{s_1 \mid \forall s_2 \in [[B]], s_1 + s_2 \in [[C]]\}$	over
$[[A \circ_k B]] = \{W_k(s_1, s_2) \mid s_1 \in [[A]] \ \& \ s_2 \in [[B]]\}$	$k > 0$
deterministic discontinuous product	
$[[A \downarrow_k C]] = \{s_2 \mid \forall s_1 \in [[A]], W_k(s_1, s_2) \in [[C]]\}$	$k > 0$
deterministic infix	
$[[C \uparrow_k B]] = \{s_1 \mid \forall s_2 \in [[B]], W_k(s_1, s_2) \in [[C]]\}$	$k > 0$
deterministic extract	
$[[A \circ B]] = \{s \mid \exists s_1 \in [[A]] \ \& \ \exists s_2 \in [[B]], W(s_1, s_2, s)\}$	
nondeterministic discontinuous product	
$[[A \downarrow C]] = \{s_2 \mid \forall s_1 \in [[A]], \forall s, W(s_1, s_2, s) \Rightarrow s \in [[C]]\}$	
nondeterministic infix	
$[[C \uparrow B]] = \{s_1 \mid \forall s_2 \in [[B]], \forall s, W(s_1, s_2, s) \Rightarrow s \in [[C]]\}$	
nondeterministic extract	

Fig. 2 Prosodic interpretation of ω -DL types, part II

$$\begin{aligned}
\mathcal{F}_i &::= \mathcal{P}_i & S(A) &= i & \text{for } A \in \mathcal{P}_i \\
\mathcal{F}_{i+1} &::= \triangleleft \mathcal{F}_i & S(\triangleleft A) &= S(A) + 1 \\
\mathcal{F}_i &::= \triangleleft^{-1} \mathcal{F}_{i+1} & S(\triangleleft^{-1} B) &= S(B) - 1 \\
\mathcal{F}_{i+1} &::= \triangleright \mathcal{F}_i & S(\triangleright A) &= S(A) + 1 \\
\mathcal{F}_i &::= \triangleright^{-1} \mathcal{F}_{i+1} & S(\triangleright^{-1} B) &= S(B) - 1 \\
\mathcal{F}_i &::= \hat{\ }^k \mathcal{F}_{i+1} & S(\hat{\ }^k A) &= S(A) - 1 \quad 1 \leq k \leq i + 1 \\
\mathcal{F}_{i+1} &::= \hat{\ }^k \mathcal{F}_i & S(\hat{\ }^k A) &= S(A) + 1 \quad 1 \leq k \leq i + 1 \\
\mathcal{F}_i &::= \hat{\ } \mathcal{F}_{i+1} & S(\hat{\ } A) &= S(A) - 1 \\
\mathcal{F}_{i+1} &::= \hat{\ } \mathcal{F}_i & S(\hat{\ } A) &= S(A) + 1
\end{aligned}$$

Fig. 3 Sorted ω -DL types and prosodic sort map for ω -DL, part I

$$\begin{aligned}
\mathcal{F}_{i+j} &::= \mathcal{F}_i \bullet \mathcal{F}_j & S(A \bullet B) &= S(A) + S(B) \\
\mathcal{F}_j &::= \mathcal{F}_i \setminus \mathcal{F}_{i+j} & S(A \setminus C) &= S(C) - S(A) \\
\mathcal{F}_i &::= \mathcal{F}_{i+j} / \mathcal{F}_j & S(C / B) &= S(C) - S(B) \\
\mathcal{F}_{i+j} &::= \mathcal{F}_{i+1} \circ_k \mathcal{F}_j & S(A \circ_k B) &= S(A) + S(B) - 1 \quad 1 \leq k \leq i + 1 \\
\mathcal{F}_j &::= \mathcal{F}_{i+1} \downarrow_k \mathcal{F}_{i+j} & S(A \downarrow_k C) &= S(C) + 1 - S(A) \quad 1 \leq k \leq i + 1 \\
\mathcal{F}_{i+1} &::= \mathcal{F}_{i+j} \uparrow_k \mathcal{F}_j & S(C \uparrow_k B) &= S(C) + 1 - S(B) \quad 1 \leq k \leq i + 1 \\
\mathcal{F}_{i+j} &::= \mathcal{F}_{i+1} \circ \mathcal{F}_j & S(A \circ B) &= S(A) + S(B) - 1 \\
\mathcal{F}_j &::= \mathcal{F}_{i+1} \downarrow \mathcal{F}_{i+j} & S(A \downarrow C) &= S(C) + 1 - S(A) \\
\mathcal{F}_{i+1} &::= \mathcal{F}_{i+j} \uparrow \mathcal{F}_j & S(C \uparrow B) &= S(C) + 1 - S(B)
\end{aligned}$$

Fig. 4 Sorted ω -DL types and prosodic sort map for ω -DL, part II

2.1 Hypersequent Calculus for ω -DL

We define the *components* of a prosodic object as its maximal subparts not containing 1. Morrill (1997)[?] introduced sequent calculus for (sorted) discontinuity in which a single discontinuous type has multiple manifestations at the loci of its expressions' components, punctuated by surds. This is called 'hypersequent calculus' in the appendix

of Morrill (2003)[?], though in a usage of the term distinct from that of A. Avron. The spirit is to maintain everything in “evaluated/spelt-out” linearized form.

The surd notation is meant to be suggestive of the (commutative) numeric law:

$$(33) \underbrace{\sqrt[i]{A} \times \cdots \times \sqrt[i]{A}}_{i \text{ times}} = A$$

For us, non-commutatively:⁸

$$(34) \sqrt[0]{A} \bullet \{1\} \cdots \{1\} \bullet {}^{S(A)}\sqrt{A} = A$$

(35) **Definition** (*figures, configurations and hypersequents of hypersequent calculus*)

In hypersequent calculus the *figures* \mathcal{Q}_i of sort i for each sort i are defined as follows (\llbracket is our *metalinguistic separator*):

$$\begin{aligned} \mathcal{Q}_0 &::= A && \text{for } S(A) = 0 \\ \mathcal{Q}_{S(A)} &::= \sqrt[0]{A}, \llbracket, \sqrt[1]{A}, \dots, {}^{S(A)-1}\sqrt{A}, \llbracket, {}^{S(A)}\sqrt{A} && \text{for } S(A) > 0 \end{aligned}$$

By the vectorial notation \vec{A} we mean the figure of sorted type A , i.e.

$$\vec{A} =_{df.} \begin{cases} A & \text{if } S(A) = 0 \\ \sqrt[0]{A}, \llbracket, \sqrt[1]{A}, \dots, {}^{S(A)-1}\sqrt{A}, \llbracket, {}^{S(A)}\sqrt{A} & \text{if } S(A) > 0 \end{cases}$$

The *configurations* \mathcal{O}_i of sort i for each sort i are defined unambiguously by mutual recursion as follows, where Λ is the empty string:

$$\begin{aligned} \mathcal{O}_0 &::= \Lambda \\ \mathcal{O}_i &::= A, \mathcal{O}_i \text{ for } S(A) = 0 \\ \mathcal{O}_{i+1} &::= \llbracket, \mathcal{O}_i \\ \mathcal{O}_{\sum_{k=0}^{S(A)} j_k} &::= \sqrt[0]{A}, \mathcal{O}_{j_0}, \sqrt[1]{A}, \dots, {}^{S(A)-1}\sqrt{A}, \mathcal{O}_{j_{S(A)-1}}, {}^{S(A)}\sqrt{A}, \mathcal{O}_{j_{S(A)}} \\ &\text{for } S(A) > 0 \end{aligned}$$

Note that not every substring of a configuration is a (well-formed) configuration because as well as containing all the segments of discontinuous types, these segments must be separated by correct configurations. We define the *components* of a

⁸ Since elements of graded prosodic algebras are in bijection with tuples, it could also be reasonable to punctuate the components of discontinuous types in hypersequents with projections $\pi_i A$. However, in TLG with bracket operators for domains (Morrill 1992)[?], it seems that eventually we will need to allow separators within bracketed domains (e.g. $s_1 + b(s_2 + 1 + s_3) + s_4$, e.g. for quantifier phrases which outscope *wh*-islands). In this case, ‘components’ would not be projective since they would not always correspond to elements of the (bracketed graded) prosodic algebra. It seems we will need surded subparts of configurations which are not well-formed terms of the configuration algebra and do not denote well-formed prosodic objects, but only ‘parts’ of them; for example, containing the left boundary of a unary operation of bracketing, but not the right boundary. And sorts themselves would seem to have to be extended to say bracketed sequences of 1’s to control for when a separator is or is not within a bracketed domain and thus is not or is available to some mode of discontinuity.

Therefore we prefer to keep the surd notation, which does not seem to imply as much as would a projective notation that there will always be projectivity, cf. that $\sqrt{-1}$ is an *imaginary* number though not a *real* number: in the complex number system a general n th degree equation has exactly n roots; for us, a prosodic object of sort i has $i + 1$ *components*. It seems that for brackets, which can perhaps be seen as a kind of negation, we could need a noncommutative monoidal analogue of the complex numbers defining roots of negative numbers.

configuration as its maximal substrings not containing the metalinguistic separator \llbracket (components *are* correct configurations).

The *hypersequents* Σ_i of sort i for each sort i are defined as follows:

$$\Sigma_i ::= \mathcal{O}_i \Rightarrow \mathcal{Q}_i$$

\mathcal{O}_i is called the *antecedent* configuration and \mathcal{Q}_i is called the *succedent* figure.

Observe that the components of discontinuous types are well-nested in configurations, i.e. that there are no crossing discontinuities, so that in configurations the enumeration of components is sufficient to define their dependencies. This corresponds to the fact that under wrapping, the infix is always kept intact within the circumfix. With a ‘shuffle’ discontinuous operation this would no longer be true, and components which belong together would need to be coindexed in some way, as in Morrill (2003)[?].

(36) **Definition** (*prosodic interpretation of configurations and validity of sequents in hypersequent calculus*)

In hypersequent calculus we extend the interpretation of types to include configurations, as follows:

$$\begin{aligned} \llbracket[A] \rrbracket &= \{0\} \\ \llbracket[A, \Gamma] \rrbracket &= \{s_1 + s_2 \mid s_1 \in \llbracket[A] \rrbracket \ \& \ s_2 \in \llbracket[\Gamma] \rrbracket\} \\ \llbracket[\llbracket, \Gamma] \rrbracket &= \{1 + s \mid s \in \llbracket[\Gamma] \rrbracket\} \\ \llbracket[\sqrt[0]{A}, \Gamma_0, \dots, \Gamma_{S(A)-1}, \sqrt[S(A)]{A}, \Gamma_{S(A)}] \rrbracket &= \{s_0 + t_0 + \dots + t_{S(A)-1} + s_{S(A)} + t_{S(A)} \mid \\ &\quad s_0 + 1 + \dots + 1 + s_{S(A)} \in \llbracket[A] \rrbracket \\ &\quad \& \ t_j \in \llbracket[\Gamma_j] \rrbracket, 0 \leq j \leq S(A)\} \end{aligned}$$

A hypersequent $\Gamma \Rightarrow X$ is *valid* iff in every interpretation, $\llbracket[\Gamma] \rrbracket \subseteq \llbracket[X] \rrbracket$.

The hypersequent calculus for ω -DL is given in Figures 5 and 6. $\Delta(\Gamma)$ means a configuration Δ in which in some distinguished positions the components of Γ appear in the given order and such that parts of Δ appearing between components of Γ are well-formed configurations. $\Delta|_k \Gamma, k > 0$ is the result of replacing the k -th separator in Δ by Γ .

Observe that the interpretation of our distinguished occurrence notation is such that the rules for continuous connectives in hypersequent calculus look just like those of the original Lambek calculus, but with the vectorial notation on the active types. Observe also that the rules for the deterministic discontinuous connectives in hypersequent calculus look just like the rules for the continuous connectives, but with metalinguistic wrapping ‘ \llbracket ’ instead of metalinguistic concatenation ‘ $,$ ’. We consider that these symmetries give some of the substance to our claim that discontinuous Lambek calculus is a natural generalisation of (continuous) Lambek calculus. But unlike with the case of deterministic discontinuity, the rules for nondeterministic discontinuity no longer follow exactly the same pattern as those for continuity because nondeterministic wrapping is no longer functional but only relational. There are an infinite number of rule schemata in the calculus since the number of premises is unbounded in $\Downarrow R$, $\Uparrow R$ and $\odot L$, though every instance is finite, and (Cut-free) only a finite number of instances can apply in derivations from a given (finite) lexicon.

(37) **Proposition** (*soundness of ω -DL*)

In ω -DL, every theorem is valid.

$$\begin{array}{c}
\frac{}{\vec{A} \Rightarrow \vec{A}} id \quad \frac{\Gamma \Rightarrow \vec{A} \quad \Delta(\vec{A}) \Rightarrow \vec{B}}{\Delta(\Gamma) \Rightarrow \vec{B}} Cut \\
\\
\frac{\Gamma(\vec{A}) \Rightarrow \vec{B}}{\Gamma(\overleftarrow{\neg}^{-1}\vec{A}, \square) \Rightarrow \vec{B}} \overleftarrow{\neg}^{-1}L \quad \frac{\Gamma, \square \Rightarrow \vec{A}}{\Gamma \Rightarrow \overleftarrow{\neg}^{-1}\vec{A}} \overleftarrow{\neg}^{-1}R \\
\\
\frac{\Gamma(\vec{A}, \square) \Rightarrow \vec{B}}{\Gamma(\overleftarrow{\neg}\vec{A}) \Rightarrow \vec{B}} \overleftarrow{\neg}L \quad \frac{\Gamma \Rightarrow \vec{A}}{\Gamma, \square \Rightarrow \overleftarrow{\neg}\vec{A}} \overleftarrow{\neg}R \\
\\
\frac{\Gamma(\vec{A}) \Rightarrow \vec{B}}{\Gamma(\square, \overrightarrow{\neg}^{-1}\vec{A}) \Rightarrow \vec{B}} \overrightarrow{\neg}^{-1}L \quad \frac{\square, \Gamma \Rightarrow \vec{A}}{\Gamma \Rightarrow \overrightarrow{\neg}^{-1}\vec{A}} \overrightarrow{\neg}^{-1}R \\
\\
\frac{\Gamma(\square, \vec{A}) \Rightarrow \vec{B}}{\Gamma(\overrightarrow{\neg}\vec{A}) \Rightarrow \vec{B}} \overrightarrow{\neg}L \quad \frac{\Gamma \Rightarrow \vec{A}}{\square, \Gamma \Rightarrow \overrightarrow{\neg}\vec{A}} \overrightarrow{\neg}R \\
\\
\frac{\Delta(\vec{B}) \Rightarrow \vec{C}}{\Delta(\overleftarrow{\neg}^k\vec{B}|_kA) \Rightarrow \vec{C}} \overleftarrow{\neg}^kL \quad \frac{\Delta|_kA \Rightarrow \vec{B}}{\Delta \Rightarrow \overleftarrow{\neg}^k\vec{B}} \overleftarrow{\neg}^kR \\
\\
\frac{\Delta(\vec{A}|_kA) \Rightarrow \vec{C}}{\Delta(\overleftarrow{\neg}^k\vec{A}) \Rightarrow \vec{C}} \overleftarrow{\neg}^kL \quad \frac{\Delta \Rightarrow \vec{A}}{\Delta|_kA \Rightarrow \overleftarrow{\neg}^k\vec{A}} \overleftarrow{\neg}^kR \\
\\
\frac{\Delta(\vec{B}) \Rightarrow \vec{C}}{\Delta(\overleftarrow{\neg}^k\vec{B}|_kA) \Rightarrow \vec{C}} \overleftarrow{\neg}L \quad \frac{\Delta|_1A \Rightarrow \vec{B} \quad \dots \quad \Delta|_{S(B)}A \Rightarrow \vec{B}}{\Delta \Rightarrow \overleftarrow{\neg}^k\vec{B}} \overleftarrow{\neg}R \\
\\
\frac{\Delta(\vec{A}|_1A) \Rightarrow \vec{C} \quad \dots \quad \Delta(\vec{A}|_{S(A)}A) \Rightarrow \vec{C}}{\Delta(\overleftarrow{\neg}^k\vec{A}) \Rightarrow \vec{C}} \overleftarrow{\neg}L \quad \frac{\Delta \Rightarrow \vec{A}}{\Delta|_kA \Rightarrow \overleftarrow{\neg}^k\vec{A}} \overleftarrow{\neg}R
\end{array}$$

Fig. 5 Hypersequent calculus for ω -DL, part I

Proof. By induction on the length of proofs. \square

(38) **Theorem** (*Cut-elimination for ω -DL*)

In ω -DL, every theorem has a Cut-free hypersequent proof.

Proof. See the Appendix. \square

(39) **Corollary** (*subformula property for ω -DL*)

In ω -DL, every theorem has a hypersequent proof containing only its subformulas.

Proof. Every rule except Cut has the property that all the types in the premises are either in the conclusion (side formulas) or are the immediate subtypes of the active formula, and Cut itself is eliminable. \square

(40) **Corollary** (*decidability of ω -DL*)

In ω -DL, it is decidable whether a hypersequent is a theorem.

Proof. By backward-chaining in the finite Cut-free hypersequent search space. \square

$$\begin{array}{c}
\frac{\Gamma \Rightarrow \vec{A} \quad \Delta(\vec{C}) \Rightarrow \vec{D}}{\Delta(\Gamma, \vec{A} \backslash \vec{C}) \Rightarrow \vec{D}} \backslash L \quad \frac{\vec{A}, \Gamma \Rightarrow \vec{C}}{\Gamma \Rightarrow \vec{A} \backslash \vec{C}} \backslash R \\
\frac{\Gamma \Rightarrow \vec{B} \quad \Delta(\vec{C}) \Rightarrow \vec{D}}{\Delta(\vec{C} / \vec{B}, \Gamma) \Rightarrow \vec{D}} /L \quad \frac{\Gamma, \vec{B} \Rightarrow \vec{C}}{\Gamma \Rightarrow \vec{C} / \vec{B}} /R \\
\frac{\Delta(\vec{A}, \vec{B}) \Rightarrow \vec{D}}{\Delta(\vec{A} \bullet \vec{B}) \Rightarrow \vec{D}} \bullet L \quad \frac{\Gamma_1 \Rightarrow \vec{A} \quad \Gamma_2 \Rightarrow \vec{B}}{\Gamma_1, \Gamma_2 \Rightarrow \vec{A} \bullet \vec{B}} \bullet R \\
\frac{\Gamma \Rightarrow \vec{A} \quad \Delta(\vec{C}) \Rightarrow \vec{D}}{\Delta(\Gamma|_k \vec{A} \downarrow_k \vec{C}) \Rightarrow \vec{D}} \downarrow_k L \quad \frac{\vec{A}|_k \Gamma \Rightarrow \vec{C}}{\Gamma \Rightarrow \vec{A} \downarrow_k \vec{C}} \downarrow_k R \\
\frac{\Gamma \Rightarrow \vec{B} \quad \Delta(\vec{C}) \Rightarrow \vec{D}}{\Delta(\vec{C} \uparrow_k \vec{B}|_k \Gamma) \Rightarrow \vec{D}} \uparrow_k L \quad \frac{\Gamma|_k \vec{B} \Rightarrow \vec{C}}{\Gamma \Rightarrow \vec{C} \uparrow_k \vec{B}} \uparrow_k R \\
\frac{\Delta(\vec{A}|_k \vec{B}) \Rightarrow \vec{D}}{\Delta(\vec{A} \odot_k \vec{B}) \Rightarrow \vec{D}} \odot_k L \quad \frac{\Gamma \Rightarrow \vec{A} \quad \Delta \Rightarrow \vec{B}}{\Gamma|_k \Delta \Rightarrow \vec{A} \odot_k \vec{B}} \odot_k R \\
\frac{\Gamma \Rightarrow \vec{A} \quad \Delta(\vec{C}) \Rightarrow \vec{D}}{\Delta(\Gamma|_k \vec{A} \downarrow_k \vec{C}) \Rightarrow \vec{D}} \downarrow L \quad \frac{\vec{A}|_1 \Gamma \Rightarrow \vec{C} \quad \dots \quad \vec{A}|_{S(A)} \Gamma \Rightarrow \vec{C}}{\Gamma \Rightarrow \vec{A} \downarrow \vec{C}} \downarrow R \\
\frac{\Gamma \Rightarrow \vec{B} \quad \Delta(\vec{C}) \Rightarrow \vec{D}}{\Delta(\vec{C} \uparrow_k \vec{B}|_k \Gamma) \Rightarrow \vec{D}} \uparrow L \quad \frac{\Gamma|_1 \vec{B} \Rightarrow \vec{C} \quad \dots \quad \Gamma|_{S(\Gamma)} \vec{B} \Rightarrow \vec{C}}{\Gamma \Rightarrow \vec{C} \uparrow \vec{B}} \uparrow R \\
\frac{\Delta(\vec{A}|_1 \vec{B}) \Rightarrow \vec{D} \quad \dots \quad \Delta(\vec{A}|_{S(A)} \vec{B}) \Rightarrow \vec{D}}{\Delta(\vec{A} \odot \vec{B}) \Rightarrow \vec{D}} \odot L \quad \frac{\Gamma_1 \Rightarrow \vec{A} \quad \Gamma_2 \Rightarrow \vec{B}}{\Gamma_1|_k \Gamma_2 \Rightarrow \vec{A} \odot \vec{B}} \odot R
\end{array}$$

Fig. 6 Hypersequent calculus for ω -DL, part II

The question of completeness of ω -DL, i.e. whether every valid hypersequent is a theorem, remains open. The question of the generative power of ω -DL also remains open. Valentín (2006)[?] observes that 1-DL can generate the non-context free but mildly context sensitive language $a^n b^n c^n$.

The *semantic type map* T for ω -DL is given in Figure 7. The unary connectives are interpreted as semantically inert. The semantic type map sends derivations into intuitionistic proofs so the usual Curry-Howard categorial type-logical semantics comes for free.

M. Moortgat has placed much emphasis on the possibility of interpreting type-logical connectives relationally (e.g. Moortgat 1997)[?], as we do here for the nondeterministic discontinuity operators. Such models can be rather austere, as van Benthem (2005)[?] puts it, i.e. being more general than functional models they are less contentful ontologically: a scientific theory should make the strongest claims possible which are not yet refuted. But in the present case we think the nondeterministic wrapping relational interpretation of discontinuity operators is motivated by its applicability to particle shift and complement alternation (see later), and perhaps to other phenomena of semi-free word order. Reape (1993)[?] appears to have been the first to propose what is (in our terms) a nondeterministic mode of discontinuous prosodic composition

$$\begin{aligned}
T(\triangleleft^{-1}B) &= T(B) \\
T(\triangleleft A) &= T(A) \\
T(\triangleright^{-1}B) &= T(B) \\
T(\triangleright A) &= T(A) \\
T(\overset{\cdot}{\leftarrow} B) &= T(B) \\
T(\overset{\cdot}{\leftarrow} A) &= T(A) \\
T(\overset{\cdot}{\rightarrow} B) &= T(B) \\
T(\overset{\cdot}{\rightarrow} A) &= T(A) \\
T(A \setminus C) &= T(A) \rightarrow T(C) \\
T(C / B) &= T(B) \rightarrow T(C) \\
T(A \bullet B) &= T(A) \& T(B) \\
T(A \downarrow_k C) &= T(A) \rightarrow T(C) \\
T(C \uparrow_k B) &= T(B) \rightarrow T(C) \\
T(A \odot_k B) &= T(A) \& T(B) \\
T(A \downarrow C) &= T(A) \rightarrow T(C) \\
T(C \uparrow B) &= T(B) \rightarrow T(C) \\
T(A \odot B) &= T(A) \& T(B)
\end{aligned}$$

Fig. 7 Semantic type map for ω -DL

(a kind of shuffle, for the German *Mittelfeld*), in the alternative categorial-like approach of Head-driven Phrase Structure Grammar.⁹

2.2 Labelled Natural Deduction for ω -DL

The Curry-Howard isomorphism between intuitionistic natural deduction and typed lambda calculus is as follows:

(41)

Intuitionistic Natural Deduction	Typed Lambda Calculus
formula	type
proof	lambda term
proof normalization	lambda reduction

The presentation of proofs in sequent calculus involves a lot of redundancy since all side formulas are copied from premises to conclusions at every step. When semantic annotation is also included there is additional redundancy with the lambda terms, and furthermore the evaluation of the Cutting of lexical semantics into derivational semantics has to be done all in one go in a phase subsequent to completing the entire derivation. Natural deduction is more economic because nodes are formulas/types rather than entire sequents, and furthermore, because of the way reasoning is carried out from lexical premises, semantic evaluation can be carried out synchronously with derivation reading from leaves to root. Therefore in this subsection we define

⁹ The extension of the present proposals to some such shuffle is problematic in that the sort of the output of shuffle might not be deterministically fixed by the sorts of its inputs; e.g. the shuffles of $s_1+1+s_2+1+s_3$ and t_1+1+t_2 could include $s_1+t_1+1+t_2+s_2+1+s_3$ and $s_1+1+s_2+t_1+1+t_2+s_3$ of sort 2, as well as $s_1+t_1+s_2+t_2+s_3$ of sort 0. Perhaps the ‘cards’ of shuffle should not be divided by separators, but just be the factors of continuous (bracketed) strings.

$$\begin{array}{c}
\vdots \\
\frac{A \rightarrow B}{B} E \rightarrow \\
\vdots \\
\frac{A^i}{B} I \rightarrow^i
\end{array}$$

$$\begin{array}{ccc}
\vdots & \vdots & \vdots \quad \vdots \\
\frac{A \wedge B}{A} E \wedge_1 & \frac{A \wedge B}{B} E \wedge_2 & \frac{A \quad B}{A \wedge B} I \wedge
\end{array}$$

$$\begin{array}{ccc}
\vdots & \vdots & \vdots \\
\frac{A \vee B \quad C \quad C}{C} E \vee^i & \frac{A}{A \vee B} I \vee_1 & \frac{B}{A \vee B} I \vee_2
\end{array}$$

Fig. 8 Natural deduction for positive propositional intuitionistic logic

(labelled, Prawitz-style) natural deduction for ω -**DL**, which will be used to illustrate some derivations with semantics in the next section.¹⁰

Natural deduction for positive propositional intuitionistic logic is as follows. A natural deduction proof is a tree of formulas with some coindexing of leaves with dominating nodes. The leaf formulas are called *hypotheses*: *open* if not indexed, *closed* if indexed. The root of the tree is the *conclusion*: a natural deduction proof asserts that the conjunction of its open hypotheses entails its conclusion. A trivial tree consisting of a single formula is a proof (from itself, as open hypothesis, to itself, as conclusion, corresponding to the identity axiom of sequent calculus). Then further proofs are those trees further generated by the rules shown in Figure 8. Note that hypotheses become indexed (closed) when the dominating inference occurs, and any number of hypotheses (including zero) can be indexed/closed in one step.

We can present type-logical calculi in a labelled deductive system (LDS) of natural deduction in which prosodic terms α and semantic terms ϕ label types A thus: $\alpha - \phi : A$; see Figures 9, 10 and 11. Like in the hypersequent calculus, prosodic terms are kept in “evaluated/spelt-out” form with only prosodic atoms of sort 0. The vectorial notation \vec{a} means $a_0 + 1 + \dots + 1 + a_i$ where i is the sort of a ; $\alpha|_k\beta, k > 0$ is the result of replacing the k -th separator in α by β .

3 Applications of Discontinuous Lambek Calculus

3.1 Linguistic Applications of **BDLC**

By Basic discontinuous Lambek calculus **BDLC** we mean ω -**DL** in which the discontinuous prosodic structure is restricted to just $+$: $L_0 \times L_0 \rightarrow L_0$ and W : $L_1 \times L_0 \rightarrow L_0$. Therefore the only discontinuous connectives it contains are sort non-polymorphic \downarrow , \odot and \uparrow . In this section we list accounts of linguistic phenomena falling within the scope of this minimal discontinuity calculus.

¹⁰ Eventually we would do even better with proof nets, where some (though not all) Cut-elimination between lexical and derivational semantics can be preevaluated once-and-for-all in a lexical compilation (Morrill 2005)[?].

$$\begin{array}{c}
\vdots \\
\frac{\alpha-\phi: \triangleleft^{-1}B}{\alpha+1-\phi: B} \triangleleft^{-1}E \quad \frac{\beta+1-\phi: B}{\beta-\phi: \triangleleft^{-1}B} \triangleleft^{-1}I \\
\vdots \\
\frac{\alpha+1-\phi: \triangleleft A}{\alpha-\phi: A} \triangleleft E \quad \frac{\alpha-\phi: A}{\alpha+1-\phi: \triangleleft A} \triangleleft I \\
\vdots \\
\frac{\alpha-\phi: \triangleright^{-1}B}{1+\alpha-\phi: B} \triangleright^{-1}E \quad \frac{1+\beta-\phi: B}{\beta-\phi: \triangleright^{-1}B} \triangleright^{-1}I \\
\vdots \\
\frac{1+\alpha-\phi: \triangleright A}{\alpha-\phi: A} \triangleright E \quad \frac{\alpha-\phi: A}{1+\alpha-\phi: \triangleright A} \triangleright I \\
\vdots \\
\frac{\alpha-\phi: \smile^k B}{\alpha|_k 0-\phi: B} E^{\smile^k} \quad \frac{\alpha|_k 0-\phi: B}{\alpha-\phi: \smile^k B} I^{\smile^k} \\
\vdots \\
\frac{\beta-\phi: \hat{}^k A \quad \frac{\overline{\alpha}-x: A^i \quad \gamma(\overline{\alpha}|_k 0)-\chi(x): C}{\gamma(\beta)-\chi(\phi): C} E^{\wedge^k i}}{\gamma(\beta)-\chi(\phi): C} \quad \frac{\alpha-\phi: A}{\alpha|_k 0-\phi: \hat{}^k A} I^{\wedge^k} \\
\vdots \\
\frac{\alpha-\phi: \smile B}{\alpha|_k 0-\phi: B} E^{\smile} \quad \frac{\alpha|_1 0-\phi: B \quad \dots \quad \alpha|_{S(B)} 0-\phi: B}{\alpha-\phi: \smile B} I^{\smile} \\
\vdots \\
\frac{\beta-\phi: \hat{} A \quad \frac{\overline{\alpha}-x: A^i \quad \gamma(\overline{\alpha}|_1 0)-\chi(x): C}{\gamma(\beta)-\chi(\phi): C} \quad \dots \quad \frac{\overline{\alpha}-x: A^i \quad \gamma(\overline{\alpha}|_{S(A)} 0)-\chi(x): C}{\gamma(\beta)-\chi(\phi): C} E^{\wedge^i}}{\gamma(\beta)-\chi(\phi): C} \quad \frac{\alpha-\phi: A}{\alpha|_k 0-\phi: \hat{} A} I^{\wedge}
\end{array}$$

Fig. 9 Labelled natural deduction for ω -DL, part I

3.1.1 Discontinuous Idioms

Idioms are complex expressions which have a meaning not compositionally attributable to the meanings of their parts (e.g. *red herring*). In grammar delivering logical semantics, they must be listed in the lexicon, because there is no other place from which their meaning can come. In discontinuous idioms, the idiomatic material is interpolated by non-idiomatic dependents, for example:

(42) Mary gave John/the man/... the cold shoulder.

Let there be the following lexical assignment:

$$\begin{array}{c}
\begin{array}{c} \vdots \\ \alpha-\phi: A \quad \gamma-\chi: A \setminus C \\ \hline \alpha+\gamma-(\chi \phi): C \end{array} E \setminus \\
\begin{array}{c} \vdots \\ \gamma-\chi: C/B \quad \beta-\psi: B \\ \hline \gamma+\beta-(\chi \beta): C \end{array} E / \\
\begin{array}{c} \overline{a}-x: A^i \quad \overline{b}-y: B^i \\ \vdots \\ \gamma-\chi: A \bullet B \quad \delta(\overline{a}+\overline{b})-\omega(x, y): D \\ \hline \delta(\gamma)-\omega(\pi_1 \chi, \pi_2 \chi): D \end{array} E \bullet^i \\
\begin{array}{c} \vdots \\ \alpha-\phi: A \quad \gamma-\chi: A \downarrow_k C \\ \hline \alpha|_k \gamma-(\chi \phi): C \end{array} E \downarrow_k \\
\begin{array}{c} \vdots \\ \gamma-\chi: C \uparrow_k B \quad \beta-\psi: B \\ \hline \gamma|_k \beta-(\chi \beta): C \end{array} E \uparrow_k \\
\begin{array}{c} \overline{a}-x: A^i \quad \overline{b}-y: B^i \\ \vdots \\ \gamma-\chi: A \odot_k B \quad \delta(\overline{a}|_k \overline{b})-\omega(x, y): D \\ \hline \delta(\gamma)-\omega(\pi_1 \chi, \pi_2 \chi): D \end{array} E \odot_k^i
\end{array}
\quad
\begin{array}{c}
\begin{array}{c} \overline{a}-x: A^i \\ \vdots \\ \overline{a}+\gamma-\chi: C \\ \hline \gamma-\lambda x \chi: A \setminus C \end{array} I \setminus^i \\
\begin{array}{c} \overline{b}-y: B^i \\ \vdots \\ \gamma+\overline{b}-\chi: C \\ \hline \gamma-\lambda y \chi: C/B \end{array} I /^i \\
\begin{array}{c} \vdots \\ \alpha-\phi: A \quad \beta-\psi: B \\ \hline \alpha+\beta-(\phi, \psi): A \bullet B \end{array} I \bullet \\
\begin{array}{c} \overline{a}-x: A^i \\ \vdots \\ \overline{a}|_k \gamma-\chi: C \\ \hline \gamma-\lambda x \chi: A \downarrow_k C \end{array} I \downarrow_k^i \\
\begin{array}{c} \overline{b}-y: B^i \\ \vdots \\ \gamma|_k \overline{b}-\chi: C \\ \hline \gamma-\lambda y \chi: C \uparrow_k B \end{array} I \uparrow_k^i \\
\begin{array}{c} \vdots \\ \alpha-\phi: A \quad \beta-\psi: B \\ \hline \alpha|_k \beta-(\phi, \psi): A \odot_k B \end{array} I \odot_k
\end{array}$$

Fig. 10 Labelled natural deduction for ω -DL, part II

$$(43) \quad \mathbf{gave+1+the+cold+shoulder} - shun := (N \setminus S) \uparrow N$$

Then our example is derived as follows in the hypersequent calculus and the labelled natural deduction calculus respectively:

$$(44) \quad \frac{\frac{N \Rightarrow N \quad S \Rightarrow S}{N, N \setminus S \Rightarrow S} \setminus L}{N, \sqrt[0]{(N \setminus S) \uparrow N}, N, \sqrt[1]{(N \setminus S) \uparrow N} \Rightarrow S} \uparrow L$$

$$(45) \quad \frac{\frac{\text{Mary} \quad \mathbf{gave+1+the+cold+shoulder-shun}: (N \setminus S) \uparrow N \quad \text{John-j}: N}{\mathbf{Mary-m}: N \quad \mathbf{gave+John+the+cold+shoulder-(shun j)}: N \setminus S} E \uparrow}{\mathbf{Mary+gave+John+the+cold+shoulder-(shun j m)}: S} E \setminus$$

$$\begin{array}{c}
\begin{array}{c} \vdots \\ \alpha-\phi: A \quad \gamma-\chi: A \downarrow C \\ \hline \alpha|_i \gamma-(\chi \phi): C \end{array} E \downarrow \quad \begin{array}{c} \overline{a}-x: A^i \\ \vdots \\ \overline{a}|_1 \gamma-\chi: C \quad \cdots \quad \overline{a}|_{S(A)} \gamma-\chi: C \\ \hline \gamma-\lambda x \chi: A \downarrow C \end{array} I \downarrow^i \\
\begin{array}{c} \vdots \\ \gamma-\chi: C \uparrow B \quad \beta-\psi: B \\ \hline \gamma|_i \beta-(\chi \beta): C \end{array} E \uparrow \quad \begin{array}{c} \overline{b}-y: B^i \\ \vdots \\ \gamma|_1 \overline{b}-\chi: C \quad \cdots \quad \gamma|_{S(C)} \overline{b}-\chi: C \\ \hline \gamma-\lambda x \chi: C \uparrow B \end{array} I \uparrow^i \\
\begin{array}{c} \overline{a}-x: A^i \quad \overline{b}-y: B^i \quad \overline{a}-x: A^i \quad \overline{b}-y: B^i \\ \vdots \quad \vdots \quad \vdots \quad \vdots \\ \gamma-\chi: A \odot B \quad \delta(\overline{a}|_1 \overline{b})-\omega(x, y): D \quad \cdots \quad \delta(\overline{a}|_{S(C)} \overline{b})-\omega(x, y): D \\ \hline \delta(\gamma)-\omega(\pi_1 \chi, \pi_2 \chi): D \end{array} E \odot^i \\
\begin{array}{c} \vdots \quad \vdots \\ \alpha-\phi: A \quad \beta-\psi: B \\ \hline \alpha|_i \beta-(\phi, \psi): A \odot B \end{array} I \odot
\end{array}$$

Fig. 11 Labelled natural deduction for ω -DL, part III

3.1.2 Quantification

Quantification is a classical instance of discontinuity, i.e. ‘syntactic’-semantic mismatch (better: prosodic-semantic mismatch): quantifier phrases occupy nominal positions prosodically but take sentential scope semantically, for example:

- (46) a. John gave every book to Mary.
b. $\forall x[(book\ x) \rightarrow (give\ m\ x\ j)]$

We treat quantification by type assignments such as the following:

- (47) **every** – $\lambda x \lambda y \forall z[(x\ z) \rightarrow (y\ z)]$
:= $((S \uparrow N) \downarrow S) / CN$

Such a composite of extraction and infixation to treat quantification was suggested in Moortgat (1991)[?], but he did not have a calculus ensuring that the extraction and infixation points would be one and the same. The first proposals to remedy this were those of Versmissen (1991)[?] and Solias (1992)[?].

An example like (46a) is derived (with the right semantics) as follows, where PTV abbreviates $(N \setminus S) / (N \bullet PP)$ or $((N \setminus S) / PP) / N$.

- (48)
$$\frac{\frac{\frac{N, PTV, N, PP \Rightarrow S}{N, PTV, [], PP \Rightarrow \forall S \uparrow N, [], \downarrow S \uparrow N} \uparrow R \quad S \Rightarrow S}{CN \Rightarrow CN \quad N, PTV, (S \uparrow N) \downarrow S, PP \Rightarrow S} \downarrow L}{N, PTV, ((S \uparrow N) \downarrow S) / CN, CN, PP \Rightarrow S} /L$$

Montague (1973)[?] presumably takes its title from its treatment of quantifiers and it is interesting to compare our treatment with his rule of term-insertion S14. Ignoring for the moment pronoun-binding aspects, S14 replaces by a noun phrase a (for us:

prosodic) variable in a nominal position in a sentence and semantically applies the noun phrase to the lambda abstraction of the sentence meaning over that of the nominal position. Our analysis splits such a step into two parts: conditionalization of the sentence over the nominal, semantically interpreted by functional abstraction over the nominal meaning, and infixing of the quantifier phrase into the conditionalized sentence, semantically interpreted by functional application of the infix to the circumfix.

Like that of Montague, our account allows quantifier phrases to take scope at the level of any embedding sentence, a feature which must eventually be constrained. However this successfully characterises the *de re*/specific and *de dicto*/nonspecific ambiguity of (49).

(49) Mary thinks someone left.

The *de dicto* reading, where the propositional attitude verb has wider scope than the existential quantifier (Mary does not necessarily have a particular person in mind), is generated by:

$$(50) \quad \frac{\frac{N, N \setminus S \Rightarrow S}{\boxed{} \uparrow R} \quad S \Rightarrow S}{(S \uparrow N) \downarrow S, N \setminus S \Rightarrow S} \downarrow L \quad \frac{N, N \setminus S \Rightarrow S}{N, (N \setminus S) / S, (S \uparrow N) \downarrow S, N \setminus S \Rightarrow S} /L$$

The *de re* reading, where the existential quantifier has wider scope than the propositional attitude verb (Mary has a particular person in mind), is generated by:

$$(51) \quad \frac{\frac{N, (N \setminus S) / S, N, N \setminus S \Rightarrow S}{N, (N \setminus S) / S, \boxed{} \uparrow R} \quad S \Rightarrow S}{N, (N \setminus S) / S, (S \uparrow N) \downarrow S, N \setminus S \Rightarrow S} \downarrow L$$

Also like the account of Montague, ours allows multiple quantifiers to scope in any order, another feature which must eventually be constrained (for example, *each* appears to always take wider scope). But this successfully characterizes the classical example of ambiguity:

(52) Everyone loves someone.

On the (dominant) subject wide scope reading, different people love, in general, different people (as in when we all love our respective mothers). On the (subordinate) object wide scope reading, different people love the same person (as in when we all love one and the same the queen). The subject wide scope ($\forall \exists$) reading is generated by:

$$(53) \quad \frac{\frac{N, (N \setminus S) / N, N \Rightarrow S}{N, (N \setminus S) / N, \boxed{} \uparrow R} \quad S \Rightarrow S}{N, (N \setminus S) / N, (S \uparrow N) \downarrow S \Rightarrow S} \downarrow L}{\frac{\boxed{} \uparrow R \quad S \Rightarrow S}{(S \uparrow N) \downarrow S, (N \setminus S) / N, (S \uparrow N) \downarrow S \Rightarrow S} \downarrow L}$$

The object wide scope ($\exists \forall$) reading is generated by:

$$(54) \quad \frac{\frac{N, (N \setminus S) / N, N \Rightarrow S}{\boxed{} \uparrow R} \quad S \Rightarrow S}{(S \uparrow N) \downarrow S, (N \setminus S) / N, N \Rightarrow S} \downarrow L}{\frac{\boxed{} \uparrow R \quad S \Rightarrow S}{(S \uparrow N) \downarrow S, (N \setminus S) / N, (S \uparrow N) \downarrow S \Rightarrow S} \downarrow L}$$

(The sooner processed, i.e. the nearer the root of the sequent proof, the wider the scope of the quantifier). Note that even assuming nondeterministic wrapping, in our account multiple quantifiers cannot get tangled up and bind each others' positions because the types driving the derivation ensure that the quantifier separator positions are only opened up and closed off one at a time, so that the only positions ever available are the unique correct ones.

For an account of the preference for $\forall\exists$ scope over $\exists\forall$ scope (i.e. left-to-right quantifier scope preference), see Morrill (2000)[?], which defines a complexity metric on analyses expressed as proof nets (Girard 1987)[?] motivated by the incrementality of processing. A range of other performance phenomena are also accounted for there in the same way.

3.1.3 VP Ellipsis

VP ellipsis refers to a class of constructions in which a form of *do* (perhaps suffixed by *too*) takes its interpretation from a preceding verb phrase, for example:

- (55) a. John slept before Mary did.
b. John slept and Mary did too.

Let there be the following lexical type assignment to the auxiliary, where VP abbreviates $N \setminus S$:

$$(56) \text{ did} - \lambda x \lambda y (x \ y \ y) \\ := ((VP \uparrow VP) / VP) \setminus (VP \uparrow VP)$$

Then an example such as (55a) is derived as follows:

$$(57) \frac{\frac{\frac{VP, (VP \setminus VP) / S, N, VP \Rightarrow VP}{[], (VP \setminus VP) / S, N, VP \Rightarrow \Downarrow VP \uparrow VP, [], \Downarrow VP \uparrow VP} \uparrow R}{[], (VP \setminus VP) / S, N \Rightarrow \Downarrow (VP \uparrow VP) / VP, [], \Downarrow (VP \uparrow VP) / VP} / R \quad \frac{VP \Rightarrow VP \quad N, VP \Rightarrow S}{N, \Downarrow VP \uparrow VP, VP, \Downarrow VP \uparrow VP \Rightarrow S} \uparrow L}{N, VP, (VP \setminus VP) / S, N, ((VP \uparrow VP) / VP) \setminus (VP \uparrow VP) \Rightarrow S} \setminus L$$

VP ellipsis can also occur intersententially, so an account must eventually be set up at the level of discourse.

3.2 Linguistic Applications of 1-DL

By 1-DL we mean ω -DL with only ever a single separator, in which case the deterministic and nondeterministic connectives collapse into the same operators, which we notate $\tilde{\cdot}$, $\hat{\cdot}$, \downarrow , \odot and \uparrow .

3.2.1 Medial Extraction

Extraction in which the gap is not at the periphery such as

- (58) dog that Mary saw today

can be modelled as follows:

$$(59) \text{ that} - \lambda x \lambda y \lambda z [(x \ z) \wedge (y \ z)] \\ := (CN \setminus CN) / \hat{\cdot} (S \uparrow N)$$

The example (58) is derived thus in the hypersequent calculus:

$$(60) \frac{\frac{N, (N \setminus S) / N, N, (N \setminus S) \setminus (N \setminus S) \Rightarrow S}{N, (N \setminus S) / N, [], (N \setminus S) \setminus (N \setminus S) \Rightarrow \downarrow S \uparrow N, [], \downarrow S \uparrow N} \uparrow R}{\frac{N, (N \setminus S) / N, (N \setminus S) \setminus (N \setminus S) \Rightarrow \wedge (S \uparrow N)}{CN, CN \setminus CN \Rightarrow CN} \wedge R} /L$$

The derivation in labelled natural deduction is as shown in Figure 12.¹¹

3.2.2 Pied-Piping

Pied-piping is the embedding of a filler such as a relative pronoun within accompanying material from the extraction site:

(61) scene the painting of which by Cezanne John sold for \$10,000,000

The depth of embedding is unbounded:

(62) thesis the height of the lettering on the first line of the second page of the third chapter of . . . of which is 0.5cm

Pied-piping can be treated by assignment as follows (cf. Morrill 1994, ch. 4; 1995)[?][?]:

(63) **which** – $\lambda x \lambda y \lambda z \lambda w [(z \ w) \wedge (y \ (x \ w))]$
 $:= (N \uparrow N) \downarrow ((CN \setminus CN) / \wedge (S \uparrow N))$

Then (61) is derived as shown in Figure 13, where PTV abbreviates say $(N \setminus S) / (N \bullet PP)$. Note that (63) can also generate relativisation in which there is no pied-piping by deriving an empty pied-piping context as $N \uparrow N$ ($[] \Rightarrow N \uparrow N$ is a theorem): once the assignment (63) is included, that of Section 3.2.1 is no longer required: the assignment (59) is derivable from, and so subsumed by, (63).

3.2.3 Appositive Relativisation

Appositive (‘nonrestrictive’) relativisation is relativisation in which the relative clause forms a lowered intonational phrase marked off by commas in writing, and modifies a noun phrase:

(64) John, who jogs, sneezed.

Semantically, the predication of the body of the appositive relative clause to the noun phrase modified is conjoined with the semantics of the embedding sentence in which the noun phrase is (also) understood. This discontinuity can be treated by the following assignment:

¹¹ This treatment captures the long-distanceness of left extraction, but something like unary bracket operators (Morrill 1992; Moortgat 1995; Fadda and Morrill 2005)[?][?][?] would be needed to express island constraints on left extraction. Also, something additional would be required to generate parasitic gaps, such as in *papers which I filed without reading*, in which the number of parasitic gaps which may accompany a host gap is unbounded (though not unconstrained). We do not see an extension of discontinuity to parasiticity because the unboundedness of the number of parasitic gaps and of their depth of embedding within islands would seem to go against the idea of finitude of a single rule of inference. We propose to treat parasiticity (and mediality) by a structural modality facilitating permutation and idempotency, see Morrill (2002)[?]. However we continue to use wrapping for left extraction in this paper, to illustrate with the machinery at hand.

$$\begin{array}{c}
\frac{N, \text{PTV}, N, \text{PP} \Rightarrow \text{S}}{\frac{N, \text{PTV}, \perp, \text{PP} \Rightarrow \downarrow \text{STN}, \perp, \downarrow \text{STN}}{\text{CN}, \text{CN} \setminus \text{CN} \Rightarrow \text{CN}} \uparrow R} \uparrow R \\
\frac{N/\text{CN}, \text{CN}/\text{PP}, \text{PP}/N, \perp, \text{CN} \setminus \text{CN} \Rightarrow \downarrow \text{NTN}, \perp, \downarrow \text{NTN}}{\text{CN}, N/\text{CN}, \text{CN}/\text{PP}, \text{PP}/N, (N \uparrow N) \downarrow ((\text{CN} \setminus \text{CN}) / \text{STN}), N, \text{PTV}, \text{PP} \Rightarrow \text{CN}} \uparrow L \\
\frac{N/\text{CN}, \text{CN}, N, \text{CN} \setminus \text{CN} \Rightarrow N}{\text{CN}, N/\text{CN}, \text{CN}/\text{PP}, \text{PP}/N, \perp, \text{CN} \setminus \text{CN} \Rightarrow \downarrow \text{NTN}, \perp, \downarrow \text{NTN}} \uparrow R \\
\frac{N, \text{PTV}, \text{PP} \Rightarrow \text{STN}}{\text{CN}, (\text{CN} \setminus \text{CN}) / \text{STN}, N, \text{PTV}, \text{PP} \Rightarrow \text{CN}} \uparrow R \\
\frac{\text{CN}, N/\text{CN}, \text{CN}/\text{PP}, \text{PP}/N, (N \uparrow N) \downarrow ((\text{CN} \setminus \text{CN}) / \text{STN}), N, \text{PTV}, \text{PP} \Rightarrow \text{CN}}{\text{CN}, N/\text{CN}, \text{CN}/\text{PP}, \text{PP}/N, \perp, \text{CN} \setminus \text{CN} \Rightarrow \downarrow \text{NTN}, \perp, \downarrow \text{NTN}} \uparrow L \\
\frac{\text{CN}, N/\text{CN}, \text{CN}/\text{PP}, \text{PP}/N, \perp, \text{CN} \setminus \text{CN} \Rightarrow \downarrow \text{NTN}, \perp, \downarrow \text{NTN}}{\text{CN}, N/\text{CN}, \text{CN}/\text{PP}, \text{PP}/N, (N \uparrow N) \downarrow ((\text{CN} \setminus \text{CN}) / \text{STN}), N, \text{PTV}, \text{PP} \Rightarrow \text{CN}} \uparrow L
\end{array}$$

Fig. 13 Hypersequent derivation of pied-piping (61)

$$\begin{aligned}
(65) \quad \mathbf{which} &- \lambda x \lambda y \lambda z [(x \ y) \wedge (z \ y)] \\
&:= (N \setminus ((S \uparrow N) \downarrow S)) / \text{STN}
\end{aligned}$$

Our example (64) is derived as follows:

$$(66)$$

$$\frac{\frac{\frac{N, N \setminus S \Rightarrow S}{\boxed{}}, N \setminus S \Rightarrow \sqrt[0]{S \uparrow N}, \boxed{}}, \sqrt[0]{S \uparrow N}}{N \setminus S \Rightarrow \wedge(S \uparrow N)} \uparrow R \quad \frac{\frac{N, N \setminus S \Rightarrow S}{\boxed{}}, N \setminus S \Rightarrow \sqrt[0]{S \uparrow N}, \boxed{}}, \sqrt[0]{S \uparrow N} \uparrow R \quad S \Rightarrow S}{\boxed{} \downarrow L} \quad \frac{N \Rightarrow N \quad (S \uparrow N) \downarrow S, N \setminus S \Rightarrow S}{N, N \setminus ((S \uparrow N) \downarrow S), N \setminus S \Rightarrow S} \setminus L}{\frac{N, N \setminus ((S \uparrow N) \downarrow S), N \setminus S \Rightarrow S}{N, (N \setminus ((S \uparrow N) \downarrow S)) / \wedge(S \uparrow N), N \setminus S, N \setminus S \Rightarrow S} /L} \wedge R$$

In a full type-logical treatment, bracket operators would be used to project the lowered intonational phrase of an appositive relative clause, and the same means as for restrictive relativisation would be used to allow pied-piping, so that the lexical assignment for an appositive relative pronoun would be:

$$(67) \text{ **which** } - \lambda w \lambda x \lambda y \lambda z [(x (w y)) \wedge (z y)] \\
:= (N \uparrow N) \downarrow ([\]^{-1} (N \setminus ((S \uparrow N) \downarrow S)) / \wedge(S \uparrow N))$$

3.2.4 Parentheticals

Parentheticals are adsentential modifiers such as *fortunately* which, to a very rough first approximation, can appear anywhere in the sentence they modify:¹²

- (68) a. Fortunately, John has perseverance.
b. John, fortunately, has perseverance.
c. John has, fortunately, perseverance.
d. John has perseverance, fortunately.

Such a distribution is captured by the following type assignment, as in Morrill and Merenciano (1996)[?].

$$(69) \text{ **fortunately** } - \textit{fortunately} \\
:= \sim S \downarrow S$$

For example, (68c) is derived as follows in the hypersequent calculus:

$$(70) \quad \frac{\frac{N, (N \setminus S) / N, N \Rightarrow S}{N, (N \setminus S) / N, \boxed{}, N \Rightarrow \sqrt[0]{\sim S}, \boxed{}}, \sqrt[0]{\sim S} \sim R \quad S \Rightarrow S}{N, (N \setminus S) / N, \sim S \downarrow S, N \Rightarrow S} \downarrow L$$

In labelled natural deduction, example (68c) is derived as shown in Figure 14.

3.2.5 Gapping

Gapping is a coordinate construction in which, in English in the simplest case, a verb missing medially in the second conjunct shares its interpretation with one present in the first conjunct:

- (71) John studies logic, and Charles, phonetics.

Coordinator types projecting such gapping were proposed in Solias (1992)[?] and Morrill and Solias (1993)[?]. P. Hendriks (1995)[?] proposed a like-type coordination assignment for gapping which we adapt as follows, where TV abbreviates $(N \setminus S) / N$.

¹² Of course, parentheticals cannot really occur *anywhere*, e.g. **The, fortunately, man left*. In the end there will have to be some kinds of domains which they cannot penetrate.

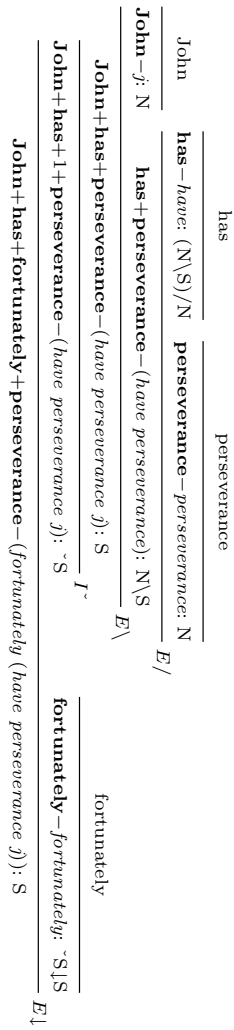


Fig. 14 Labelled natural deduction derivation of parenthesization (68c)

$$\begin{aligned}
(72) \quad \mathbf{and} &- \lambda x \lambda y \lambda z [(y z) \wedge (x z)] \\
&:= ((S \uparrow TV) \setminus (S \uparrow TV)) / \wedge (S \uparrow TV)
\end{aligned}$$

That the coordination is (almost) like-type is attractive, since it narrows the distance between gapping and constituent coordination (cf. Steedman 1990)[?]. The example (71) is derived as shown in Figure 15.

3.2.6 Comparative Subdeletion

Comparative subdeletion refers to comparisons in which the *than*-clause is missing a determiner:

(73) John ate more donuts than Mary bought bagels.

$$\begin{array}{c}
\frac{\frac{\frac{N, TV, Q, CN \Rightarrow S}{\downarrow R} \quad \frac{CP/S, N, TV, Q, CN \Rightarrow CP}{\uparrow R}}{CP/S, N, TV, \perp, CN \Rightarrow \downarrow CP\uparrow Q, \perp, \downarrow CP\uparrow Q} \quad \frac{CP/S, N, TV, CN \Rightarrow \uparrow CP\uparrow Q}{\downarrow R}}{N, TV, \perp, CN \Rightarrow \downarrow ST\uparrow Q, \perp, \downarrow ST\uparrow Q} \quad \frac{S/\uparrow CP\uparrow Q, CP/S, N, TV, CN \Rightarrow S}{\downarrow L}}{N, TV, (ST\uparrow Q)\downarrow(S/\uparrow CP\uparrow Q), CN, CP/S, N, TV, CN \Rightarrow S}
\end{array}$$

Fig. 16 Hypersequent derivation of comparative subdeletion (73)

even English. Although some have defended context-freeness (Gazdar et al.)[?], Huybregts (1976, 1985)[?][?] argued that Dutch is not context-free and Shieber (1985)[?] formally proved that Swiss German is not context-free.¹³ The relevant feature of both languages is semantic cross-serial dependency in subordinate clauses, and the formal (i.e. string set) proof is enabled by the morphological case-marking of dependents by verbs in Swiss German (but not Dutch). Cross-serial dependency in Swiss German is illustrated by the following examples:

- (75) a. ... das mer em Hans es huus hälfed aastriche
that we Hans-dat the house-acc helped paint
“that we helped Hans paint the house”
b. ... das mer d’chind em Hans es huus lönd hälfe aastriche
that we the children-acc Hans-dat the house-acc let help paint
“that we let the children help Hans paint the house”

¹³ See also Culy (1985)[?] for the non-context-freeness of Bambara morphology.

Calcagno (1995)[?] provides an analysis of cross-serial dependencies which is a close precedent to ours, but in terms of categorial head-wrapping of headed strings. In that account, all expressions are of the same datatype (headed string) and there is no prosodic sorting. Here we present a similar account, but using the same sorted discontinuity calculus also motivated by our other linguistic applications.

Dutch subordinate clauses are verb final:

- (76) (... dat) Jan boeken las
 (... that) J. books read
 CP/S N N $N \setminus (N \setminus S)$ \Rightarrow CP
 ‘(... that) Jan read books’

Modals and control verbs, so-called verb raising triggers, appear in a clause-final verb cluster but in the English word order relative to one another:

- (77) (... dat) Jan boeken kan lezen
 (... that) J. books is able read
 CP/S N N $(N \setminus Si) \downarrow (N \setminus S)$ $\triangleright^{-1} (N \setminus (N \setminus Si))$ \Rightarrow CP
 ‘(... that) Jan is able to read books’

- (78) (... dat) Jan boeken wil kunnen
 (... that) J. books wants be able
 CP/S N N $(N \setminus Si) \downarrow (N \setminus S)$ $\triangleright^{-1} ((N \setminus Si) \downarrow (N \setminus Si))$
 lezen
 read
 $\triangleright^{-1} (N \setminus (N \setminus Si))$ \Rightarrow CP
 ‘(... that) Jan wants to be able to read books’

The basic idea of our analysis (Morrill 2000; Morrill et al. 2007)[?][?] is to mark the left edge of the subordinate clause verb cluster with a separator, and to have successive verb-raising triggers infixing at this point and inserting another separator to their own left (if they are infinitive) or closing off the point of discontinuity (if they are finite). The labelled natural deduction derivation of the subordinate clause verb phrase in (78) is given in Figure 17.

However, caution needs to be taken in relation to the interaction of verb-raising with our account of quantification:

- (79) a. ... dat Jan alles las
 that Jan everything read
 “that Jan read everything”
 b. ... dat Jan alles kan lezen
 that Jan everything is-able read
 “that Jan can read everything”

Such a quantifier phrase presumably at least sometimes needs to be allowed to take scope which is intermediate with respect to the clause-final verbs, i.e. within clauses that still contain the verb-raising separator, and the process of quantification introduces its own separator in addition. The two positions of discontinuity must not be confused.

We thus propose to allow a new additional separator for the verb cluster left edge. Thus Si is of sort 1_v where 1_v is the new verb cluster left edge separator, etc. All the machinery is just duplicated for the additional separator. The ease of this iteration (which can apparently be repeated at will) would appear to be a virtue of the separator approach. But for the purposes of illustration here we use the same notation as always.

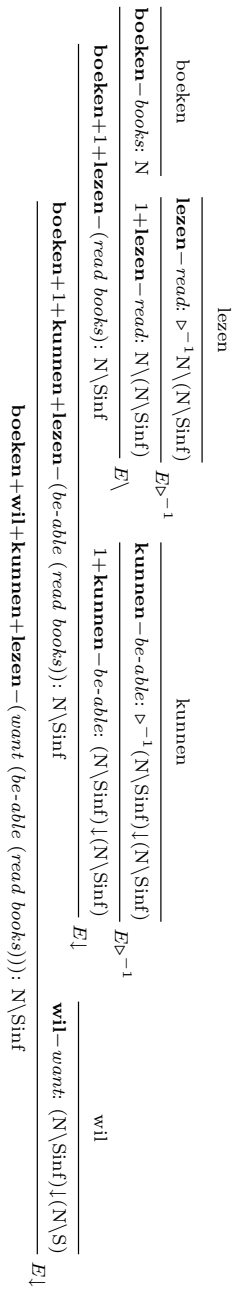


Fig. 17 Labelled natural deduction derivation of Dutch verb-raising (78)

To generate alternative quantifier scopings of examples like (79b) we require a quantifier assignment for quantification in Si of sort 1, in addition to the standard one for quantification in $S(\text{fin})$ of sort 0 for (79a); we would want to collapse these into a sort-polymorphic type:

$$(80) \quad \text{alles} \quad - \quad \lambda x \forall y [(thing\ y) \rightarrow (x\ y)] \\ := \quad (S\alpha \uparrow N) \downarrow S\alpha$$

When the infinitival complement verbs also take objects, cross-serial dependencies are generated.¹⁴

$$(81) \quad \begin{array}{l} (\dots \text{ dat}) \quad \text{Jan} \quad \text{Cecilia}_1 \quad \text{Henk}_2 \quad \text{de} \quad \text{nijlpaarden}_3 \\ (\dots \text{ that}) \quad \text{J.} \quad \text{C.} \quad \text{H.} \quad \text{the} \quad \text{hippos} \\ \text{CP/S} \quad \text{N} \quad \text{N} \quad \text{N} \quad \text{N/CN} \quad \text{CN} \\ \text{zag}_1 \quad \quad \quad \text{helpen}_2 \quad \quad \quad \text{voeren}_3 \\ \text{saw} \quad \quad \quad \text{help} \quad \quad \quad \text{feed} \\ (N \setminus Si) \downarrow (N \setminus (N \setminus S)) \quad \triangleright^{-1}((N \setminus Si) \downarrow (N \setminus (N \setminus Si))) \quad \triangleright^{-1}(N \setminus (N \setminus Si)) \quad \Rightarrow \quad \text{CP} \\ \text{'}(\dots \text{ that}) \text{ Jan saw}_1 \text{ Cecilia}_1 \text{ help}_2 \text{ Henk}_2 \text{ feed}_3 \text{ the hippos}_3 \text{'} \end{array}$$

These are generated by assignments with the verbs seeking objects cross-serially to the far left after infixing at the separator-marked left edge of the verb cluster.

Main clause yes/no interrogative word order, V1, is derived from subordinate clause word order by fronting the finite verb. We therefore propose a lexical rule mapping (subordinate clause) finite verb types V to $Q/\wedge(S \uparrow V)$, cf. Hepple (1990)[?].

$$(82) \quad \begin{array}{l} \text{Wil} \quad \quad \quad \text{Jan} \quad \text{boeken} \quad \text{lezen?} \\ \text{wants} \quad \quad \quad \text{J.} \quad \text{books} \quad \text{read} \\ Q/\wedge(S \uparrow ((N \setminus Si) \downarrow (N \setminus S))) \quad \text{N} \quad \text{N} \quad \triangleright^{-1}(N \setminus (N \setminus Si)) \quad \Rightarrow \quad Q \\ \text{'Does Jan want to read books?'} \end{array}$$

Main clause declarative word order, V2, is further derived from V1 by fronting a major constituent. We propose to achieve this by allowing complex distinguished types (cf. Morrill and Gavarró 1992)[?].

$$(83) \quad \begin{array}{l} \text{Jan} \quad \text{wil} \quad \quad \quad \text{boeken} \quad \text{lezen.} \\ \text{J.} \quad \text{wants} \quad \quad \quad \text{books} \quad \text{read} \\ \text{N} \quad Q/\wedge(S \uparrow ((N \setminus Si) \downarrow (N \setminus S))) \quad \text{N} \quad \triangleright^{-1}(N \setminus (N \setminus Si)) \quad \Rightarrow \quad \text{N} \bullet \wedge (Q \uparrow N) \\ \text{'Jan wants to read books.'} \end{array}$$

A hypersequent calculus derivation of the main clause *Jan wil boeken lezen* is given in Figure 18, where VP abbreviates $N \setminus S$ and VPi abbreviates $N \setminus Si$. A hypersequent calculus derivation of the main clause *Marie zegt dat Jan Cecilia Henk de nijlpaarden zag helpen voeren* ('Marie says that Jan saw Cecilia help Henk feed the hippos'), with subordinate clause cross-serial dependencies, is given in Figure 19.

3.3 Linguistic Applications of Deterministic 2-DL

Here we consider deterministic discontinuity allowing two separators.

¹⁴ 'An increasing load in processing makes such multiple embeddings increasingly unacceptable', Steedman (1985)[?], fn. 29, p.546]

- (89) a. Mary talked to John_i about himself_i.
 b. *Mary talked about himself_i to John_i

Such a feature can be captured using second-position deterministic wrapping (VP abbreviates N\S):

- (90) **himself** – $\lambda x \lambda y (x \ y \ y)$
 $:= ((VP \uparrow N) \uparrow_2 N) \downarrow_2 (VP \uparrow N)$

(The unsubscripted operators can be considered the first-position deterministic wrapping varieties, or equally the nondeterministic operators, since the sorting of the types ensures that the separator is always unique.) Then (84b) is derived as shown in Figure 20.

We assume that clause-locality can again be ensured by intensionalization as before, but (90) additionally overgenerates in allowing an antecedent which does not c-command the reflexive such as in:

- (91) *Mary talked to the friends of John_i about himself_i.

On the other hand, a requirement that the antecedent always c-command a reflexive does not seem right either, since in (89a) the antecedent does not do so. Chomskyan syntax salvages c-command by claiming that in such an example, the preposition *to* is assimilated to the verb *talked* (“reanalysis”), but the issue remains a mystery to us. Likewise a mystery is the fact that a reflexive can sometimes precede a (non c-commanding) antecedent, or even apparently take its antecedent in another sentence:

- (92) That photofit poster of himself_i hanging in every postoffice was really beginning to worry Clyde_i.
 (93) Clyde_i was really beginning to get worried. That photofit poster of himself_i in every postoffice was making it harder and harder to get around unnoticed.

Pollard and Sag (1994)[?] claim that such “logophora” is possible just when the reflexive falls within the least “oblique” complement of a verb. Note that an antecedent can occur in an adverbial phrase, and even be split:

- (94) When Bonnie met up with Clyde, they really began to get carried away with themselves.

Finally, note that a reflexive can also take a deictic antecedent:

- (95) Know yourself!

Our sketch of intrasentential reflexivisation can only be a beginning.

3.3.2 Anaphora

Of course the possibility of intersentential dependencies is even more extensive with personal pronouns than with reflexive pronouns. In the end a theory of sentential syntax must be integrated with a theory of discourse in this and other relations. But for the time-being we approximate intrasentential anaphora.

Anaphora divides into backward anaphora or anaphora proper (antecedent precedes pronoun) and forward anaphora or cataphora (pronoun precedes ‘antecedent’). At first blush it might seem that a single nondeterministic wrapping pronoun type $((S \uparrow N) \uparrow N) \downarrow (S \uparrow N)$ would conveniently allow the alternation, but reflection reveals that the very nondeterminism would make it impossible not to violate case restrictions when these are different between antecedent and pronoun positions:

$$(97) \text{ him} - \lambda x \lambda y (x y y) \\ := ((S \uparrow \text{Nsg}(3(m)) C) \uparrow_2 \text{Nsg}(3(m)) \text{acc}) \downarrow_2 (S \uparrow \text{Nsg}(3(m)) C)$$

This interacts with our treatment of quantification to produce essentially the same characterisation of quantification and bound anaphora that Montague’s (1973)[?] rule S14 of term insertion gives in PTQ. But that rule is infinitary: a term-inserted antecedent noun phrase looks to bind unboundedly many following pronouns in one go; in that respect the rule is not computational in the sense of being a finitary step. Our treatment is computationally finitary in that it is the responsibility of each pronoun to find in turn its single preceding antecedent, but unboundedly many pronouns can find the same antecedent, so the effect is basically the same as in PTQ.

As a result, our account suffers the same limitations as Montague’s: it undergenerates in not producing any forward anaphora (cataphora), for example *Near him_i, [every cowboy]_i kept a gun*, it overgenerates with respect to Principle B, for example **John_i likes him_i*,¹⁵ and of course the whole issue of intersentential anaphora is left hanging. Finally, the only reason why there are no Principle C violations is that the account generates no cataphora at all and English is right-branching. We could attempt to treat cataphora by, say:

$$(98) \text{ he} - \lambda x \lambda y (x y y) \\ := ((S \uparrow \text{Nsg}(3(m)) \text{nom}) \uparrow_2 \text{Nsg}(3(m)) C) \downarrow_1 (S \uparrow \text{Nsg}(3(m)) C)$$

But then we would even overgenerate simultaneous Principle B and C violations like **He_i likes John_i*.

Jacobson (1999)[?] and Jäger (2001)[?] both give accounts of anaphora invoking a new binary type constructor such that B^A (Jacobson’s notation) or $B|A$ (Jäger’s notation) is an expression of type B containing an unbound anaphor of type A ; the meaning is the functional abstraction of the expression meaning over the anaphor meaning, i.e. the semantic type is $T(A) \rightarrow T(B)$.

Although the introduction of new type-constructors is type-logical in spirit, Jacobson’s account is couched in terms of combinatory categorial grammar, i.e. the characterisation of derivability by a small number of axiomatic combinatory reduction schemata.¹⁶

The account of Jäger is type-logical in giving a sequent calculus, but we know of no straightforward prosodic interpretation of the type-constructor $|$. Like Jacobson’s account, it has the attractive feature that a sentence with n free pronouns is simply of type $(\dots (S|N^{(1)}) \dots)|N^{(n)}$ with semantics the functional abstraction of the sentence meaning over the meanings of the pronoun positions. This seems like a good point of departure to interface syntax with discourse, but as we say, prosodic interpretation is pending.

¹⁵ Grodzinsky and Reinhart (1993)[?] suggest that Principle B may be a pragmatic constraint rather than a syntactic one: that if the local interpretation was intended, the less ambiguous reflexive form would have been used. Morrill (2003)[?] attempts to substantiate such an idea further by the principles on the incremental complexity of proof nets of Morrill (2000)[?].

¹⁶ There is the following argument as to why combinatory categorial grammar is by definition incomplete. The product-free Lambek calculus is complete with respect to free semigroups, i.e. concatenation, (Buszkowski 1982)[?], but it is not finitely axiomatizable (Zielonka 1981)[?]. So combinatory categorial grammar qua finite (‘small’) number of combinatory schemata cannot be complete with respect to concatenation. The essence of Zielonka’s result is that the properties of recursively defined types cannot be fully captured without recursive rules, which is not really surprising.

3.4 Linguistic Applications of Nondeterministic 2-DL

Here we consider nondeterministic discontinuity allowing two separators.

3.4.1 Complement Alternation

By complement alternation we mean the free alternation in the order of the PPs in examples such as the following:¹⁷

- (99) a. John talked to Mary about Bill.
b. John talked about Bill to Mary.

The following single lexical assignment generates the alternation; this improves on Morrill (2002)[?] in using the left projection type-constructor rather than a complex prosodic form containing separators.

$$(100) \text{ talked} - \text{talk} \\ := \triangleleft^{-1} \triangleleft^{-1} ((N \setminus S) \uparrow PP_{to}) \uparrow PP_{about}$$

The derivations of (99a, b) are as follows, where VP abbreviates $N \setminus S$.

$$(101) \frac{\frac{\frac{PP_{about} \Rightarrow PP_{about}}{\frac{\frac{PP_{to} \Rightarrow PP_{to}}{\frac{VP \Rightarrow VP}{\frac{\emptyset \sqrt{VP \uparrow PP_{to}}, PP_{to}, \downarrow \sqrt{VP \uparrow PP_{to}} \Rightarrow VP} \uparrow L} \uparrow L} \frac{\emptyset \sqrt{(VP \uparrow PP_{to}) \uparrow PP_{about}}, PP_{about}, \downarrow \sqrt{(VP \uparrow PP_{to}) \uparrow PP_{about}}, PP_{to}, \uparrow \sqrt{(VP \uparrow PP_{to}) \uparrow PP_{about}} \Rightarrow VP} \uparrow L} \triangleleft^{-1} L} \frac{\emptyset \sqrt{\triangleleft^{-1}((VP \uparrow PP_{to}) \uparrow PP_{about}), PP_{about}}, \downarrow \sqrt{\triangleleft^{-1}((VP \uparrow PP_{to}) \uparrow PP_{about}), PP_{to}} \Rightarrow VP} \triangleleft^{-1} L} \triangleleft^{-1} \triangleleft^{-1} ((VP \uparrow PP_{to}) \uparrow PP_{about}), PP_{about}, PP_{to} \Rightarrow VP$$

$$(102) \frac{\frac{\frac{PP_{about} \Rightarrow PP_{about}}{\frac{\frac{PP_{to} \Rightarrow PP_{to}}{\frac{VP \Rightarrow VP}{\frac{\emptyset \sqrt{VP \uparrow PP_{to}}, PP_{to}, \downarrow \sqrt{VP \uparrow PP_{to}} \Rightarrow VP} \uparrow L} \uparrow L} \frac{\emptyset \sqrt{(VP \uparrow PP_{to}) \uparrow PP_{about}}, PP_{to}, \downarrow \sqrt{(VP \uparrow PP_{to}) \uparrow PP_{to}}, PP_{about}, \uparrow \sqrt{(VP \uparrow PP_{to}) \uparrow PP_{about}} \Rightarrow VP} \uparrow L} \triangleleft^{-1} L} \frac{\emptyset \sqrt{\triangleleft^{-1}((VP \uparrow PP_{to}) \uparrow PP_{about}), PP_{to}}, \downarrow \sqrt{\triangleleft^{-1}((VP \uparrow PP_{to}) \uparrow PP_{about}), PP_{about}} \Rightarrow VP} \triangleleft^{-1} L} \triangleleft^{-1} \triangleleft^{-1} ((VP \uparrow PP_{to}) \uparrow PP_{about}), PP_{to}, PP_{about} \Rightarrow VP$$

3.4.2 Particle Shift

Particle shift is the alternation in the order of a particle verb's object and its particle:¹⁸

- (103) a. John called up Mary.
b. John called Mary up.

The alternation is generated by the following single lexical assignment, which also refines Morrill (2002)[?] in using the left projection type-constructor instead of a peripheral lexical separator, though the medial lexical separator must remain:

¹⁷ We have no explanation for the discrepancy: *John talked to Mary about herself* / **John talked about Mary to herself*. Perhaps it has to do with an obliqueness ordering on the complement thematic roles and that a reflexive must have a less oblique antecedent (Pollard and Sag 1994)[?].

¹⁸ We have no grammatical explanation of why the pronoun must be focused in *John called up HER*. or why ?*John called the very heavy man up*. It seems that for the object to appear to the right it is necessary and sufficient for it to be 'informative'. This is the same issue as heavy noun phrase shift.

$$\begin{array}{c}
\text{called } (\dots) \text{ up} \\
\hline
\text{called}+1+\text{up}-\text{phone}: \triangleleft^{-1} \cdot (N \setminus S) \uparrow N \\
\hline
\text{called}+1+\text{up}+1-\text{phone}: \cdot (N \setminus S) \uparrow N \quad \triangleleft^{-1} E \quad \frac{\text{Mary}}{\text{Mary}-m: N} \\
\hline
\text{called}+1+\text{up}+\text{Mary}-(\text{phone } m): \cdot (N \setminus S) \quad E \uparrow \\
\hline
\text{called}+\text{up}+\text{Mary}-(\text{phone } m): N \setminus S \quad E \cdot
\end{array}$$

$$\begin{array}{c}
\text{called } (\dots) \text{ up} \\
\hline
\text{called}+1+\text{up}-\text{phone}: \triangleleft^{-1} \cdot (N \setminus S) \uparrow N \\
\hline
\text{called}+1+\text{up}+1-\text{phone}: \cdot (N \setminus S) \uparrow N \quad \triangleleft^{-1} E \quad \frac{\text{Mary}}{\text{Mary}-m: N} \\
\hline
\text{called}+\text{Mary}+\text{up}+1-(\text{phone } m): \cdot (N \setminus S) \quad E \uparrow \\
\hline
\text{called}+\text{Mary}+\text{up}-(\text{phone } m): N \setminus S \quad E \cdot
\end{array}$$

Fig. 21 Labelled natural deduction derivations of particle shift (103a, b).

$$(104) \quad \text{called}+1+\text{up} - \text{phone} \\
:= \triangleleft^{-1} \cdot (N \setminus S) \uparrow N$$

The derivations of (103a, b) are given in Figure 21.

4 Conclusion

Starting around 1990, the most promising line of enrichment of Lambek calculus into Type Logical Grammar appeared to be multimodality with structural rules of interaction and perhaps inclusion communicating between distinct modes of adjunction (Moortgat and Morrill 1991, Morrill 1994, Moortgat 1997)[?][?][?]. The idea of multimodality seems right enough, but its mediation by structural rules led in some cases to analyses in which whole families of connectives and attendant structural rules were introduced to deal with a single linguistic construction. That does not seem to be very good value for money in terms of the number of linguistic phenomena addressed per technical device introduced. We hope to have shown here that by contrast, the Discontinuous Lambek Calculus, while remaining multimodal, has a wide range of linguistic applications at the same time that it does away with cumbersome structural manipulations (the hypersequent calculus contains no structural rules).

Sequent calculus is an elegant and revealing deductive formalism because it fully modularizes the inferential properties of single connective occurrences. However, it is not canonical, i.e. even amongst Cut-free proofs there are still distinct equivalent derivations which only differ in irrelevant rule orderings. There is a pun here: *sequent* proofs, although of tree-form, are still to some extent *sequential* in the sense of serial. (Morrill 2003)[?] defines syntactic structure as ‘the essential geometrical structure in virtue of which an expression is deemed to be grammatical’. For the proclamation of such syntactic structure, and also to broach optimally efficient (incremental) processing, we need to identify maximally *parallelised* representations. These appear to be noncommutative versions of the proof nets of Girard (1987)[?], introduced for the Lambek calculus in Roorda (1991)[?]. On the assumption that our language faculty is optimal, such optimized representations would then have the most psychological credibility; see

Morrill (2000)[?] on psycholinguistic complexity and the incremental complexity of proof nets. M. Steedman (p.c. 1987) pointed out that we want for categorial grammar what phrase-markers are for context-free grammar. We think that proof nets are what was sought, that proof nets are the syntactic structures, and the term of reference, both technologically and psychologically, for the processing of categorial grammar.

To our mind then, the next step is the development of proof nets for discontinuous Lambek calculus. An initial proposal is developed in Morrill (1999)[?] and Morrill and Fadda (2008)[?], based on a generalisation to discontinuity of a first-order encoding of the binary relational models of van Benthem (1991)[?] (see also Moot and Piazza 2001)[?].

5 Appendix: Cut-Elimination

The main goal of this section is to prove the Cut-elimination theorem for $\omega\text{-DL}$. A syntactic proof of the Cut-elimination property in the hypersequent calculus is obtained via consideration of Cut-elimination in two variant calculi: a hypersequent variant $\omega\text{-DL}_\epsilon$ of $\omega\text{-DL}$ in which units replace the unary connectives, and a variant $\text{mm-}\omega\text{-DL}_\epsilon$ of this calculus which is a sorted multimodal sequent calculus.

A few words on notation: 1) in hypersequent calculi, Greek capital letters (Δ, Γ, \dots) denote not just configurations but also preconfigurations, where a preconfiguration is any substring of a configuration (configurations are particular cases of preconfigurations); 2) where \mathbf{C} is a sequent or hypersequent calculus, then \mathbf{C}^- denotes the corresponding calculus without the Cut rule, so, for example, $\omega\text{-DL}^-$ is the hypersequent calculus for $\omega\text{-DL}$ without the Cut rule; and 3) where \mathbf{C} is a calculus and n a natural number, then $\vdash_{\mathbf{C}}^n$ is the provability relation of at most n derivation steps.

5.1 The Discontinuous Lambek Calculus with Units: $\omega\text{-DL}_\epsilon$

This variant of $\omega\text{-DL}$ preserves the binary type-constructors, but replaces the unary type-constructors by two units (nullary type-constructors) in terms of which the unary connectives can be defined. The units are the continuous unit I of sort 0 such that $[[I]] = \{0\}$ and the discontinuous unit J of sort 1 such that $[[J]] = \{1\}$. The set of types and the sort map for $\omega\text{-DL}_\epsilon$ are obtained from Figures 3 and 4 by dropping the clauses for the unary connectives and adding:

$$\begin{aligned}\mathcal{F}_0 &::= I \quad S(I) = 0 \\ \mathcal{F}_1 &::= J \quad S(J) = 1\end{aligned}$$

The configurations of $\omega\text{-DL}_\epsilon$ are the same as those of $\omega\text{-DL}$ and the hypersequent calculus is obtained from Figures 5 and 6 by dropping the rules for the unary connectives and adding:

$$\begin{array}{c} \frac{\Delta_0, \Delta_1 \Rightarrow \vec{A}}{\Delta_0, I, \Delta_1 \Rightarrow \vec{A}} IL \quad \frac{}{\Rightarrow I} IR \\ \\ \frac{\Delta_0, \Delta_1, \Delta_2 \Rightarrow \vec{A}}{\Delta_0, \sqrt[0]{J}, \Delta_1, \sqrt[1]{J}, \Delta_2 \Rightarrow \vec{A}} JL \quad \frac{}{\Box \Rightarrow \sqrt[0]{J}, \Box, \sqrt[1]{J}} JR \end{array}$$

We define the following translation τ from the types of $\omega\text{-DL}$ to the types of $\omega\text{-DL}_\epsilon$:

$$\begin{aligned}
\tau(\hat{\ }A) &= \tau(A) \odot I \\
\tau(\check{\ }A) &= \tau(A) \uparrow I \\
\tau(\hat{\ }^i A) &= \tau(A) \odot_i I \\
\tau(\check{\ }^i A) &= \tau(A) \uparrow_i I \\
\tau(\triangleleft A) &= \tau(A) \bullet J \\
\tau(\triangleleft^{-1} A) &= \tau(A) / J \\
\tau(\triangleright A) &= J \bullet \tau(A) \\
\tau(\triangleright^{-1} A) &= J \setminus \tau(A) \\
\tau(A \star B) &= \tau(A) \star \tau(B) \text{ for } \star \in \{\setminus, \bullet, /, \downarrow_k, \odot_k, \uparrow_k, \downarrow, \odot, \uparrow\}
\end{aligned}$$

It is straightforward to extend τ to map $\omega\text{-DL}$ hypersequents to $\omega\text{-DL}_\epsilon$ hypersequents. The following lemma states that this is an embedding translation, and that if the image of a $\omega\text{-DL}$ hypersequent has a Cut-free proof in $\omega\text{-DL}_\epsilon$ then the $\omega\text{-DL}$ hypersequent has a Cut-free proof in $\omega\text{-DL}$.

Lemma 1 *Let $\Delta \Rightarrow \vec{A}$ be a $\omega\text{-DL}$ hypersequent. Then:*

- (1) If $\vdash_{\omega\text{-DL}} \Delta \Rightarrow \vec{A}$ then $\vdash_{\omega\text{-DL}_\epsilon} \tau(\Delta) \Rightarrow \tau(\vec{A})$
- (2) If $\vdash_{\omega\text{-DL}_\epsilon} \tau(\Delta) \Rightarrow \tau(\vec{A})$, then $\vdash_{\omega\text{-DL}} \Delta \Rightarrow \vec{A}$

The proof of lemma 1 is on page 47.

5.2 A Sorted Multimodal Sequent Calculus for $\omega\text{-DL}_\epsilon$: $\text{mm-}\omega\text{-DL}_\epsilon$

We consider now a sorted multimodal calculus which displays the implicit structural rules in $\omega\text{-DL}$. This is like the calculi of standard multimodal type logical grammar with structural terms in antecedents, but with the novelties that: 1) structural terms, like types, are sorted; 2) there are two structural term constants which stand respectively for the continuous unit and discontinuous unit; 3) there are infinitely many terms constructors (modes); and 4) there are some rules of unbounded arity.

There are two kinds of structural term constructors: \circ (which stands for term concatenation), and $\circ_i, i > 0$, (which stands for term wrapping at the i -th position). The sorted structural terms are defined by mutual recursion as follows:

$$\begin{aligned}
\mathbf{StructTerm}_0 &::= \mathbb{I} \\
\mathbf{StructTerm}_1 &::= \mathbb{J} \\
\mathbf{StructTerm}_i &::= \mathcal{F}_i \\
\mathbf{StructTerm}_{i+j} &::= \mathbf{StructTerm}_i \circ \mathbf{StructTerm}_j \\
\mathbf{StructTerm}_{i+j} &::= \mathbf{StructTerm}_{i+1} \circ_k \mathbf{StructTerm}_j
\end{aligned}$$

Their *sort map* is computed in the obvious way.

$X[Y]$ denotes a structural term with a distinguished position occupied by the structural term Y . If A, X are respectively a type and a structural term, then a and x denote their sorts.

The multimodal discontinuity calculus is as follows. The identity rules are:

$$\frac{}{A \rightarrow A} \textit{id} \qquad \frac{\Gamma \Rightarrow A \quad \Delta(A) \Rightarrow B}{\Delta(\Gamma) \Rightarrow B} \textit{Cut}$$

The deterministic logical rules are:

$$\begin{array}{c}
\frac{T[\mathbb{I}] \rightarrow A}{T[I] \rightarrow A} IL \quad \frac{}{\mathbb{I} \Rightarrow I} IR \\
\\
\frac{T[\mathbb{J}] \rightarrow A}{T[J] \rightarrow A} JL \quad \frac{}{\mathbb{J} \Rightarrow J} JR \\
\\
\frac{X \rightarrow A \quad Y[B] \rightarrow C}{Y[X \circ A \setminus B] \rightarrow C} \setminus L \quad \frac{A \circ X \rightarrow B}{X \rightarrow A \setminus B} \setminus R \\
\frac{X \rightarrow A \quad Y[B] \rightarrow C}{Y[B / A \circ X] \rightarrow C} /L \quad \frac{X \circ A \rightarrow B}{X \rightarrow B / A} /R \\
\frac{X[A \circ B] \rightarrow C}{X[A \bullet B] \rightarrow C} \bullet L \quad \frac{X \rightarrow A \quad Y \rightarrow B}{X \circ Y \rightarrow A \bullet B} \bullet R \\
\frac{X \rightarrow A \quad Y[B] \rightarrow C}{Y[X \circ_i A \downarrow_i B] \rightarrow C} \downarrow_i L \quad \frac{A \circ_i X \rightarrow B}{X \rightarrow A \downarrow_i B} \downarrow_i R \\
\frac{X \rightarrow A \quad Y[B] \rightarrow C}{Y[B \uparrow_i A \circ_i X] \rightarrow C} \uparrow_i L \quad \frac{X \circ_i A \rightarrow B}{X \rightarrow B \uparrow_i A} \uparrow_i R \\
\frac{X[A \circ_i B] \rightarrow C}{X[A \odot_i B] \rightarrow C} \odot_i L \quad \frac{X \rightarrow A \quad Y \rightarrow B}{X \circ_i Y \rightarrow A \odot_i B} \odot_i R
\end{array}$$

The nondeterministic logical rules are:

$$\begin{array}{c}
\frac{X \rightarrow A \quad Y[B] \rightarrow C}{Y[X \circ_i A \downarrow_i B] \rightarrow C} \downarrow L \quad \frac{A \circ_1 X \rightarrow B \quad \dots \quad A \circ_a X \rightarrow B}{X \rightarrow A \downarrow B} \downarrow R \\
\frac{X \rightarrow A \quad Y[B] \rightarrow C}{Y[B \uparrow_i A \circ_i X] \rightarrow C} \uparrow L \quad \frac{X \circ_1 A \rightarrow B \quad \dots \quad X \circ_x A \rightarrow B}{X \rightarrow B \uparrow A} \uparrow R \\
\frac{X[A \circ_1 B] \rightarrow C \quad \dots \quad X[A \circ_a B] \rightarrow C}{X[A \odot B] \rightarrow C} \odot L \quad \frac{X \rightarrow A \quad Y \rightarrow B}{X \circ_i Y \rightarrow A \odot B} \odot R
\end{array}$$

Note that there is no structural term constructor for the nondeterministic wrapping and that the nondeterministic rules, some of which have an unbounded number of premises, are explicated in terms of deterministic wrapping. Let's see now the structural rules.

Structural rules for units

-Continuous unit:

$$\frac{T[X] \rightarrow A}{T[X \circ \mathbb{I}] \rightarrow A} \quad \frac{T[X \circ \mathbb{I}] \rightarrow A}{T[X] \rightarrow A} \quad \frac{T[X] \rightarrow A}{T[\mathbb{I} \circ X] \rightarrow A} \quad \frac{T[\mathbb{I} \circ X] \rightarrow A}{T[X] \rightarrow A}$$

- Discontinuous unit:

$$\frac{T[X] \rightarrow A}{T[X \circ_i \mathbb{J}] \rightarrow A} \quad \frac{T[X \circ_i \mathbb{J}] \rightarrow A}{T[X] \rightarrow A} \quad \frac{T[X] \rightarrow A}{T[\mathbb{J} \circ_1 X] \rightarrow A} \quad \frac{T[\mathbb{J} \circ_1 X] \rightarrow A}{T[X] \rightarrow A}$$

Continuous associativity

$$\frac{X[(T_1 \circ T_2) \circ T_3] \rightarrow D}{X[T_1 \circ (T_2 \circ T_3)] \rightarrow D} \text{Assc}_c \quad \frac{X[T_1 \circ (T_2 \circ T_3)] \rightarrow D}{X[(T_1 \circ T_2) \circ T_3] \rightarrow D} \text{Assc}_c$$

Split-wrap

$$\frac{T_1[T_2 \circ T_3] \rightarrow D}{T_1[(\mathbb{J} \circ T_3) \circ_1 T_2] \rightarrow D} \text{SW} \quad \frac{T_1[(\mathbb{J} \circ T_3) \circ_1 T_2] \rightarrow D}{T_1[T_2 \circ T_3] \rightarrow D} \text{SW}$$

$$\frac{T_1[T_2 \circ T_3] \rightarrow D}{T_1[(T_2 \circ \mathbb{J}) \circ_{t_2+1} T_3] \rightarrow D} \text{SW} \quad \frac{T_1[(T_2 \circ \mathbb{J}) \circ_{t_2+1} T_3] \rightarrow D}{T_1[T_2 \circ T_3] \rightarrow D} \text{SW}$$

Definition 1 Given the term $(T_1 \circ_i T_2) \circ_j T_3$, we say that:

- (P1) $T_2 \prec_{T_1} T_3$ iff $i + t_2 - 1 < j$.
- (P2) $T_3 \prec_{T_1} T_2$ iff $j < i$.
- (O) $T_2 \not\prec_{T_1} T_3$ iff $i \leq j \leq i + t_2 - 1$.

If $T_2 \prec_{T_1} T_3$ (respectively $T_3 \prec_{T_1} T_2$), we say that T_2 and T_3 (respectively T_3 and T_2) *permute* in T_1 . Otherwise, if (O) holds, we say that T_2 *wraps* T_3 in T_1 . Note that these three conditions are mutually exclusive and are exhaustive.

Discontinuous associativity if $T_2 \not\prec_{T_1} T_3$

$$\frac{X[(T_1 \circ_i T_2) \circ_{i+j-1} T_3] \rightarrow D}{X[T_1 \circ_i (T_2 \circ_j T_3)] \rightarrow D} \text{Assc}_d \quad \frac{X[T_1 \circ_i (T_2 \circ_j T_3)] \rightarrow D}{X[(T_1 \circ_i T_2) \circ_{i+j-1} T_3] \rightarrow D} \text{Assc}_d$$

Mixed permutation if $T_3 \prec_{T_1} T_2$

$$\frac{X[(T_1 \circ_i T_2) \circ_j T_3] \rightarrow D}{X[(T_1 \circ_j T_3) \circ_{i+t_3-1} T_2] \rightarrow D} \text{MP} \quad \frac{X[(T_1 \circ_j T_3) \circ_{i+t_3-1} T_2] \rightarrow D}{X[(T_1 \circ_i T_2) \circ_j T_3] \rightarrow D} \text{MP}$$

Mixed permutation if $T_2 \prec_{T_1} T_3$

$$\frac{X[(T_1 \circ_i T_2) \circ_j T_3] \rightarrow D}{X[(T_1 \circ_{j-t_2+1} T_3) \circ_i T_2] \rightarrow D} \text{MP} \quad \frac{X[(T_1 \circ_{j-t_2+1} T_3) \circ_i T_2] \rightarrow D}{X[(T_1 \circ_i T_2) \circ_j T_3] \rightarrow D} \text{MP}$$

We define the following translation \sharp from **mm- ω -DL $_\epsilon$** sequents to **ω -DL $_\epsilon$** sequents:

$$\begin{aligned} (T \rightarrow A)^\sharp &= T^\sharp \Rightarrow A^\sharp \\ (T_1 \circ T_2)^\sharp &= T_1^\sharp, T_2^\sharp \\ (T_1 \circ_i T_2)^\sharp &= T_1^\sharp |_i T_2^\sharp \\ \mathbb{I}^\sharp &= \Lambda \\ \mathbb{J}^\sharp &= [] \\ A^\sharp &= \sqrt[0]{A}, [], \sqrt[1]{A}, \dots, \sqrt[a]{A}, [], \sqrt[A]{A} \quad \text{if } A \text{ is a type of at least sort } 1 \\ A^\sharp &= A \quad \text{if } A \text{ is of sort } 0 \end{aligned}$$

The following lemma states that the \sharp translation is a faithful embedding and that if a multimodal sequent has a Cut-free proof in **mm- ω -DL $_\epsilon$** then its \sharp -image has a Cut-free proof in **ω -DL $_\epsilon$** .

Lemma 2 Let $\Delta \Rightarrow \vec{A}$ be a $\omega\text{-DL}_\epsilon$ hypersequent. Then, for any structural term X such that $X^\sharp = \Delta$:

$$\vdash_{\omega\text{-DL}_\epsilon} \Delta \Rightarrow \vec{A} \text{ iff } \vdash_{\mathbf{mm}\text{-}\omega\text{-DL}_\epsilon} X \rightarrow A$$

Moreover, if $\vdash_{\mathbf{mm}\text{-}\omega\text{-DL}_\epsilon^-} X \rightarrow A$ then $\vdash_{\omega\text{-DL}_\epsilon^-} \Delta \Rightarrow \vec{A}$

The proof is on page 50.

Theorem 1 $\mathbf{mm}\text{-}\omega\text{-DL}_\epsilon$ enjoys Cut-elimination.

The proof is on page 44.

Theorem 2 $\omega\text{-DL}$ enjoys Cut-elimination.

Proof Let $\Delta \Rightarrow \vec{A}$ be a provable $\omega\text{-DL}$ hypersequent. Then, by the τ embedding lemma, $\tau(\Delta) \Rightarrow \vec{\tau(A)}$ is a provable $\omega\text{-DL}_\epsilon$ hypersequent. Let T be such that $T^\sharp = \tau(\Delta)$. By the $(\cdot)^\sharp$ embedding lemma, $\vdash_{\mathbf{mm}\text{-}\omega\text{-DL}_\epsilon} T \rightarrow A$. By Cut-elimination of $\mathbf{mm}\text{-}\omega\text{-DL}_\epsilon$, this sequent has a cut-free derivation which can be lifted via $(\cdot)^\sharp$ (without using Cut!) to a cut-free $\omega\text{-DL}_\epsilon$ derivation. So, by the τ embedding lemma $\Delta \Rightarrow \vec{A}$ has a $\omega\text{-DL}$ cut-free derivation. ■

Remark 1 We have proved also the Cut elimination theorem for $\omega\text{-DL}_\epsilon$.

Proof of Theorem 1

Cut-Elimination Steps for $\mathbf{mm}\text{-}\omega\text{-DL}_\epsilon$

The proof of Cut-elimination for $\mathbf{mm}\text{-}\omega\text{-DL}_\epsilon$ is by induction on the weight of the cut formula and a subinduction on the sum of the heights of derivations of the two premises. This sum is called the *cut-height*.

Principal Cuts

– \odot case:

$$\frac{\frac{X \rightarrow A \quad Y \rightarrow B}{X \circ_i Y \rightarrow A \odot B} \odot R \quad \frac{Z[A \circ_1 B] \rightarrow C \quad \cdots \quad Z[A \circ_a B] \rightarrow C}{Z[A \odot B] \rightarrow C} \odot L}{Z[X \circ_i Y] \rightarrow C} \text{Cut}$$

\rightsquigarrow

$$\frac{\frac{X \rightarrow A \quad Z[A \circ_i B] \rightarrow C}{Z[X \circ_i B] \rightarrow C} \text{Cut} \quad Y \rightarrow B}{Z[X \circ_i Y] \rightarrow C} \text{Cut}$$

– \uparrow case:

$$\frac{\frac{X_{\circ_1}A \rightarrow B \quad \dots \quad X_{\circ_x}A \rightarrow B}{X \rightarrow B \uparrow A} \uparrow L \quad \frac{Y \rightarrow A \quad Z[B] \rightarrow C}{Z[B \uparrow A_{\circ_i}Y] \rightarrow C} \uparrow L}{Z[X_{\circ_i}Y] \rightarrow C} Cut$$

\rightsquigarrow

$$\frac{\frac{X_{\circ_i}A \rightarrow B \quad Z[B] \rightarrow B}{Z[X_{\circ_i}A] \rightarrow C} Cut \quad Y \rightarrow A}{Z[X_{\circ_i}Y] \rightarrow C} Cut$$

– \downarrow case:

$$\frac{\frac{A_{\circ_1}Y \rightarrow B \quad \dots \quad A_{\circ_a}Y \rightarrow B}{Y \rightarrow A \downarrow B} \downarrow R \quad \frac{X \rightarrow A \quad Z[B] \rightarrow C}{Z[X_{\circ_i}A \downarrow B] \rightarrow C} \downarrow L}{Z[X_{\circ_i}Y] \rightarrow C} Cut$$

\rightsquigarrow

$$\frac{\frac{A_{\circ_i}Y \rightarrow B \quad Z[B] \rightarrow C}{Z[A_{\circ_i}Y] \rightarrow C} Cut \quad X \rightarrow A}{Z[X_{\circ_i}Y] \rightarrow C} Cut$$

– Units:

$$\frac{\frac{\frac{}{\mathbb{I} \rightarrow I} IR \quad \frac{T[\mathbb{I}] \rightarrow A}{T[I] \rightarrow A} IL}{T[\mathbb{I}] \rightarrow A} Cut}{T[\mathbb{I}] \rightarrow A}$$

\rightsquigarrow

$$T[\mathbb{I}] \rightarrow A$$

The case of discontinuous unit is exactly like the continuous case.

Permutation Conversions

We show some permutation conversion cases. Remaining cases are completely similar. The reader should notice that in some conversions the logical rule and the Cut rule are permuted by several Cuts and the logical rule.

– \odot :

$$\frac{\frac{X[B_{\circ_1}C] \rightarrow A \quad \dots \quad X[B_{\circ_b}C] \rightarrow A}{X[B \odot C] \rightarrow A} \odot L \quad Y[A] \rightarrow D}{Y[X[B \odot C]] \rightarrow D} Cut$$

\rightsquigarrow

$$\frac{\frac{X[B \circ_1 C] \rightarrow A \quad Y[A] \rightarrow D}{Y[X[B \circ_1 C]] \rightarrow D} \text{Cut} \quad \dots \quad \frac{X[B \circ_b C] \rightarrow A \quad Y[A] \rightarrow D}{Y[X[B \circ_b C]] \rightarrow D} \text{Cut}}{Y[X[B \circ C]] \rightarrow D} \odot L$$

– ↓:

$$\frac{\frac{X \rightarrow B \quad Y[C] \rightarrow A}{Y[X \circ_i B \downarrow C] \rightarrow A} \downarrow L \quad Z[A] \rightarrow D}{Z[Y[X \circ_i B \downarrow C]] \rightarrow D} \text{Cut}$$

↔

$$\frac{X \rightarrow B \quad \frac{Y[C] \rightarrow A \quad Z[A] \rightarrow D}{Z[Y[C]] \rightarrow D} \text{Cut}}{Z[Y[X \circ_i B \downarrow C]] \rightarrow D} \downarrow L$$

– ↑:

$$\frac{X \rightarrow A \quad \frac{Y[A] \circ_1 B \rightarrow C \quad \dots \quad Y[A] \circ_S(Y[A]) B \rightarrow C}{Y[A] \rightarrow C \uparrow B} \uparrow R}{Y[X] \rightarrow C \uparrow B} \text{Cut}$$

↔

$$\frac{\frac{X \rightarrow A \quad Y[A] \circ_1 B \rightarrow C}{Y[X] \circ_1 B \rightarrow C} \text{Cut} \quad \dots \quad \frac{X \rightarrow A \quad Y[A] \circ_S(Y[A]) B \rightarrow C}{Y[X] \circ_c B \rightarrow C} \text{Cut}}{Y[X] \rightarrow C \uparrow B} \uparrow R$$

– Units:

$$\frac{\frac{X[\mathbb{I}] \rightarrow A}{X[I] \rightarrow A} \text{IL} \quad T[A] \rightarrow B}{T[X[I]] \rightarrow A} \text{Cut}$$

↔

$$\frac{X[\mathbb{I}] \rightarrow A \quad T[A] \rightarrow B}{T[X[\mathbb{I}]] \rightarrow A} \text{Cut} \quad \frac{T[X[\mathbb{I}]] \rightarrow A}{T[X[I]] \rightarrow A} \text{IL}$$

 Proofs of Main Lemmas

Proof Lemma 1.

(1) Easy induction on the length of $\omega\text{-DL}$ derivations.

(2) Induction on the length of $\omega\text{-DL}_\epsilon$ derivations. The interesting cases are those the last rule application of which has as active formula $\tau(\star A)$, where $\star \in \{\hat{\cdot}, \check{\cdot}, \hat{\cdot}^i, \check{\cdot}^i, \triangleleft^{-1}, \triangleright, \triangleright^{-1}\}$.

- Suppose the last step of the derivation is a left inference which contains as active formula $\tau(A) \odot I (= \tau(\hat{A}))$:

$$\frac{\tau(\Delta)(\overrightarrow{\tau(A)}|_1 I) \Rightarrow \overrightarrow{\tau(C)} \quad \cdots \quad \tau(\Delta)(\overrightarrow{\tau(A)}|_a I) \Rightarrow \overrightarrow{\tau(C)}}{\tau(\Delta)(\overrightarrow{\tau(A)} \odot I) \Rightarrow \overrightarrow{\tau(C)}} \odot L$$

It is straightforward to see that for every i , $1 \leq i \leq a$, if $\vdash_{\omega\text{-DL}_\epsilon}^n \tau(\Delta)(\overrightarrow{\tau(A)}|_i I) \Rightarrow \overrightarrow{\tau(C)}$, then for $1 \leq i \leq a$, $\vdash_{\omega\text{-DL}_\epsilon}^{n-1} \tau(\Delta)(\overrightarrow{\tau(A)}|_i A) \Rightarrow \overrightarrow{\tau(C)}$. Therefore by the induction hypothesis (i.h.), for every i , $\vdash_{\omega\text{-DL}} \Delta(A|_i A) \Rightarrow \overrightarrow{C}$. Then, by application of $\hat{\cdot}L$ we get the desired result.

- Suppose the last step of the derivation is a right inference with active formula $\tau(A) \odot I (= \tau(\hat{A}))$:

$$\frac{\tau(\Delta) \Rightarrow \overrightarrow{\tau(A)} \quad A \Rightarrow I}{\tau(\Delta)|_k A \Rightarrow \overrightarrow{\tau(A)} \odot I} \odot R$$

By i.h., $\vdash_{\omega\text{-DL}} \Delta \Rightarrow \overrightarrow{A}$, whence $\vdash_{\omega\text{-DL}} \Delta|_k A \Rightarrow \overrightarrow{A}$.

- Suppose the last step of the derivation is a left inference with active formula $\tau(A) \uparrow I (= \tau(\check{A}))$:

$$\frac{\tau(\Delta)(\overrightarrow{\tau(A)}) \Rightarrow \overrightarrow{\tau(C)} \quad \tau(\Gamma) \Rightarrow I}{\tau(\Delta)(\overrightarrow{\tau(A)} \uparrow I|_k \tau(\Gamma)) \Rightarrow \overrightarrow{\tau(C)}}$$

We see that if $\vdash_{\omega\text{-DL}_\epsilon} \tau(\Gamma) \Rightarrow I$, then $\Gamma = A$, for otherwise $\tau(\Gamma)$ wouldn't be a configuration of $\omega\text{-DL}$. By i.h., $\vdash_{\omega\text{-DL}} \Delta(\overrightarrow{A}) \Rightarrow \overrightarrow{C}$, and by application of the left rule for $\check{\cdot}$, $\vdash_{\omega\text{-DL}} \Delta(\overrightarrow{\check{A}}|_k A) \Rightarrow \overrightarrow{C}$.

- Suppose the last step of the derivation is a right inference with active formula $\tau(A) \uparrow I (= \tau(\check{A}))$:

$$\frac{\tau(\Gamma)|_I \Rightarrow \overrightarrow{\tau(A)} \quad \cdots \quad \tau(\Gamma)|_c I \Rightarrow \overrightarrow{\tau(A)}}{\tau(\Gamma) \Rightarrow \overrightarrow{\tau(A)} \uparrow I}$$

As before, for every i , $1 \leq i \leq c$, if $\vdash_{\omega\text{-DL}_\epsilon}^n \tau(\Gamma)|_i I \Rightarrow \overrightarrow{\tau(A)}$ then for $1 \leq i \leq c$, $\vdash_{\omega\text{-DL}_\epsilon}^{n-1} \tau(\Gamma)|_i A \Rightarrow \overrightarrow{\tau(A)}$. By, i.h., $\vdash_{\omega\text{-DL}}^{n-1} \Gamma|_i A \Rightarrow \overrightarrow{A}$. Then, by application of $\check{\cdot}R$, we get the desired result.

- Suppose the last step of the derivation is a left inference with active formula $\tau(A) \bullet J (= \tau(\triangleleft A))$:

$$\frac{\tau(\Delta_0), \sqrt[0]{\tau(A)}, \tau(\Delta_1), \dots, \overset{a-1}{\sqrt{\tau(A)}}, \tau(\Delta_a), \sqrt[0]{\tau(A)}, \sqrt[0]{J}, \tau(\Delta_{a+1}), \sqrt[1]{J}, \tau(\Delta_{a+2}) \Rightarrow \overrightarrow{\tau(C)}}{\tau(\Delta_0), \sqrt[0]{\tau(A)} \bullet J, \tau(\Delta_1), \dots, \tau(\Delta_a), \sqrt[0]{\tau(A)} \bullet J, \tau(\Delta_{a+1}), \overset{a+1}{\sqrt{\tau(A)} \bullet J}, \tau(\Delta_{a+2}) \Rightarrow \overrightarrow{\tau(C)}} \bullet L$$

As before, we see that if, for a given $n > 0$,

$$\vdash_{\omega\text{-DL}_\epsilon}^{-n} \tau(\Delta_0), \sqrt[0]{\tau(A)}, \tau(\Delta_1), \dots, \overset{a-1}{\sqrt{\tau(A)}}, \tau(\Delta_a), \sqrt[0]{\tau(A)}, \sqrt[0]{J}, \tau(\Delta_{a+1}), \sqrt[1]{J}, \tau(\Delta_{a+2}) \Rightarrow \overrightarrow{\tau(C)}$$

then:

$$\vdash_{\omega\text{-DL}_\epsilon} \tau(\Delta_0), \sqrt[0]{\tau(A)}, \tau(\Delta_1), \dots, \sqrt[a-1]{\tau(A)}, \tau(\Delta_a) \sqrt[a]{\tau(A)}, \tau(\Delta_{a+1}), \tau(\Delta_{a+2}) \Rightarrow \overrightarrow{\tau(C)}$$

Thus, by i.h, we have $\vdash_{\omega\text{-DL}^-} \Delta_0, \sqrt[0]{A}, \Delta_1, \dots, \sqrt[a-1]{A}, \Delta_a \sqrt[a]{A}, \Delta_{a+1}, \Delta_{a+2} \Rightarrow \overrightarrow{C}$.
By application of the left rule of \triangleleft , we get the desired result.

- Suppose the last step of the derivation is a right inference with active formula $\tau(A) \bullet J (= \tau(\triangleleft A))$:

$$\frac{\tau(\Delta) \Rightarrow \overrightarrow{\tau(A)} \quad \square \Rightarrow \overrightarrow{J}}{\tau(\Delta), \square \Rightarrow \overrightarrow{\tau(A) \bullet J}} \bullet R$$

By i.h, $\vdash_{\omega\text{-DL}^-} \Delta \Rightarrow \overrightarrow{A}$, and by application of the right rule for \triangleleft we get the desired result.

- Suppose the last step of the derivation is a left inference with active formula $\tau(A)/J (= \tau(\triangleleft^{-1} A))$:

$$\frac{\tau(\Delta_0), \sqrt[0]{A}, \tau(\Delta_1), \dots, \tau(\Delta_a) \sqrt[a]{A}, \tau(\Delta_{a+1}), \sqrt[a+1]{A}, \tau(\Delta_{a+2}) \Rightarrow \overrightarrow{\tau(C)} \quad \tau(\Gamma_0), \square, \tau(\Gamma_1) \Rightarrow \overrightarrow{J}}{\tau(\Delta_0), \sqrt[0]{A/J}, \tau(\Delta_1), \dots, \tau(\Delta_a) \sqrt[a]{A/J}, \tau(\Gamma_0), \tau(\Delta_{a+1}), \tau(\Gamma_1), \tau(\Delta_{a+2}) \Rightarrow \overrightarrow{\tau(C)}} /L$$

Clearly, $\tau(\Gamma_0) = \tau(\Gamma_1) = A$ for otherwise $\Gamma_0, \square, \Gamma_1$ wouldn't be a $\omega\text{-DL}$ configuration. We have then by i.h:

$$\vdash_{\omega\text{-DL}^-} \Delta_0, \sqrt[0]{A}, \Delta_1, \dots, \Delta_a \sqrt[a]{A}, \Delta_{a+1}, \sqrt[a+1]{A}, \Delta_{a+2} \Rightarrow \overrightarrow{C}$$

Thus, by application of the left rule for \triangleleft we get the desired result.

- Suppose the last step of the derivation is a right rule with active formula $\tau(A)/J (= \tau(\triangleleft^{-1} A))$:

$$\frac{\tau(\Delta), \sqrt[0]{J}, \square, \sqrt[1]{J} \Rightarrow \overrightarrow{\tau(A)}}{\tau(\Delta) \Rightarrow \overrightarrow{A/J}} /R$$

As before it's easy to see that if $\vdash_{\omega\text{-DL}_\epsilon}^n \tau(\Delta), \sqrt[0]{J}, \square, \sqrt[1]{J} \Rightarrow \overrightarrow{\tau(A)}$ for a given $n > 0$, then $\vdash_{\omega\text{-DL}_\epsilon}^{n-1} \tau(\Delta), \square \Rightarrow \overrightarrow{\tau(A)}$. By i.h., $\vdash_{\omega\text{-DL}^-} \Delta, \square \Rightarrow \overrightarrow{A}$, and then by application of the right rule for \triangleleft we get the desired result. ■

Preliminaries for the Proof of Lemma 2

Lemma 3 (Term typability) *For every configuration Δ , there exists a structural term T_Δ such that $T_\Delta^\# = \Delta$.*

Proof Lemma 3. By induction on the hypersequent configurations. ■

A normal form for structural terms

We define the relation \leftrightarrow^* on **StructTerm** as the relation such that $T \leftrightarrow^* S$, iff $S = T$ or S can be derived from T by the use of structural rules in a finite number of steps. We see that \leftrightarrow^* is reflexive, symmetric and transitive, i.e. it is an equivalence relation.

Now, consider the following reduction relation at the level of structural terms by application of structural rules:

- R1** Unit elimination: if $T = S \circ \mathbb{I}$ or $T = \mathbb{I} \circ S$ transform T to S : $T \triangleright S$. If $T = S \circ_i \mathbb{J}$ or $T = \mathbb{J} \circ_1 S$ transform T to S : $T \triangleright S$.
- R2** If $T = S \circ K$ then if $S = S_1 \circ S_2$, then $T \triangleright S_1 \circ (S_2 \circ K)$.
- R3** If $T = S \circ_i K$. If $S = S_1 \circ S_2$, then $T \triangleright (S_1 \circ_i K) \circ S_2$, if $1 \leq i \leq s_1$, or $T \triangleright S_1 \circ (S_2 \circ_{i-s_1} K)$, if $s_1 < i$.
- R4** If $T = S \circ_i K$. If $S = S_1 \circ_k S_2$, then if S_2 wraps K in S_1 , then $T \triangleright S_1 \circ_k (S_2 \circ_l K)$. Otherwise, if S_2 and K permute in S_1 , then $T \triangleright (S_1 \circ_i K) \circ_{k+s_2-1} S_2$, if $K \prec_{S_1} S_2$, i.e., if $i < k$.

\triangleright^* is defined to be the reflexive and transitive closure of \triangleright . The first reduction at **R3** is motivated by the following derivations:

$$(S_1 \circ S_2) \circ_i K \leftrightarrow^{*SW} ((\mathbb{J} \circ S_2) \circ_1 S_1) \circ_i K \leftrightarrow^{*Assd} (\mathbb{J} \circ S_2) \circ_1 (S_1 \circ_i K) \leftrightarrow^{*SW} (S_1 \circ_i K) \circ S_2$$

The second reduction at **R3** has a similar reasoning.

Normal form for structural terms: given a structural term T , we say that T is in *normal form* if no reduction can be applied to it.

Remark 2 \mathbb{I} and \mathbb{J} are in normal form. If a term T is in normal form, then all its subterms are in normal form. Moreover T is unit free. \mathbb{J} can only appear in \circ contexts, e.g. $T[S \circ \mathbb{J}]$, whereas \mathbb{I} can only appear in wrapping contexts, e.g. $T[S \circ_i \mathbb{I}]$, for $i > 0$. Terms in normal form are \circ right-associative. If $T = R \circ S$ then either $R = A$, with $A \in \mathcal{F}$, or $R = L \circ_i K$, for otherwise rule **R2** could apply. On the other hand, if $T = R \circ_i S$, then $R = L \circ_j K$, for otherwise, if R 's main term constructor was \circ , then rule **R3** would apply. Repeating these steps of reasoning (given the fact that the number of subterms of T is finite), we get that $T = (\dots (K \circ_{i_1} S_1) \circ_{i_2} S_2 \dots) \circ_{i_n} S_n$, with $K = A \in \mathcal{F}_\epsilon$, for if $K = \mathbb{J}$ then **R1** could apply. In this case applying **R4** we get $S_1 \prec_A S_2$, $S_2 \prec_{A \circ_{i_1} T_1} S_3$, \dots , $S_{n-1} \prec_{(\dots (A \circ_{i_1} S_1) \dots \circ_{i_{n-1}} S_{n-1})} S_n$. It's not difficult to see that if $(\dots (A \circ_{i_1} S_1) \circ_{i_2} S_2 \dots) \circ_{i_n} S_n$ is in normal form, then $n \leq S(A)$.

If S is a term in normal form such that $S^\sharp = A$, then $S = \mathbb{I}$. Suppose that $T = R \circ S$, then either R is a type A , or $T = ((\dots (B \circ_{i_1} S_1) \circ_{i_2} S_2 \dots) \circ_{i_n} S_n) \circ S$, with $A, B \in \mathcal{F}$. In both cases, T^\sharp couldn't be equal to the empty string Λ . Finally, if $T = A$ (with A a type) or $T = \mathbb{J}$, then $T^\sharp \neq \Lambda$. So, $T = \mathbb{I}$.

Analogously, if S is a term in normal form such that $S^\sharp = \square$ or $S^\sharp = \overline{A}$ for a given type A , then $S = \mathbb{J}$, or $S = A$ respectively.

The construction of a normal form T^* for a given term T is carried out inductively.¹⁹ We compute the normal forms of every subterm and do some reductions if necessary. As the length l of T is finite and the subterms we consider are strictly shorter than l , the computation of a normal form is terminating:

- NF1** Render T unit-free.
- NF2** If $T = \mathbb{I}$, then $T^* = \mathbb{I}$. If $T = \mathbb{J}$, then $T^* = \mathbb{J}$.
- NF3** Else if $T = A$ for a given type $A \in \mathcal{F}_\epsilon$, then $T^* = A$.
- NF4** Else if $T = S \circ K$ and $S = S_1 \circ S_2$, then do **R2** reductions until either $T \triangleright^* A \circ L$, where A is a type, or $T \triangleright^* (L_1 \circ_i L_2) \circ L$, for $i > 0$. In case $T \triangleright^* A \circ L$, $T^* = A \circ L^*$. In case $T \triangleright^* (L_1 \circ_i L_2) \circ L$, $T^* = (L_1 \circ_i L_2)^* \circ L^*$.
- NF5** Else $T = S \circ_i K$, and S 's main term constructor is \circ , then apply reduction **R3** until $T \triangleright^* (\dots (A \circ_{i_1} S_1) \circ_{i_2} S_2 \dots) \circ_{i_n} S_n$, for a given type $A \in \mathcal{F}_\epsilon$. In that case,

¹⁹ Later we see the unicity of normal forms.

apply *R4* until $T \triangleright^* (\dots (A \circ_{j_1} L_1) \circ_{j_2} S_2 \dots) \circ_{j_m} S_m$, such that $m \leq S(A)$. Then $T^* = (\dots (A \circ_{i_1} S_1^*) \circ_{i_2} S_2^* \dots) \circ_{i_m} S_m^*$.

Where Δ is a preconfiguration, we define $l(\Delta)$ to be the length of the string Δ . For example, $l(\sqrt[A]{B}) = 2$.

Lemma 4 *Where $M, N \in \mathbf{StructTerm}$, if $M^\sharp = N^\sharp$ then $M \leftrightarrow^* N$.*

Proof We prove that for every T, S in normal form and such that $T^\sharp = S^\sharp$, then $T = S$. We proceed by induction on $l(M^\sharp)$. The base case arises with $l(T^\sharp) = 0$. By remark 2, $S = \mathbb{I}$.

Inductive case: by remark 2, if one of the terms, say T , is a type A, \mathbb{I} or \mathbb{J} , then $S = T$. Now, suppose $T = T_1 \circ T_2$; then $S = S_1 \circ S_2$ for otherwise we couldn't have $T^\sharp = S^\sharp$. Suppose now that $T_1 \neq S_1$. Then, T_1^\sharp is a prefix of S_1^\sharp or inversely. Suppose the first case, then $S_1^\sharp = T_1^\sharp \cdot \Gamma$, with Γ a configuration. Let L be in normal form such that $L^\sharp = \Gamma$ (this is possible thanks to the lemma 3). $T_1 \circ L$ is in normal form, and $(T_1 \circ L)^\sharp = S_1^\sharp$. By induction hypothesis (i.h.), $S_1 = T_1 \circ L$. Contradiction, because in that case S is not in normal form (it violates right continuous associativity). If S_1 is a prefix of T_1 , again (by a similar argument) we get a contradiction. So, $T_1 = S_1$, therefore $T_2^\sharp = S_2^\sharp$, and by i.h. $T_2 = S_2$, whence $T = S$.

If $T = T_1 \circ_i T_2$ it follows that $T = (\dots ((B \circ_{i_1} Z_1) \circ_{i_2} Z_2) \dots Z_l)$ where B must be a type, because T_1 cannot contain subterms of the form $(Z_1 \circ Z_2) \circ_i Z_3$, and B couldn't be \mathbb{J} . This is due to the fact that T is in normal form. Similarly, $S = (\dots ((C \circ_{j_1} Z'_1) \circ_{j_2} Z'_2) \dots Z'_m)$. As $T^\sharp = S^\sharp$, it follows that the types B and C are equal. Again, it follows that $l = m$, $i_n = j_n$ for $1 \leq n \leq l$, and then by i.h., $Z_i = Z'_i$. So, we get the desired result.

Suppose now M, N are arbitrary terms such that $M^\sharp = N^\sharp$. Compute their normal forms M^*, N^* . As $M^{*\sharp} = M^\sharp = N^\sharp = N^{*\sharp}$, we have $M^* = N^*$, and in particular, $M^* \leftrightarrow^* N^*$. Since $M \leftrightarrow^* M^*$, $M^* \leftrightarrow^* N^*$, and $N \leftrightarrow^* N^*$, then $M \leftrightarrow^* N$ by transitivity of \leftrightarrow^* . ■

Remark 3 If T and Z are such that $T^\sharp = Z^\sharp$ then, for every type A , $T \rightarrow A$ iff $Z \rightarrow A$. By lemma 4, $T \leftrightarrow^* S$. Thus, $T \leftrightarrow S$ and by *Cut* we get $T \rightarrow A$ iff $Z \rightarrow A$.

Proof Lemma 2. If part: by induction on the length of $\mathbf{mm}\text{-}\omega\text{-DL}_\epsilon$ sequent derivations.

Only if part: the proof is by induction on the length of $\omega\text{-DL}_\epsilon$ hypersequent derivations. In the proof we will use two facts:

- For any $\Delta \in \mathcal{O}_\epsilon$, by the typability lemma 3, there exists T_Δ such that $(T_\Delta)^\sharp = \Delta$.
- By the structural identity lemma 4, if $\vdash_{\mathbf{mm}\text{-}\omega\text{-DL}_\epsilon} X \rightarrow A$, and $Y^\sharp = X^\sharp$, then $\vdash_{\mathbf{mm}\text{-}\omega\text{-DL}_\epsilon} Y \rightarrow A$.

– Axiom case:

We have: $\vdash_{\omega\text{-DL}_\epsilon} \vec{A} \Rightarrow \vec{A}$, for a given atomic type A . Let X be a structural term X such that $X^\sharp = \vec{A}$. We have $\vdash_{\mathbf{mm}\text{-}\omega\text{-DL}_\epsilon} A \rightarrow A$. Thus, $\vdash_{\mathbf{mm}\text{-}\omega\text{-DL}_\epsilon} X \rightarrow A$.

– Units:

Consider the *IL* rule:

$$\frac{\Delta_0, \Delta_1 \Rightarrow \vec{A}}{\Delta_0, I, \Delta_1} IL$$

Δ_i ($i = 1, 2$) are preconfigurations. Let $Y[I]$ be such that $(Y[I])^\sharp = \Delta_0, I, \Delta_1$. We have that $(Y[\mathbb{I}])^\sharp = \Delta_0, \Delta_1$. By induction hypothesis (i.h.), $\vdash_{\mathbf{mm}\text{-}\omega\text{-}\mathbf{DL}_\epsilon} Y[\mathbb{I}] \rightarrow A$. By application of the left rule for the continuous unit $\vdash_{\mathbf{mm}\text{-}\omega\text{-}\mathbf{DL}_\epsilon} Y[I] \rightarrow A$.

Consider now the left discontinuous unit rule:

$$\frac{\Delta_0, \Delta_1, \Delta_2 \Rightarrow \vec{A}}{\Delta_0, \sqrt[0]{J}, \Delta_1, \sqrt[1]{J}, \Delta_2 \vec{A}}$$

$\Delta_0, \square, \Delta_2$ and Δ_1 are configurations. Let T and S be such that $T^\sharp = \Delta_0, \square, \Delta_2$ and $S^\sharp = \Delta_1$. Then, there is an i with $1 \leq i \leq t$ such that $(T \circ_i S)^\sharp = \Delta_0, \Delta_1, \Delta_2$. By i.h., $\vdash_{\mathbf{mm}\text{-}\omega\text{-}\mathbf{DL}_\epsilon} A. T \circ_i S \leftrightarrow^* T \circ_i (J \circ_1 S)$ by the structural discontinuous unit rule. Thus $\vdash_{\mathbf{mm}\text{-}\omega\text{-}\mathbf{DL}_\epsilon} T \circ_i (J \circ_1 S) \rightarrow A$ by the left rule of the discontinuous unit, and moreover $(T \circ_i (J \circ_1 S))^\sharp = \Delta_0, \sqrt[0]{J}, \Delta_1, \sqrt[1]{J}, \Delta_2$.

– Continuous case:

We consider the $/$ case. The cases of the remaining continuous proceed in a similar way.

Suppose the last rule is $/L$:

$$\frac{\Delta \Rightarrow \vec{A} \quad \Gamma(\vec{B}) \Rightarrow \vec{C}}{\Gamma(\vec{B}/\vec{A}, \Delta) \Rightarrow \vec{C}}$$

Let $T[B]$ be such that $(T[B])^\sharp = \Delta(\vec{B})$. By i.h., $T[B] \rightarrow C$ and $T_\Delta \rightarrow A$. By the left rule of $/$, $\vdash T[B/A \circ T_\Delta] \rightarrow C$, and $(T[B/A \circ T_\Delta])^\sharp = \Gamma(\vec{B}/\vec{A}, \Delta)$.

Suppose the last rule is $/R$:

$$\frac{\Delta, \vec{A} \Rightarrow \vec{B}}{\Delta \Rightarrow \vec{B}/\vec{A}}$$

Let T_Δ be such that $(T_\Delta)^\sharp = \Delta$. Since $(T_\Delta \circ A)^\sharp = \Delta, \vec{A}$, A is at the periphery, and then by the right rule of $/$, $\vdash T_\Delta \rightarrow B/A$.

– Discontinuous connectives:

We consider the nondeterministic cases, of which we see representative examples.

Deterministic continuous connectives \downarrow_j, \odot_k and \uparrow_i are similar.

Suppose the last rule is nondeterministic \uparrow on the right:

$$\frac{\Delta|_1 \vec{A} \Rightarrow \vec{B} \quad \dots \quad \Delta|_b \vec{A} \Rightarrow \vec{B}}{\Delta \Rightarrow \vec{B} \uparrow \vec{A}} \uparrow R$$

Let T be such that $T^\sharp = \Delta$. $(T \circ_i A)^\sharp = \Delta|_i \vec{A}$. Let Z_i be such that $(Z_i)^\sharp = \Delta|_i \vec{A}$. Then $Z_i \leftrightarrow^* T \circ_i A$. By i.h., $\vdash_{\mathbf{mm}\text{-}\omega\text{-}\mathbf{DL}_\epsilon} Z_i \rightarrow A$, so that $\vdash_{\mathbf{mm}\text{-}\omega\text{-}\mathbf{DL}_\epsilon} T \circ_i A \rightarrow B$, for all $i, 1 \leq i \leq b$. By application of $\uparrow R$, we get $\vdash_{\mathbf{mm}\text{-}\omega\text{-}\mathbf{DL}_\epsilon} T \rightarrow B \uparrow A$.

The discontinuous product on the left:

$$\frac{\Delta(\vec{A}|_1 \vec{B}) \Rightarrow \vec{C} \quad \dots \quad \Delta(\vec{A}|_a \vec{B}) \Rightarrow \vec{C}}{\Delta(\vec{A} \odot \vec{B}) \Rightarrow \vec{C}}$$

Let $T[A \odot B]$ be such that $(T[A \odot B])^\sharp = \Delta(\vec{A} \odot \vec{B})$. For every $i, 1 \leq a$ $(T[A \circ_i B])^\sharp = \Delta(\vec{A}|_i \vec{B})$. By i.h., $\vdash_{\mathbf{mm}\text{-}\omega\text{-}\mathbf{DL}_\epsilon} T[A \circ_i B] \rightarrow C$ for every $1 \leq i \leq a$. By application of the left rule of \odot we get $\vdash_{\mathbf{mm}\text{-}\omega\text{-}\mathbf{DL}_\epsilon} T[A \odot B] \rightarrow C$.

– Finally, the Cut rule case is obvious. ■

Acknowledgments

Thanks to the audiences at King's College, London, September 2006 and June 2008, School of Informatics, University of Edinburgh, December 2006, International Workshop on Computational Semantics, Tilburg University, January 2007, Rencontres autour des formalismes de Lambek, Toulouse University, June 2007, and Second Workshop on Types, Logic and Grammar, Universitat Politècnica de Catalunya, December 2007.

References

1. Emmon Bach. Discontinuous constituents in generalized categorial grammars. In V.A. Burke and J. Pustejovsky, editors, *Proceedings of the 11th Annual Meeting of the North Eastern Linguistics Society, New York*, pages 1–12. GLSA Publications, Department of Linguistics, University of Massachusetts at Amherst, Amherst, Massachusetts, 1981.
2. Emmon Bach. Some Generalizations of Categorial Grammars. In Fred Landman and Frank Veltman, editors, *Varieties of Formal Semantics: Proceedings of the Fourth Amsterdam Colloquium*, pages 1–23. Foris, Dordrecht, 1984. Reprinted in Walter J. Savitch, Emmon Bach, William Marsh and Gila Safran-Naveh, editors, 1987, *The Formal Complexity of Natural Language*, D. Reidel, Dordrecht, 251–279.
3. W. Buszkowski. Compatibility of categorial grammar with an associated category system. *Zeitschrift für mathematische Logik und Grundlagen der Mathematik*, 28:539–548, 1982.
4. Mike Calcagno. A Sign-Based Extension to the Lambek Calculus for Discontinuous Constituency. *Bulletin of the IGPL*, 3(4):555–578, 1995.
5. Noam Chomsky. *The Logical Structure of Linguistic Theory*. PhD thesis, Harvard University, 1955. Excerpt from 1956 revision published by Plenum, New York, 1975, and University of Chicago Press, Chicago, 1985.
6. Noam Chomsky. *Syntactic Structures*. Mouton, The Hague, 1957.
7. Noam Chomsky. *Aspects of the Theory of Syntax*. MIT Press, Cambridge, MA, 1965.
8. C.D. Culy. The Complexity of the Vocabulary of Bambara. *Linguistics and Philosophy*, 8:345–351, 1985. Reprinted in Walter J. Savitch, Emmon Bach, William Marsh and Gila Safran-Naveh, editors, 1987, *The Formal Complexity of Natural Language*, D. Reidel, Dordrecht, 349–357.
9. Mario Fadda and Glyn Morrill. The Lambek Calculus with Brackets. In C. Casadio, P.J. Scott, and R.A.G. Seely, editors, *Language and Grammar: Studies in Mathematical Linguistics and Natural Language*, number 168 in CSLI Lecture Notes, pages 113–128. CSLI Publications, Stanford, 2005.
10. Gerald Gazdar, Ewan Klein, Geoffrey Pullum, and Ivan Sag. *Generalized Phrase Structure Grammar*. Basil Blackwell, Oxford, 1985.
11. J.-Y. Girard. Linear logic. *Theoretical Computer Science*, 50:1–102, 1987.
12. Yosef Grodzinsky and Tanya Reinhart. The Innateness of Binding and Coreference. *Linguistic Inquiry*, 24(1):69–101, 1993.
13. Petra Hendriks. *Comparatives and Categorial Grammar*. PhD thesis, Rijksuniversiteit Groningen, Groningen, 1995.
14. Mark Hepple. *The Grammar and Processing of Order and Dependency*. PhD thesis, University of Edinburgh, 1990.
15. M.A.C. Huybregts. Overlapping dependencies in Dutch. *Utrecht Working Papers in Linguistics*, 1:24–65, 1976.
16. Riny Huybregts. The weak inadequacy of Context-Free Phrase Structure Grammars. In Ger Jan de Haan, Mieke Trommelen, and Wim Zonneveld, editors, *Van Periferie naar Kern*, pages 81–99. Foris Publications, Dordrecht, 1985.
17. Pauline Jacobson. Towards a variable-free semantics. *Linguistics and Philosophy*, 22(2):117–184, 1999.
18. Gerhard Jäger. Anaphora and quantification in categorial grammar. In Michael Moortgat, editor, *Logical Aspects of Computational Linguistics*, number 2014 in Lecture Notes in Artificial Intelligence, pages 70–89, Berlin, 2001. Springer.

-
19. J. Lambek. Categorical and Categorical Grammars. In Richard T. Oehrle, Emmon Bach, and Deidre Wheeler, editors, *Categorical Grammars and Natural Language Structures*, volume 32 of *Studies in Linguistics and Philosophy*, pages 297–317. D. Reidel, Dordrecht, 1988.
 20. Joachim Lambek. The mathematics of sentence structure. *American Mathematical Monthly*, 65:154–170, 1958. Reprinted in Buszkowski, W., W. Marciszewski, and J. van Benthem, editors, 1988, *Categorical Grammar*, Linguistic & Literary Studies in Eastern Europe volume 25, John Benjamins, Amsterdam, 153–172.
 21. Richard Montague. Universal grammar. *Theoria*, 36:373–398, 1970. Reprinted in R.H. Thomason, editor, 1974, *Formal Philosophy: Selected Papers of Richard Montague*, Yale University Press, New Haven, 222–246.
 22. Richard Montague. The Proper Treatment of Quantification in Ordinary English. In J. Hintikka, J.M.E. Moravcsik, and P. Suppes, editors, *Approaches to Natural Language: Proceedings of the 1970 Stanford Workshop on Grammar and Semantics*, pages 189–224. D. Reidel, Dordrecht, 1973. Reprinted in R.H. Thomason, editor, 1974, *Formal Philosophy: Selected Papers of Richard Montague*, Yale University Press, New Haven, 247–270.
 23. Michael Moortgat. *Categorical Investigations: Logical and Linguistic Aspects of the Lambek Calculus*. Foris, Dordrecht, 1988. PhD thesis, Universiteit van Amsterdam.
 24. Michael Moortgat. Generalized Quantification and Discontinuous Type Constructors. *Manuscript, Universiteit Utrecht*, 1991. Published in H. Bunt, editor, *Discontinuous Constituency*, De Gruyter, Berlin, 1996.
 25. Michael Moortgat. Multimodal linguistic inference. *Journal of Logic, Language and Information*, 5:349–385, 1995. Also in *Bulletin of the IGPL*, 3(2,3):371–401, 1995.
 26. Michael Moortgat. In situ binding: a modal analysis. In P. Dekker and M. Stokhof, editors, *Proceedings of the 10th Amsterdam Colloquium*, pages 235–240, Amsterdam, 1996. ILLC, Universiteit van Amsterdam.
 27. Michael Moortgat. Categorical Type Logics. In Johan van Benthem and Alice ter Meulen, editors, *Handbook of Logic and Language*, pages 93–177. Elsevier Science B.V. and The MIT Press, Amsterdam and Cambridge, Massachusetts, 1997.
 28. Michael Moortgat. Constants of grammatical reasoning. In G. Bouma, E. Hinrichs, G.-J. Kruijff, and R.T. Oehrle, editors, *Constraints and Resources in Natural Language Syntax and Semantics*, pages 195–219. CSLI Publications, Stanford, 1999.
 29. Michael Moortgat and Glyn Morrill. Heads and phrases: Type calculus for dependency and constituent structure. *Manuscript, Universiteit Utrecht*, 1991.
 30. R. Moot and M. Piazza. Linguistic applications of first order intuitionistic linear logic. *Journal of Logic, Language and Information*, 10:211–232, 2001.
 31. Glyn Morrill. Intensionality and Boundedness. *Linguistics and Philosophy*, 13(6):699–726, 1990.
 32. Glyn Morrill. Categorical Formalisation of Relativisation: Pied Piping, Islands, and Extraction Sites. *Report de Recerca LSI-92-23-R, Departament de Llenguatges i Sistemes Informàtics, Universitat Politècnica de Catalunya*, 1992.
 33. Glyn Morrill. Discontinuity in Categorical Grammar. *Linguistics and Philosophy*, 18(2):175–219, 1995.
 34. Glyn Morrill. Proof Syntax of Discontinuity. In Paul Dekker, Martin Stokhof, and Yde Venema, editors, *Proceedings of the 11th Amsterdam Colloquium*, pages 235–240, Universiteit van Amsterdam, 1997. Institute for Logic, Language and Computation, ILLC.
 35. Glyn Morrill. Relational interpretation and geometrical form. In V. Michele Abrusci and Claudia Casadio, editors, *Dynamic Perspectives in Logic and Linguistics*, pages 145–182. Bulzoni, 1999.
 36. Glyn Morrill. Dutch Word Order and Binding. *Report de Recerca LSI-00-59-R, Departament de Llenguatges i Sistemes Informàtics, Universitat Politècnica de Catalunya*, 2000.
 37. Glyn Morrill. Incremental Processing and Acceptability. *Computational Linguistics*, 26(3):319–338, 2000.
 38. Glyn Morrill. Islands, Coordination and Parasitic Gaps. In V.M. Abrusci and C. Casadio, editors, *New Perspectives in Logic and Formal Linguistics, Proceedings Vth Roma Workshop*. Bulzoni Editore, Roma, 2002. Also Report de Recerca LSI-02-16-R, Departament de Llenguatges i Sistemes Informàtics, Universitat Politècnica de Catalunya.
 39. Glyn Morrill. Towards Generalised Discontinuity. In Gerhard Jäger, Paula Monachesi, Gerald Penn, and Shuly Wintner, editors, *Proceedings of the 7th Conference on Formal Grammar*, pages 103–111, Trento, 2002. ESSLLI.

40. Glyn Morrill. On Bound Anaphora in Type Logical Grammar. In Geert-Jan M. Kruijff and Richard T. Oehrle, editors, *Resource-Sensitivity, Binding and Anaphora*, volume 80 of *Studies in Linguistics and Philosophy*, pages 159–177. Kluwer Academic Press, Dordrecht, 2003.
41. Glyn Morrill. Syntax and Semantics: Formal Approaches. In Lynn Nadal, editor, *Encyclopedia of Cognitive Science*, volume 4, pages 338–345. Nature Publishing Group, Macmillan Publishers Ltd, London, 2003.
42. Glyn Morrill. Geometry of Language and Linguistic Circuitry. In C. Casadio, P.J. Scott, and R.A.G. Seely, editors, *Language and Grammar: Studies in Mathematical Linguistics and Natural Language*, number 168 in CSLI Lecture Notes, pages 237–264. CSLI Publications, Stanford, 2005.
43. Glyn Morrill and Mario Fadda. Proof Nets for Basic Discontinuous Lambek Calculus. *Logic and Computation*, pages 239–256, 2008.
44. Glyn Morrill, Mario Fadda, and Oriol Valentín. Nondeterministic Discontinuous Lambek Calculus. In Jeroen Geertzen, Elias Thijsse, Harry Bunt, and Amanda Schiffrin, editors, *Proceedings of the Seventh International Workshop on Computational Semantics, IWCS-7*, pages 129–141. Tilburg University, 2007.
45. Glyn Morrill and Anna Gavarró. Catalan Clitics. In Alain Lecomte, editor, *Word Order in Categorical Grammar / L'Ordre des mots dans les grammaires catégorielles*, pages 211–232. Éditions Adosa, Clermont-Ferrand, 1992.
46. Glyn Morrill, Neil Leslie, Mark Hepple, and Guy Barry. Categorical deductions and structural operations. In G. Barry and G. Morrill, editors, *Studies in Categorical Grammar*, volume 5 of *Edinburgh Working Papers in Cognitive Science*. Centre for Cognitive Science, University of Edinburgh, 1990. Revised version published as G. Barry, M. Hepple, N. Leslie and G. Morrill, Proof figures and structural operators for categorial grammar, in *Proceedings of the Fifth Conference of the European Chapter of the Association for Computational Linguistics*, Berlin, 1991.
47. Glyn Morrill and Josep-Maria Merenciano. Generalising discontinuity. *traitement automatique des langues*, 37(2):119–143, 1996.
48. Glyn Morrill and M. Teresa Solias. Tuples, Discontinuity and Gapping in Categorical Grammar. In *Proceedings of the European Chapter of the Association for Computational Linguistics, EACL93*, pages 287–297, Utrecht, 1993.
49. Glyn V. Morrill. *Type Logical Grammar: Categorical Logic of Signs*. Kluwer Academic Press, Dordrecht, 1994.
50. A. Ojeda. Discontinuous Dependencies. In Keith Brown, editor, *Encyclopedia of Language & Linguistics, Second Edition*, volume 3, pages 624–630. Elsevier, Amsterdam, 2006.
51. M. Pentus. Lambek calculus is L-complete. *ILLC Report, University of Amsterdam*, 1993. Shortened version published as Language completeness of the Lambek calculus, *Proceedings of the Ninth Annual IEEE Symposium on Logic in Computer Science*, Paris, pages 487–496, 1994.
52. Carl Pollard. *Generalized Phrase Structure Grammars, Head Grammars and Natural Language*. PhD thesis, Stanford University, 1984.
53. Carl Pollard and Ivan A. Sag. *Head-Driven Phrase Structure Grammar*. The University of Chicago Press, Chicago, 1994.
54. M. Reape. *A Formal Theory of Word Order: A Case Study in West Germanic*. PhD thesis, University of Edinburgh, 1993.
55. D. Roorda. *Resource Logics. Proof-theoretical Investigations*. PhD thesis, Universiteit van Amsterdam, 1991.
56. Stuart Shieber. Evidence Against the Context-Freeness of Natural Language. *Linguistics and Philosophy*, 8:333–343, 1985. Reprinted in Walter J. Savitch, Emmon Bach, William Marsh and Gila Safran-Naveh, editors, 1987, *The Formal Complexity of Natural Language*, D. Reidel, Dordrecht, 320–334.
57. M. Teresa Solias Arís. *Gramáticas Catoriales, Coordinación Generalizada y Elisión*. PhD thesis, Universidad Autónoma de Madrid, 1992. Revised version published as *Gramática categorial: Modelos y aplicaciones*, Editorial Síntesis, Madrid, 1996.
58. Mark Steedman. Dependency and Coordination in the Grammar of Dutch and English. *Language*, 61:523–568, 1985.
59. Mark Steedman. Gapping as Constituent Coordination. *Linguistics and Philosophy*, 13(2):207–263, 1990.
60. Oriol Valentín. 1-Discontinuous Lambek Calculus: Type Logical Grammar and discontinuity in natural language. *DEA dissertation, Universitat Autònoma de Barcelona*, 2006. <http://seneca.uab.es/ggt/tesis.htm>.

-
61. J. van Benthem. *Language in action: Categories Lambdas, and Dynamic Logic*, volume 130 of *Studies in Logic and the Foundations of Mathematics*. North-Holland, Amsterdam, 1991.
 62. Johan van Benthem. The Semantics of Variety in Categorical Grammar. *Report 83-29*, 1983, Department of Mathematics, Simon Fraser University. Published in Buszkowski, W., W. Marciszewski, and J. van Benthem, editors, 1988, *Categorical Grammar*, Linguistic & Literary Studies in Eastern Europe Volume 25, John Benjamins, Amsterdam, 37–55.
 63. Johan van Benthem. The Categorical Fine-Structure of Natural Language. In C. Casadio, P.J. Scott, and R.A.G. Seely, editors, *Language and Grammar: Studies in Mathematical Linguistics and Natural Language*, number 168 in CSLI Lecture Notes, pages 3–29. CSLI Publications, Stanford, 2005.
 64. Koen Versmissen. Discontinuous Type Constructors in Categorical Grammar. *MSc. dissertation, Universiteit Utrecht*, 1991.
 65. V. Yngve. A model and an hypothesis for language structures. In *Proceedings of the American Philosophical Society*, volume 104, pages 444–466, 1960.
 66. W. Zielonka. Axiomatizability of Ajdukiewicz-Lambek Calculus by Means of Cancellation Schemes. *Zeitschrift für mathematische Logik und Grundlagen der Mathematik*, 27:215–224, 1981.