# **Towards Generalised Discontinuity**

#### **GLYN MORRILL**

ABSTRACT. We present a formulation of discontinuity in type logical grammar in which separation points are encoded not by a binary operation, as in earlier treatments, but by a nullary operation (placeholder); this novelty appears to provide the key to enabling multiple discontinuities in a natural algebra.

### 1.1 Introduction

A number of proposals have been made for discontinuity in type logical grammar, e.g. Moortgat (1988, 1991/96), Versmissen (1991), Solias (1992), Morrill and Solias (1993), Morrill (1994, 1995), Calcagno (1995), Hendriks (1995) and Morrill and Merenciano (1996). Notwithstanding their various merits, none is fully "generalised" in the sense of supporting multiple discontinuities and non-deterministic wrapping. Here, we present a formulation which is generalised in this respect. Separation points are encoded not by a binary operation, as in earlier treatments, but by a nullary operation (placeholder); this novelty appears to provide the key to enabling multiple discontinuities in a natural algebra. We work on a design of *sorted residuation* and formulate a labelled natural deduction calculus.

A minimal example of discontinuity is proffered by a discontinuous idiom such as 'give ... the cold shoulder':

(1.1) Mary gave  $\{John/the man\}$  the cold shoulder.

We want to associate with the form a unitary lexical meaning ("shun") and characterise it as wrapping around its object. A similar discontinuity can occur with particle verbs:

(1.2) Mary rang  $\{John/the man\}$  up.

We want to associate the form 'ring up' with a unitary lexical meaning ("phone") and allow it to wrap around its object. However, in the case of particle verbs, the object can also come after the particle:

(1.3) Mary rang up  $\{John/the man\}$ .

Thus we would like to allow the object to appear in either of two positions.<sup>1</sup> In this paper we develop type logical grammar of discontinuity which allows such 'non-deterministic' wrapping, at the same time as preserving the previous accounts of discontinuous phenomena such as medial extraction, quantification, pied-piping, and gapping.

#### 1.2 Residuation

A structure  $(S, \mathbf{B}, \mathbf{B}^{-1}; \leq)$  of arity (1, 1; 2) where  $\leq$  is a partial order comprises a *residuated pair* if and only if:

$$(1.4) \quad \mathbf{B}\mathbf{B}^{-1}A \leq A \leq \mathbf{B}^{-1}\mathbf{B}A$$

A structure  $(S, \mathbf{T}^{-l}, \mathbf{T}, \mathbf{T}^{-r}; \leq)$  of arity (2, 2, 2; 2) where  $\leq$  is a partial order comprises a *residuated triple* if and only if:

(1.5)  $A\mathbf{T}(A\mathbf{T}^{-l}C) \leq C \leq A\mathbf{T}^{-l}(A\mathbf{T}C)$  and  $(C\mathbf{T}^{-r}B)\mathbf{T}B \leq C \leq (C\mathbf{T}B)\mathbf{T}^{-r}B$ 

Moortgat (1997) emphasizes the possibility of defining type-constructors in type logical grammar which are residuated families. Let W be a set and  $\mathcal{P}(W)$  the powerset of W. Let R be a binary relation on W. We define unary operations on  $\mathcal{P}(W)$  by:

(1.6) 
$$\langle \rangle A = \{ w_2 | \exists w_1 [w_1 \in A \& R(w_1, w_2)] \}$$
  
 $[]^{-1}B = \{ w_1 | \forall w_2 [R(w_1, w_2) \Rightarrow w_2 \in B] \}$ 

Then  $(\mathcal{P}(W), \langle \rangle, []^{-1}; \subseteq)$  comprises a residuated pair. Let R be a ternary relation on W. We define binary operations on  $\mathcal{P}(W)$  by:

(1.7)  $\begin{array}{rcl} A \cdot B &=& \{w_3 | \exists w_1, w_2[w_1 \in A \& w_2 \in B \& R(w_1, w_2, w_3)]\} \\ A \to C &=& \{w_2 | \forall w_1, w_3[w_1 \in A \& R(w_1, w_2, w_3) \Rightarrow w_3 \in C]\} \\ C \leftarrow B &=& \{w_1 | \forall w_2, w_3[w_2 \in B \& R(w_1, w_2, w_3) \Rightarrow w_3 \in C]\} \end{array}$ 

Then  $(\mathcal{P}(W), \rightarrow, \cdot, \leftarrow; \subseteq)$  comprises a residuated triple.

#### **1.3** Sorted residuation

Let W be a set and  $\{W_i\}_{i \in \Sigma}$  a *partition* of W, i.e. an indexed family of pairwise disjoint sets, the union of which is W.

<sup>&</sup>lt;sup>1</sup>In the case that the object is a pronoun, it is only acceptable after the particle under certain prosodic and/or semantic conditions, for example with heavy stress and/or deictic use. We do not address this here; we take as our objective the characterisation of a free alternation.

Let R be a binary relation on W of *relationality*  $\sigma_1, \sigma_2$ , i.e.:

(1.8)  $R \subseteq \bigcup_{\sigma_1, \sigma_2 \in \Sigma} (W_{\sigma_1} \times W_{\sigma_2})$ 

We define sorted unary operations on  $\{\mathcal{P}(W_i)\}_{i \in \Sigma}$  by:

(1.9) 
$$\langle \rangle_{\sigma_1 \to \sigma_2} A = \{ w_2 \in W_{\sigma_2} | \exists w_1 [w_1 \in A \& R(w_1, w_2)] \}$$
  
 $[]_{\sigma_2 \to \sigma_1}^{-1} B = \{ w_1 \in W_{\sigma_1} | \forall w_2 [R(w_1, w_2) \Rightarrow w_2 \in B] \}$ 

We refer to the unary operations as a sorted residuated pair.

Let R be a ternary relation on W of *relationality*  $\sigma_1, \sigma_2, \sigma_3$ , i.e.:

$$(1.10)R \subseteq \bigcup_{\sigma_1, \sigma_2, \sigma_3 \in \Sigma} (W_{\sigma_1} \times W_{\sigma_2} \times W_{\sigma_3})$$

We define sorted binary operations on  $\mathcal{P}(W)$  by:

$$\begin{array}{rcl} (1.11) & A \cdot_{\sigma_{1},\sigma_{2}\to\sigma_{3}} B &= & \{w_{3}\in W_{\sigma_{3}} | \\ & & \exists w_{1},w_{2}[w_{1}\in A \& w_{2}\in B \& R(w_{1},w_{2},w_{3})]\} \\ & A \to_{\sigma_{1},\sigma_{3}\to\sigma_{2}} C &= & \{w_{2}\in W_{\sigma_{2}} | \\ & & \forall w_{1},w_{3}[w_{1}\in A \& R(w_{1},w_{2},w_{3})\Rightarrow w_{3}\in C]\} \\ & & C \leftarrow_{\sigma_{3},\sigma_{2}\to\sigma_{1}} B &= & \{w_{1}\in W_{\sigma_{1}} | \\ & & \forall w_{2},w_{3}[w_{2}\in B \& R(w_{1},w_{2},w_{3})\Rightarrow w_{3}\in C]\} \end{array}$$

We refer to the binary operations as a *sorted residuated triple*.

Sorting in type logical grammar is suggested in Morrill and Merenciano (1996). Sorted residuation adds structure to the residuation of section 1.2, which can be seen as the special case of sorted residuation in which there is one sort. We will formulate type logical grammar of discontinuity within the framework of sorted residuation.

#### 1.4 Discontinuity

Let there be a *vocabulary* V which is a set with a distinguished *separator*  $\$ \in V$ . Then there is the algebra (L, +, \$) where L is the set of non-empty strings over V and + is the operation of concatenation. We have  $s_1+(s_2+s_3) = (s_1+s_2)+s_3$ , i.e. concatenation is *associative* and we can omit its parentheses.<sup>2</sup>

Let  $|s|_{\$}$  be the number of \$'s in  $s \in L$ .<sup>3</sup> We define a partition  $\{L_i\}_{i \in \mathcal{N}}$  of L thus:

 $(1.12)L_i = \{s \in L \mid |s|_{\$} = i\}$ 

 $<sup>^{2}(</sup>L, +)$  is a free semigroup.

 $<sup>||</sup>_{\$}$  is a homomorphism from the algebra (L, +, \$) to the algebra of naturals  $(\mathcal{N}, +, 1)$ .

We define a binary relation U ('fusion') on L by:

(1.13)  $U(s_1, s_2)$  if and only if  $s_1 = \$ + s_2 \lor \exists s, s' \in L[s_1 = s + \$ + s' \land s_2 = s + s'] \lor s_1 = s_2 + \$$ 

I.e  $U(s_1, s_2)$  means that  $s_2$  is the result of removing a \$ from  $s_1$ . The relation U is of relationality i+1, i.

Let there be some *atomic types*, e.g. S for declarative sentence, N for proper name and CN for count noun. Let there be a *sort map* S from atomic types to sorts, e.g. S(S) = S(N) = S(CN) = 0. We define type formulas  $\mathcal{F}_i$  for each sort *i* by unary operators  $\hat{}$  ("bridge") and  $\hat{}$  ("split"):

(1.14) 
$$\mathcal{F}_0 ::= \mathbf{S} \mid \mathbf{N} \mid \mathbf{CN}$$
  
 $\mathcal{F}_i ::= \hat{\mathcal{F}}_{i+1} \qquad \mathcal{F}_{i+1} ::= \hat{\mathcal{F}}_i$ 

Let each atomic type P have an interpretation  $\llbracket P \rrbracket \subseteq L_{S(P)}$ . Then we define the interpretation of types preserving the relation  $\llbracket A \rrbracket \subseteq L_{S(A)}$  between the sort map and interpretation for all types A:

(1.15) 
$$\llbracket^{A}\rrbracket = \{s_{2} | \exists s_{1} \in \llbracket A \rrbracket, U(s_{1}, s_{2})\}$$
  
=  $\{s_{2} | s_{2}$  is the result of removing a \$ from an  $A\}$ ;  
$$\llbracket^{B}\rrbracket = \{s_{1} | \forall s_{2}, U(s_{1}, s_{2}) \Rightarrow s_{2} \in \llbracket B \rrbracket\}$$
  
=  $\{s_{1} | \text{ every removal of a $ from } s_{1} \text{ results in a } B\}$ ;

We see that bridge and split are a sorted residuated pair.

As a ternary relation, the binary operation + of concatenation is of relationality i, j, i+j. We define further types by binary operators • ("product"), \ ("under") and / ("over"):

(1.16)  $\mathcal{F}_{i+j} ::= \mathcal{F}_i \bullet \mathcal{F}_j \quad \mathcal{F}_j ::= \mathcal{F}_i \setminus \mathcal{F}_{i+j} \quad \mathcal{F}_i ::= \mathcal{F}_{i+j} / \mathcal{F}_j$ 

$(1.17) [[A \bullet B]]$	=	$\{s_1+s_2 \mid s_1 \in [[A]] \text{ and } s_2 \in [[B]] \}$
	=	$\{s_3   s_3 \text{ is the concatenation of an } A \text{ and a } B\};$
$\llbracket A \setminus C \rrbracket$	=	$\{s_2   \forall s_1 \in [\![A]\!], s_1 + s_2 \in [\![C]\!]\}$
	=	$\{s_2   s_2 \text{ concatenated on the left with any } A \text{ forms a } C\};$
$\llbracket C/B \rrbracket$	=	$\{s_1   \forall s_2 \in [\![B]\!], s_1 + s_2 \in [\![C]\!]\}$
	=	$\{s_1   s_1 \text{ concatenated on the right with any } B \text{ forms a } C\};$

We see that under, product and over are a sorted residuated triple. We define a ternary relation W ('wrap') on L by:

(1.18)  $W(s_1, s_2, s_3)$  if and only iff  $(s_1 = \$ \land s_2 = s_3) \lor \exists s, s' \in L[(s_1 = \$+s' \land s_3 = s_2+s') \lor (s_1 = s+\$+s' \land s_3 = s+s_2+s') \lor (s_1 = s+\$ \land s_3 = s+s_2)]$ 

I.e.  $W(s_1, s_2, s_3)$  means that  $s_3$  is the result of replacing a \$ in  $s_1$  by  $s_2$ . The

ternary relation W is of relationality i+1, j, i+j. We define further types by binary operators  $\odot$  ("discontinuous product")  $\downarrow$  ("infix") and  $\uparrow$  ("extract"):

(1.19)  $\mathcal{F}_{i+j} ::= \mathcal{F}_{i+1} \odot \mathcal{F}_j \quad \mathcal{F}_j ::= \mathcal{F}_{i+1} \downarrow \mathcal{F}_{i+j} \quad \mathcal{F}_{i+1} ::= \mathcal{F}_{i+j} \uparrow \mathcal{F}_j$ 

We see that infix, discontinuous product and extract are a sorted residuated triple.

### 1.5 Grammar

Let there be prosodic identifiers of sort 0. A *prosodic term* is a term built over the identifiers by the binary operator + and the nullary operator \$. Prosodic terms are sorted and interpreted in the obvious way.

A type assignment statement  $\alpha$ : A comprises a prosodic term  $\alpha$  and a type formula A of the same sort and is read as stating that  $\llbracket \alpha \rrbracket \in \llbracket A \rrbracket$ . A set S of type assignment statements *entails* a type assignment statement  $\sigma$ ,  $S \models \sigma$ , iff every interpretation making true every type assignment statement in S also makes  $\sigma$  true.

A *lexicon* is a set of type assignment statements. The *language model*  $\mathcal{L}$  defined by a lexicon *Lex* is the set of all type assignment statements entailed by *Lex*, i.e. the closure of *Lex* under entailment:

 $(1.21)\mathcal{L} = \{\sigma \mid Lex \models \sigma\}$ 

For example, let there be the following lexical assignments:

(1.22)	everyone	:	(S↑N)↓S
	gave+\$+the+cold+shoulder	:	(N\S)↑N
	John	:	Ν
	loves	:	(N S)/N
	man	:	CN
	Mary	:	Ν
	rang+\$+up+\$	:	`(N\S)↑N
	someone	:	(S↑N)↓S
	that	:	$(CN CN)^{(S\uparrow N)}$
	thinks	:	(N S)/S
	whom	:	$(N\uparrow N)\downarrow((CN\backslash CN)/^{(S\uparrow N)})$

Then the language model will include medial extraction, quantification and piedpiping.

### 1.6 Calculus

We present a labelled Prawitz-style natural deduction type assignment calculus. In the following,  $\Pi_{\mathcal{C}}(\mathcal{D})$  signifies a derivation  $\mathcal{D}$  for each of the (finite number of) instances satisfying the condition  $\mathcal{C}$ ; and  $\alpha |\beta|$  signifies a prosodic term  $\alpha$  containing a distinguished occurrence of  $\beta$ .

$$(1.23) \qquad \vdots \\ \frac{\alpha|\$|:A}{\alpha||:A} \uparrow \mathbf{I} \\ \frac{\beta:A}{\beta:A} \qquad \Pi_{\alpha \in (\mathbf{a}_{1}+)\$ \cdots \$(+\mathbf{a}_{S(A)+1})} \begin{pmatrix} \overline{\alpha:A}^{i} \\ \vdots \\ \gamma(\beta'):C \end{pmatrix}}{\gamma(\beta'):C} \uparrow \mathbf{E}^{i} \\ \frac{\Pi_{\beta s.t. U(\alpha,\beta)} \begin{pmatrix} \vdots \\ \beta:B \end{pmatrix}}{\alpha: {}^{*}B} \uparrow \mathbf{E} \\ \frac{\beta|\$|:B}{\beta||:B} \uparrow \mathbf{E} \\ (1.24) \\ \vdots \\ \frac{\alpha:A - \beta:B}{\alpha+\beta:A \bullet B} \bullet \mathbf{I} \\ \frac{\beta(\$|:A \bullet B - \Pi_{\alpha \in (\mathbf{a}_{1}+)\$ \cdots \$(+\mathbf{a}_{S(A)+1})}}{\beta \in (\mathbf{b}_{1}+)\$ \cdots \$(+\mathbf{b}_{S(B)+1})} \begin{pmatrix} \overline{\alpha:A}^{i} \\ \overline{\beta:B}^{i} \\ \vdots \\ \gamma(\alpha+\beta):C \end{pmatrix}}_{\gamma(\alpha+\beta):C} \bullet \mathbf{E}^{i} \\ \bullet \mathbf{E}^{i} \end{cases}$$

$$\begin{array}{c} \Pi \\ \alpha \in (\mathbf{a}_{1}+) \$ \cdots \$ (+\mathbf{a}_{S(A)+1}) & \begin{pmatrix} \overline{\alpha:A} & i \\ \vdots \\ \alpha+\beta:C \\ \end{pmatrix} \\ \hline \beta: A \setminus C \\ \hline \beta: A \setminus C \\ \hline \\ \frac{\alpha:A - \beta:A \setminus C}{\alpha+\beta:C} \setminus E \\ \Pi \\ \beta \in (\mathbf{b}_{1}+) \$ \cdots \$ (+\mathbf{b}_{S(B)+1}) & \begin{pmatrix} \overline{\beta:B} & i \\ \vdots \\ \alpha+\beta:C \\ \end{pmatrix} \\ \hline \\ \frac{\alpha:C/B}{\alpha+\beta:C} / I^{i} \\ \hline \\ \frac{\alpha:C/B - \beta:B}{\alpha+\beta:C} / E \\ \end{array}$$

$$\begin{array}{c} (1.25) \\ \vdots \\ \frac{\alpha:S - \beta:B}{\alpha+\beta:C} / E \\ \hline \\ \frac{\alpha:S - \beta:B}{\alpha+\beta:C} & - I \\ \hline \\ \frac{\beta:B - \beta:B}{\alpha+\beta:C} & - I \\ \hline \\ \frac{\beta:B - \beta:B}{\alpha+\beta:C} & - I \\ \hline \\ \frac{\beta:B - \beta:B}{\alpha+\beta:C} & - I \\ \hline \\ \frac{\beta:B - \beta:B}{\beta+1} & - I \\ \hline \\ \gamma(\delta):C & - E^{i} \\ \hline \end{array}$$

$$\frac{\Pi}{\alpha \in (\mathbf{a}_{1}+) \$ \cdots \$ (+\mathbf{a}_{S(A)+1})}{\gamma \ s.t. \ W(\alpha, \beta, \gamma)} \begin{pmatrix} \overline{\alpha:A}^{i} \\ \vdots \\ \gamma:C \\ \gamma:C \\ \end{pmatrix}}{\beta:A \downarrow C} \downarrow \mathbf{I}^{i}$$

$$\frac{\alpha|\$|:A \quad \beta:A \downarrow C}{\alpha|\beta|:C} \downarrow \mathbf{E}$$

$$\frac{\Pi}{\beta \in (\mathbf{b}_{1}+) \$ \cdots \$ (+\mathbf{b}_{S(B)+1})}{\gamma \ s.t. \ W(\alpha, \beta, \gamma)} \begin{pmatrix} \overline{\beta:B}^{i} \\ \vdots \\ \gamma:C \\ \gamma:C \\ \end{pmatrix}}{\gamma:C} \uparrow \mathbf{I}^{i}$$

$$\frac{\alpha|\$|:C \uparrow B \quad \beta:B}{\alpha|\beta|:C} \uparrow \mathbf{E}$$

## 1.7 Examples

The discontinuous idiom word order in (1.1) is generated from the lexical assignment gave++the+cold+shoulder : (N\S) $\uparrow$ N thus:

$$(1.26) \\ \underline{Mary: N} \frac{gave+\$+the+cold+shoulder: (N \setminus S) \uparrow N \qquad John: N}{gave+John+the+cold+shoulder: N \setminus S} \uparrow E \\ \underline{Mary+gave+John+the+cold+shoulder: S} \setminus E$$

The two particle-verb words orders in (1.2) and (1.3) are generated from a single lexical assignment  $ring+\$+up+\$: (N \setminus S)\uparrow N$  as follows:

$$(1.27) \frac{\operatorname{rang}+\$+up+\$:\check{}(N\backslash S)\uparrow N \quad John: N}{\operatorname{rang}+John+up+\$:\check{}(N\backslash S)} \uparrow E} \frac{\mathsf{Mary}: N}{\operatorname{rang}+John+up: N\backslash S} \downarrow E}{\mathsf{Mary}+\operatorname{rang}+John+up: S}$$

$$(1.28) \frac{\operatorname{rang}+\$+\operatorname{up}+\$:\check{}(N\backslash S)\uparrow N \quad \operatorname{John:} N}{\operatorname{rang}+\$+\operatorname{up}+\operatorname{John:}\check{}(N\backslash S)} \stackrel{}{}_{E}} \uparrow E}{\operatorname{Mary:} N \quad \operatorname{rang}+\operatorname{up}+\operatorname{John:} N\backslash S}_{\operatorname{Mary}+\operatorname{rang}+\operatorname{up}+\operatorname{John:} S} \backslash E}$$

Previous accounts under deterministic wrapping, which cover a wide variety of phenomena, are all preserved in our generalisation, being the special case that there is a single \$-placeholder.

#### **Bibliography**

- Calcagno, Mike (1995). A Sign-Based Extension to the Lambek Calculus for Discontinuous Constituency. *Bulletin of the Interest Group in Propositional and Predicate Logic* 3, pp. 555–578.
- Hendriks, Petra (1995). *Comparatives and Categorial Grammar*. Ph.D. dissertation, Rijksuniversiteit Groningen.
- Moortgat, Michael (1988). Categorial Investigations: Logical and Linguistic Aspects of the Lambek Calculus. Foris, Dordrecht.
- Moortgat, Michael (1991/96). Generalized Quantification and Discontinuous Type Constructors. In Bunt and van Horck, eds., *Discontinuous Constituency*, pp. 181–208. Mouton de Gruyter, Berlin.
- Moortgat, Michael (1997). Categorial Type Logics. In J. van Benthem and A. ter Meulen, eds., *Handbook of Logic and Language*, pp. 93–177. Elsevier, Amsterdam.
- Morrill, Glyn (1994). *Type Logical Grammar: Categorial Logic of Signs*. Kluwer Academic Publishers, Dordrecht.
- Morrill, Glyn (1995). Discontinuity in Categorial Grammar. *Linguistics and Philosophy* 18, pp. 175–219.
- Morrill, Glyn and Josep Maria Merenciano (1996). Generalising Discontinuity. *traitement automatique des langues* **37** 2, pp. 119–143.
- Morrill, Glyn and Teresa Solias (1993). Tuples, Discontinuity and Gapping. In *Proceedings of the Sixth Conference of the European Chapter of the Association for Computational Linguistics*, pp. 287–297. Utrecht.

- Solias, Teresa (1992). *Gramáticas Categoriales, Coordinación Generalizada y Elisión*. Ph.D. dissertation, Universidad Autónoma de Madrid.
- Versmissen, Koen (1991). Discontinuous Type Constructors in Categorial Grammar. Ms. Rijksuniversiteit Utrecht.