Modal Merge and Minimal Move for Dislocation and Verb Clustering

Crit Cremers, *Department of General Linguistics, Leiden University P.O. box 9515, 2300 RA Leiden, The Netherlands. E-mail: cremers@rullet.leidenuniv.nl*

Abstract

Apparently non-related phenomena like wh-movement and Dutch verb clustering can be handled as bounded varieties of generalized composition in a categorial framework. This supports the unification of *merge* and *move* from an operational point of view. It abstracts away from hypothetical reasoning as an instrument of grammatical deduction.

Keywords: categorial grammar, minimalism, verb clustering, wh-movement

1 Introduction

Mankind[1] owes to Generative Grammar the discovery and subsequent theoretical exploitation of at least two phenomena in natural language: leftward dislocation or fronting (1.1), in particular wh-movement, and Dutch verb clustering (1.2), in particular cross-serial dependencies. Leftward dislocation typically creates non-local dependencies between a structural position at the left edge of a sentence and a structural position somewhere in the sentence's kernel. Verb clustering groups together a bunch of verbal heads, leaving a sequence of motherless arguments to these verbs.

1.1
Hier heb ik op gewacht
'here have I on waited'
This is what I waited for
Waar heb jij hem op laten wachten?
'where have you him for let wait'
What did you let him wait for

1.2
Geen geleerde heeft *haar mij deze stelling ooit uit laten leggen*
'No scholar has her me this proposition ever ex- let -plain'
No scholar has ever allowed her to explain this proposition to me

Leftward dislocation is considered to be common to all languages and to be steered and restricted by mechanisms that reside in the very core of human grammar. This one can learn from almost any study in generative linguistics in the last decades. The study of the nature of the non-local relation created by fronting feeds the whole grammatical enterprise.

[1]Ton van der Wouden, Jeroen van de Weijer, Lisa Cheng, and an anonymous reviewer made valuable comments on earlier drafts. Maarten Hijzelendoorn co-constructed the DELILAH system. They all are innocent, though.

Verb clustering and resulting forms of cross-serial dependencies are intrinsic to only some languages, and the processes underlying it from a generative point of view are almost epiphenomena of more fundamental choices one makes. See, for instance, [Zwart 1993], [Barbiers 1998] and [Den Dikken 1996] for diverging but minimalistically inspired analyses of the clustering phenomena in Dutch. One of the major parameters here is the basic order of objects and predicates.

The categorial grammar exploited in the parsing system DELILAH (http://fonetiek-6.leidenuniv.nl/hijzlndr/delilah.html) deals with both phenomena using a limited set of combinatory instruments, under very restricted modal variation. This grammar emanates from Generalized Composition by some restrictions and additions which are discussed in this paper. It subsequently reports on the way in which the DELILAH grammar syntactically captures leftward dislocation and verb clustering phenomena, thereby demonstrating that phenomena which seem to be separate in the theory of grammar, may be connected formally. The grammar formalism introduced here exploits a concept of categories as combinatorial agendas which is far from hostile to notions like *feature checking*, *merge* and *move* in the minimalist framework (cf. [Stabler 1996]). As a matter of fact, in the rules of generalized composition proposed below notions like *merge* and *move* explicitly converge. This approach therefore suggests that projecting minimalism on multimodality may play down any fundamental distinction between these two styles of structure building.

2 Merging Move

[Stabler 1996] formalizes the two structure building operations *move* and *merge* suggested in [Chomsky 1995]. In this formalization, they differ mainly as to their arity. *Merge* operates on two phrases, *move* on just one. Both are triggered by pairs of features, construed as a negative and a positive one. *Move* is overt or covert (or partial), the difference being whether or not phonological features are involved. Stabler tends to consider this distinction as derived but [Cheng *t.a*] argues in favor of partial movement on independent grounds. [Vermaat 1999] projects Stabler's formalization on multimodal categorial grammar. She maps *merge* on functional application, and captures movement logically with hypothetical reasoning, deductively with abstraction and structurally with modalized reordering postulates. The latter process is typically induced by higher order types reflecting dislocation.
[Stabler 1996] observes instances of count invariancy in converging minimalist derivations. In the same vain, [Stabler 1999] appeals to proof net representations of these derivations, proof nets being resource sensitive construals of reasoning. In accordance with these analogies, [Vermaat 1999] translates the triggering feature pairs of minimalistic operations in terms of indexed residuation.
The distinction between the operations *merge* and *move* is rather clear-cut in Stabler's formalization. It is obscured, however, in revisions of 'the minimalistic program' emanating from exegeses of [Chomsky 1998]. Here we find more or less explicit reductions of *move* to *merge*. *Move* is reviewed as *merge* plus agree; the latter involves the determination of a local target for the cancellation of a special brand of features. Traces, the once proud markers of dislocation dependencies, are orphanized. Interestingly, [Chomsky 1998] recasts *merge* as an asymmetric operation, driven by 'probe'

and 'goal' mechanisms. This modelling reflects the emphasis in Stabler's analysis on the need for positive and negative feature pairs in both structure building rules.
If *merge* is the combinatorial engine of *move*, it is hard to see why *move* and *merge* should reflect dual operations like elimination and introduction in a categorial or deductive setting. Movement in its newest dress does not seem to be particular or distinct enough to be decorated with hypothetical reasoning as an exclusive asset. Either hypothetical reasoning is a viable move in every derivation of complex phrases - as was already immanent in [Montague 1973]'s approach to quantification - or it is obsolete in the account of linguistic phenomena. Of course, hypothetical reasoning remains an interesting mode of deduction, but grammar does not have to take resort to it in order to become relevant or adequate or faithful.
The system presented here treats *merge* and *move* phenomena as predictable varieties of one process of composing complex symbols. All structural positions are projected from the lexicon and are preserved and transferred in this process. Consequently, there are no structure building rules at all, but only resource sensitive compositional procedures steering unification: the ultimate licensing.

3 Categorial Grammar made context-sensitive

Categorial Grammar defines the set of phrases of a language by the closure of a set of lexical items with explicit combinatorial agendas under a fixed set of operations. It has been made relevant to the analysis of natural language by [Ades and Steedman 1983], [Hoeksema 1984], [Zwarts 1986] and [Moortgat 1988], among many others. Its logical dimensions have been exploited by [Lambek 1958], [Bar-Hillel *et al* 1960] and in various studies by Van Benthem (e.g. [Van Benthem 1991]) and Buszkowski (e.g. [Buszkowski 1984]). A comprehensive perspective on Categorial Grammar is offered by [Morrill 1994] and [Carpenter 1997]. The canonical forms in which the categorial approach materializes in our days include the Lambek calculus and its variations and extensions as well as Combinatory Categorial Grammar ([Steedman 1987], [Steedman 1996], [Steedman 2000]).

[Pentus 1992] first proved the longstanding conjecture that the Lambek calculus recognizes exactly the class of context-free languages. The rich variety of categorial combinations which the Lambek calculus offers does not push its recognizing capacity beyond the boundaries of context-free phrase structure grammar. This concerns in particular the status of the theorem of harmonious composition, which has the following grammatical effect:

3.1
a term *t1* of category a/b ($a\backslash b$) and a term *t2* of category b/c ($b\backslash c$) combine to form a term *t1+t2* (*t2+t1*) of category a/c ($a\backslash c$)[2]
$a/b \quad b/c \quad \Rightarrow \quad a/c$
$b\backslash c \quad a\backslash b \quad \Rightarrow \quad a\backslash c$

This option combines two 'incomplete' or functional expressions into one, but does not, according to Pentus' proof, stretch recognizing power. This is reflected in the fact that for every context-free Lambek grammar G using harmonious composition

[2] The reader may notice that this paper uses the triple notation (head,direction,argument) for complex categories

($x/y\ y/z \Rightarrow x/z$ etc.) there is a weakly equivalent context-free grammar G' using only application $x/yy \Rightarrow x$. Purely applicative systems are known to be context-free since [Bar-Hillel *et al* 1960].

Though some have suggested that natural language is '.. for the most part ..' ([Mcgee Wood 1993, p. 14]) context-free, from a linguistic point of view more than context-free power seems mandatory to capture the essence of dependencies that languages weave. [Friedman *et al.* 1986] have shown that parenthesis-free categorial grammar, i.e. a grammar in which the internal structure of categories is not marked, may have recognizing power beyond context-freeness. Similar findings were made by [Joshi *et al.* 1991], considering Combinatory Categorial Grammar in its emanation as Generalized Composition:

3.2

x/y		$(...(y \mid z_1) \mid z_2... \mid z_n)$	$\Rightarrow$	$(...(x \mid z_1) \mid z_2... \mid z_n)$
$(...(y \mid z_1) \mid z_2... \mid z_n)$		$x\backslash y$	$\Rightarrow$	$(...(x \mid z_1) \mid z_2... \mid z_n)$

x/y and $x\backslash y$ are dubbed the primary category; the other ones are secondary. A term $\mid z_i$ is a unit representing a typologically fixed argument of a category with a fixed directional marker, which may be $\backslash$ or $/$; for the persistence of directionality under composition, see [Steedman 1990]. If n = 0 in (3.2), we have simple application. Otherwise, we have the cancellation of a deeply embedded head y of the secondary category while the arguments of this category are transmitted to the reduction result, preserving their directionality (head-directed cancellation). Since the directionality of these arguments can be left or right in both instances of Generalized Composition, 3.2 merges the lambekian harmonious composition 3.1 with the non-lambekian disharmonious or mixed composition 3.3. In disharmonious composition, the 'transferred' argument is directed against the direction in which the cancellation took place.

3.3

x/y	$y\backslash z$	$\Rightarrow$	$x\backslash z$
y/z	$x\backslash y$	$\Rightarrow$	x/z

Note that the symbiosis of harmony and disharmony in 3.2 results from the rule's sensitivity to the full internal structure of the secondary category. [Joshi *et al.* 1991] shows that grammars with the combinatory potential of 3.2 are, under some provisos, weakly equivalent to Tree-Adjoining Grammars and other 'mildly' context-sensitive formalisms. [Van de Woestijne 1999] proves that the DELILAH grammar falls into that class. Consequently, the combination of disharmonious composition and head-directed cancellation, as combined in 3.2 adds new recognizing power to categorial grammar. [Moortgat 1988] and [Van Benthem 1991] notice that merely adding disharmony to the Lambek calculus is counterproductive, as the resulting system runs into permutation closure.

DELILAH's grammar exploits generalized composition a bit further by also taking into consideration the full internal structure of the primary category. In (3.2), the head x of the primary category is left unaffected by the composition. As a consequence, in the resulting category the relative ordering of the arguments stemming from the secondary category and those (hidden ones) stemming from the primary category, is fixed: each of the former ones is to be cancelled before any of the latter ones. If

we assume that x is not necessarily primitive and spell that assumption out, (3.2) becomes

3.4

$$(..(x \mid w_1) \mid .. \mid w_m)/y \quad (..(y \mid z_1) \mid .. \mid z_n) \quad \Rightarrow \quad (.((.(x \mid w_1) \mid .. \mid w_m) \mid z_1) \mid .. \mid z_n)$$

$$(..(y \mid z_1) \mid .. \mid z_n) \quad (..(x \mid w_1) \mid .. \mid w_m)\backslash y \quad \Rightarrow \quad (.((.(x \mid w_1) \mid .. \mid w_m) \mid z_1) \mid .. \mid z_n)$$

DELILAH's grammar adds two more instances of generalized composition, by just considering the possibility that the 'row' of arguments w_i, stemming from the primary category, can be more 'peripheral' in the resulting category than the row of arguments z_j, stemming from the secondary category. Under this option, each of the arguments w_i must be cancelled before any of the arguments z_j can be the target of composition.

3.5

$$(..(x \mid w_1) \mid .. \mid w_m)/y \quad (..(y \mid z_1) \mid .. \mid z_n) \quad \Rightarrow \quad (.((.(x \mid z_1) \mid .. \mid z_n) \mid w_1) \mid .. \mid w_m)$$

$$(..(y \mid z_1) \mid ... \mid z_n) \quad (..(x \mid w_1) \mid ... \mid w_m)\backslash y \quad \Rightarrow \quad (.((.(x \mid z_1) \mid .. \mid z_n) \mid w_1) \mid .. \mid w_m)$$

The additional assumption as to the transparency of the internal structure of the primary category does not induce order permutations of arguments: rows of arguments are internally unaffected and are not interwoven. There doesn't seem to be any reason to believe that natural language takes resort to 'popping' arguments from one row into the other, reversing the order of at least one stack of arguments, or with mixing rows up just because of composition. Maintaining the integrity of argument stacks under composition may be seen as a combinatorial invariant: generalized composition is conservative. [Vijay-Shanker and Weir 1993] observes that in order to prove these grammars to be reducible to Linear Indexed Grammars the indices n and m in (3.4) and (3.5) should be bound by a constant. The application of compositional modalities as described below (section 3ff.), may have this effect to some degree. Whether or not this can be made intrinsic to the grammar as such, is currently a topic of examination. Furthermore, DELILAH restricts y, the target of the cancellation, in (3.4) and (3.5) to atomic types, i.e. types from a fixed set of argument types (see also [Vijay-Shanker and Weir 1993]. [Hepple 1996] considers the compilation of general categorial grammars to this effect, from a parsing perspective. The resulting grammar may be viewed as linear or first order (cf. [König 1990]), operating on flat types in the sense of [Buszkowski 1988].

The formalisms (3.4) and (3.5) abstracts away from directionality within the rows of arguments. As noted before, however, directionality is considered to be part of the specification of an argument ([Steedman 1990]). For the purpose of operational grammar, then, it is useful to separate left and right arguments of a category. Thus, a category is considered to be a triple consisting of a head, a stack of left arguments and a stack of right arguments. The stacks are ordered, by definition, and may be empty. Only the top element of a stack is available for cancelling. The stacks of arguments are represented as Prolog lists, with the top element being the leftmost; these list are referred to by abbreviations like PL - *l*eft argument stack of the *p*rimary category) -

and SR - *r*ight argument stack (of the *s*econdary category). Categories will appear in the following formats.

3.6

$head\backslash LeftArgList/RightArgListA$
$head\backslash LeftArgList/[top|RestRightArgList]$
$head\backslash[l_1...l_m]/[r_1...r_n]$

Appended stacks are written *(Upper + Lower)*, such that [a,b] + [c,d] = [a,b,c,d]. Using the triple representation (3.6), the combinatory engine of DELILAH's grammar is stated as in (3.7), restyling 3.4 and 3.5.

3.7

EXTENDED GENERALIZED COMPOSITION

$prim\backslash PL/[sec\|PR]$	$sec\backslash SL/SR$	$\Rightarrow$	$prim\backslash PL+SL/PR+SR$
$prim\backslash PL/[sec\|PR]$	$sec\backslash SL/SR$	$\Rightarrow$	$prim\backslash SL+PL/PR+SR$
$prim\backslash PL/[sec\|PR]$	$sec\backslash SL/SR$	$\Rightarrow$	$prim\backslash PL+SL/SR+PR$
$prim\backslash PL/[sec\|PR]$	$sec\backslash SL/SR$	$\Rightarrow$	$prim\backslash PL+SL/SR+PR$
$sec\backslash SL/SR$	$prim\backslash[sec\|PL]/PR$	$\Rightarrow$	$prim\backslash PL+SL/PR+SR$
$sec\backslash SL/SR$	$prim\backslash[sec\|PL]/PR$	$\Rightarrow$	$prim\backslash SL+PL/PR+SR$
$sec\backslash SL/SR$	$prim\backslash[sec\|PL]/PR$	$\Rightarrow$	$prim\backslash PL+SL/SR+PR$
$sec\backslash SL/SR$	$prim\backslash[sec\|PL]/PR$	$\Rightarrow$	$prim\backslash SL+PL/SR+PR$

Although the stacks are marked for direction and composition is sensitive to this marking, the arguments lists in one direction do not interfere with each other. They conserve their integrity under append.

This formalism gives rise to a family of parameters: languages are not indifferent with respect to the modes of composition captured here. They may be restricted to certain patterns of merging argument stacks, for the nature of a top element, for the emptiness of stacks, and so on. In the DELILAH grammar, these parameters are operationalized as *modes*, clusters of constraints on composition. Every argument in a stack of a (lexical) category is obligatorily equipped with such a fixed mode, determining the pre- end post-conditions of its cancellation. The mode at least specifies the way of merging the argument lists 'after' cancelling the argument. It can do so by requiring one or more list to be empty 'before' cancellation and/or by specifying the merge order of non-empty stacks. Thus, a mode m at an argument A in a category C constrains the composition of C and a (secondary) category headed by A by imposing conditions on A, the other category and the composition itself. m defines a partial map of pairs of categories onto categories. Moreover, the mode may be sensitive to the history of the input categories in the sense that they may require argument stacks to be affected (non-lexical) or affected (lexical) at cancellation. The number of modes is arguably finite. Still, languages will select just a subset of these to be operative in their grammar. The broad-coverage grammar of Dutch which DELILAH aims at, actually supports only about 10 different modes; some of these modes occur in both directions, some only in one. Clearly, this system ([Cremers 1993]) is cognate with the concept of indexed operators in multimodal categorial type logic ([Moortgat 1997]. In the next sections, some aspects of global and lexical parametrization for Dutch will be discussed.

In the DELILAH system, the grammar format is parsed in a deterministic shift-reduce fashion to obtain one analysis; the set of all possible analyses for adjuncts is obtained by backtracking.
The resulting logic is essentially non-lambekian in that the only engine of inference is Modus Ponens; hypothetical reasoning is not supported. The logic consists of the identity axiom and a finite set of modalized eliminations. Capitals A, B, C stand for nonempty sequences $< A_1, ..., A_n >$ of categories A_i; capitals L, LL, R and RR stand for possibly empty lists of terms $y^{\wedge}j$, where y is a type and j a mode.

3.8

(id) $\quad x\backslash L/R \vdash x\backslash L/R$

($\backslash_i$) $\quad \dfrac{A \vdash \mathbf{y}\backslash L/R \quad B \vdash \mathbf{x}\backslash[\mathbf{y}^{\wedge}i \mid LL]/RR}{\mathrm{A} \quad , B \ \vdash \ \mathbf{x}\backslash L \oplus_i LL/R \oplus_i RR}$

($/_j$) $\quad \dfrac{A \vdash \mathbf{y}\backslash L/R \quad B \vdash \mathbf{x}\backslash LL/[\mathbf{y}^{\wedge}j \mid RR]}{B, A \vdash \mathbf{x}\backslash L \oplus_j LL/R \oplus_j RR}$

For all modes i, $\mathrm{X}\oplus_{\mathrm{i}}[\,] = [\,]\oplus_{\mathrm{i}}\mathrm{X} = \mathrm{X}$. For a given mode m, such that $\mathrm{X}\oplus_{\mathrm{m}}\mathrm{Y} = \mathrm{append(Y,X)} = \mathrm{[Y+X]}$, ($\backslash_m$) might be as in (3.9), requiring the secondary category's left argument list SL to be empty; its righthand counterpart ($/_{\mathrm{m}}$) might be as in (3.10), requiring the secondary category's right argument list SR to be non-empty and the primary category's left argument list PL to be empty.

3.9

$$\frac{A \vdash \mathbf{y}\backslash[\,]/R \quad B \vdash \mathbf{x}\backslash[\mathbf{y}^{\wedge}m \mid LL]/RR}{A, B \vdash \mathbf{x}\backslash LL/[RR + R]}$$

3.10

$$\frac{A \vdash \mathbf{y}\backslash L/[T \mid R] \quad B \vdash \mathbf{x}\backslash[\,]/[\mathbf{y}^{\wedge}m \mid RR]}{\mathrm{B},\ A \vdash \mathbf{x}\backslash L/[RR + [T \mid R]]}$$

Because of lacking operator introduction schemata, the system is not blessed with theorems like Lifting, Geach Division, Restructuring, Permutation and Exchange (cf. [Van Benthem 1991], [Moortgat 1997]). Its resource management, however, is as sensitive as in the lambekian calculi (cf. [Cremers 1993] where the present grammar format was - accidentally - introduced as 'minimal categorial grammar'). Modalized forms of disharmonious composition were introduced by [Moortgat and Oehrle 1994] in a lambekian context.
In the next sessions, the literals representing modes may be suppressed whenever the nature of a mode is topic of the argument.

4 Disharmony put to work

Given the format of extended generalized composition one can identify certain restrictions that languages may exploit. In particular, disharmonious instances of generalized composition can be held responsible for the patterns which arise from verb clustering and are known as crossing dependencies.

4.1

$[_{\mathrm{S}}..\ \mathrm{XP_1..XP_i..XP_j..XP_n}\ \ \mathrm{V_1..V_k..V_l..V_m..}]$

If XP_i is licensed by V_k and XP_j is licensed by V_l and $i < j$ and $k < l$, string (4.1) contains crossing dependencies. This type of linking is operative in Dutch sentences like 1.2, repeated below.

4.2
Geen geleerde heeft *haar mij deze stelling ooit uit laten leggen*
'No scholar has her me this proposition ever ex- let -plain'
No scholar has ever allowed her to explain this proposition to me

To the extent that crossing dependencies are productive, as they are in Dutch, they cannot be adequately captured by context-free grammars.
Let us assume - or just: observe, at the level of admissible strings - that in Dutch verbal complements are mainly right arguments and nominal complements are mainly left arguments. Now suppose we have two verbal categories adjacent to each other as in

4.3
$vp\backslash[np_1 np_2]/[vp] \quad vp\backslash[np_3]/[\,]$

The indices on *np* are for differentiation purposes only. Applying extended generalized composition to 4.3 may result in one of two reductions:

4.4
$vp\backslash[np_1 np_2]/[vp] \quad vp\backslash[np_3]/[\,] \quad \Rightarrow \quad vp\backslash[np_1 np_2 np_3]/[\,]$
$vp\backslash[np_1 np_2]/[vp] \quad vp\backslash[np_3]/[\,] \quad \Rightarrow \quad vp\backslash[np_3\ np_1 np_2]/[\,]$

As the order in the stack inverts the linear order in the string, the second option reflects the pattern of a crossing dependency: the arguments licensed by the rightmost verb have to appear to the right of the arguments licensed by the leftmost verb, the primary category. Disharmonious composition - merging stacks in the direction that was not affected by the cancellation - in this particular mode adds weak and strong recognizing power to the grammar, as was shown in [Zinn 1993].
To get Dutch verb clustering patterns right, we need at least the following merger of left argument stacks for rightward cancellation in Dutch:

4.5

$prim\backslash PL/[sec\|PR]$	$sec\backslash SL/SR$	$\Rightarrow$	$prim\backslash(SL+PL)/(SR+PR)$
laten: $vp\backslash[np]/[vp]$	*gaan*: $vp\backslash[pp]/[\,]$	$\Rightarrow$	*laten gaan*: $vp\backslash[pp, np]/[\,]$
'let'	'go'		'let go'

Now let us have a look at long-distance dependencies, i.e. the relation between a dislocated element (in [Spec, CP]) and the position in which it is licensed. Under generalized composition, this relation is established if the grammar can cancel the dislocated element against the argument that marks the gap or the trace. This trace can only occur as a left argument. It is transported leftward by merging stacks under generalized composition, but it has to be 'suppressed' with respect to other leftward arguments involved in the composition. Consider the following configuration:

4.6
wh ... $prim\backslash PL/[sec|PR] \quad sec\backslash[wh]/SR$ ...

Clearly, for the secondary category to cancel its argument *wh* (its 'gap') against the dislocated *wh*, that argument has to be carried leftward by composition. Also one can

see that the argument stack PL has to be cancelled before the dislocated argument is: wh-elements are left peripheral, when dislocated. Consequently, (4.6) can be recognized iff in the new left stack the arguments of the primary category are stacked before the gap argument. For the sake of leftward dislocation we need, along merger mode (4.5), also

4.7
$prim\backslash PL/[sec|PR] \quad sec\backslash SL/SR \quad \Rightarrow \quad prim\backslash(PL + SL)/(SR + PR)$

This merger mode might be restricted to cases where SL = *[wh]*, *i.e.* a left argument list containing only a dislocated argument. Applying composition in (4.6), then, yields as an intermediate step the category $prim\backslash(PL + [wh])/(SR + PR)$.
Thus, disharmonious composition, as an instance of Generalized Composition, derives crossing dependencies as well as long-distance discontinuities. Languages will not, however, apply composition blindly. In particular, to the extent that composition derives discontinuity, composition will be restricted by local conditions, as indicated in (3.8)-(3.10).
It must be stated that some theoreticians have refuted disharmonious composition as a legitimate way of dealing with discontinuities, for instance, [Jacobson 1992] and [Carpenter 1997, p. 201ff]. Evidently, the effects of deriving disharmony depend on its categorial environment. Disharmony is known to induce 'collapse of order' ([Carpenter 1997]) and 'damaging loss of control over the grammatical resources' ([Moortgat 1997]), when it is adopted in Lambek categorial grammar, in which it cannot be derived. Under the present assumptions, however, every composition must be constrained as to the way of merging argument lists. Here, disharmony is just one of the options out of which one has to be specified for composition to take place. It is neither added to an independently motivated repertoire nor must it be additionally restricted. Disharmony is a specified well-defined move in the game of generalized composition. In the next section, it is shown that the patterns arising from verb clustering are steered by mechanisms which also compute islandhood.

5 Verb clustering patterns captured

Given Generalized Composition, verb clustering varieties can be described by lexically assigned restrictions on the argument stacks of the primary and secondary category at composition time. To facilitate comparison, the patterns are stated in terms of a left verb with a vp-head selecting a vp complement:

5.1
$vp\backslash PL/[vp|PR] \quad vp\backslash SL/SR \quad \Rightarrow \quad vp\backslash(SL + PL)/(SR + PR)$

The following patterns must be available. It is assumed that argument stacks are marked for having been affected by a cancellation, but no such labels are specified here, for readability.

5.2
OBLIGATORY EXTRAPOSITION
SL is empty: the secondary category must be fully applied as to its left arguments; they cannot be taken over by the resulting category; no crossing dependencies can

arise

...	*ontkennen*	*mij te hebben gezien*		
	vp\[]/[*vp*]	*vp*\[]/[]	⇒	*vp*\[]/[]
	'deny	me to have seen'		

5.3

OBLIGATORY VERB RAISING

PL and SL are not yet affected in the derivation (but may be lexically empty); crossing dependencies arise when possible

...	laten	geven		
	vp\[*np*]/[*vp*]	*vp*\[*np*, *np*]/[]	⇒	*vp*\[*np*, *np*, *np*]/[]
	'let	give'		
...	*laten*	*slapen*		
	vp\[np]/[*vp*]	vp\[]/[]	⇒	vp\[np]/[]
	'let	sleep'		

5.4

INTERMEDIATE CLUSTERING, INCLUDING THIRD CONSTRUCTION

No absolute conditions on stacks are imposed; but if SL is not empty and some of its arguments are inherited by the resulting category under disharmonious composition, PL has to be unaffected hitherto; in that case, crossing dependencies arise

...	*proberen*	*een boek te geven*		
	vp\[]/[*vp*]	*vp*\[*np*]/[]	⇒	*vp*\[*np*]/[]
	'try	a book to give'		
...	*proberen*	*te geven*		
	vp\[]/[*vp*]	*vp*\[*np*, *np*]/[]	⇒	*v*\[*np*, *np*]/[]
	'try	to give'		
...	*proberen*	*haar een boek te geven*		
	vp\[]/[*vp*]	*vp*\[]/[]	⇒	*vp*\[]/[]
	'try	her a book to give'		

The variety of composition is marked in the lexicon as a property of the vp-argument in the primary category's right stack. For example: the argument of an extraposing verb is marked for cancellation against a category with an empty left stack only. The specification concerning affectedness of a stack at the Verb Raising variety is necessary: obligatory verb raising presupposes that neither the secondary nor the primary category consumed any left-hand side argument prior to the composition (for similar notions see [Houtman 1984]. The 'calculus of affectedness' is basically rather simple: lexical stacks are unaffected; a stack is affected *iff* one of its members has been cancelled; the merge of two unaffected stacks is unaffected.

Note that the third option ('anything goes *salve* crossing dependencies') also marks the way the argument of adverbials is cancelled. They will be categorized as automorphisms, e.g. *vp*\[]/[*vp*], and go along with any SL, as is argued by, e.g., [Zwart 1993] - from a generative point of view - and [Cremers 1993].

Furthermore, disharmony can be adapted for auxiliary inversion and *infinitivum-pro-participio* (*ipp*) phenomena by additional conditions on the right-hand stack of the secondary category. The first structure involves inversion of an auxiliary and the first (participle) verb of its complement. In the DELILAH grammar, this is dealt with by the requirement that the right-hand stack of the secondary category - headed by the

main verb of the complement - be unaffected. This requirement is lexically imposed upon the left *vp* argument of some auxiliaries. *Ipp*-effects follow from the requirement that the right-hand stack of the secondary category - the one associated with the infinitive - has been affected and is empty.
(5.5) and (5.6) provide, as an example, a lexical category for *willen* 'to want' and the instance of Generalized Composition dealing with the cancellation mode for its *vp*-argument. This instance of the composition rule is triggered by the mode $^\wedge vr$ on the relevant argument of the primary category. These modes have effects comparable to the effects of modal operators in categorial type logics with hypothetical reasoning (cf. [Hepple 1990], [Moortgat 1997]).

5.5
willen: $vp\backslash[np]/[vp^\wedge vr]$

5.6
$prim\backslash PL/[sec^\wedge vr|PR] \quad sec\backslash SL/SR \quad \Rightarrow \quad prim\backslash(SL+PL)/(SR+PR)$
iff PL and SL are marked 'hitherto unaffected'

Thus, DELILAH subsumes the full range of Dutch verbal combinatorics under conditions on the state of argument stacks at extended generalized composition. In fact, it describes verb clustering patterns and related phenomena by lexical stipulation in terms of the integrity of arguments: arguments (secondary categories) for Generalized Composition may, must or may not have fully cancelled their own argument stacks, i.e. may or may not be incomplete.

6 Leftward dislocation captured

The above view on the variety of verb clustering patterns in terms of completeness of arguments also applies to fronting. It is a well-known and extensively studied phenomenon that long-distance dependencies cannot cross some borderlines. If a constituent should not contain a gap created by left-peripheral dislocation without containing the dislocated element itself, it is called a (strong) island. To the extent that islands are arguments selected by other categories, we can mark arguments for islandhood. For instance, noun phrases in Dutch generally will be marked as islands. An np argument, then, can only be cancelled against a category with an np head, the left-hand argument stack of which is empty. On the other hand, the sentential complement of verbs like *zeggen* ('to say') may contain a (to be precise: one) free dislocated argument. This is also marked at the relevant argument in the lexical category of *zeggen*. Moreover, this argument will not accept cancellation against a secondary category whose left hand stack contains anything other than at most one argument marked for dislocation (mode: $^\wedge d$). Here are some relevant lexical categories and the corresponding instances of extended generalized composition.

6.1
met ('with'): $n\backslash[n]/[np^\wedge isl]$
$prim\backslash PL/[sec^\wedge isl|PR] \quad sec\backslash[\,]/SR] \quad \Rightarrow \quad prim\backslash PL/(SR+PR)$

6.2
zeggen ('to say'): $vp\backslash[\,]/[s^\wedge ni]$
$prim\backslash PL/[sec^\wedge ni|PR] \quad sec\backslash SL/SR \quad \Rightarrow \quad prim\backslash(PL+SL)/(SR+PR)$

iff SL = [ ] or SL = [$x^\wedge d$] for some x

The combinatorics of long-distance dependencies are steered, again, by conditions on the state of argument stacks, imposed by lexical markings on arguments.

7 Unifying Verb Clustering and Leftward dislocation

The exposition above shows that extended generalized composition (3.7) offers, within the limits of a context-sensitive grammar, the instruments needed to describe the fine structure of both long distance dependencies and Dutch verb clustering. For this purpose, extended generalized composition makes the following devices available:

- two controlled forms of disharmonious composition
- two-valued parameters as for the emptiness of local argument stacks at merging time
- two-valued parameters as for the affectedness of local argument stacks at merging time.

In order to assure correct handling of the relevant configurations, these devices need only be effective in a strictly local fashion, to wit at the composition of two adjacent categories.
DELILAH parses the resulting grammar deterministically in a shift-reduce rhythm. From an instrumental point of view, then, there is no fundamental difference in the way verb clusters and long-distance dependencies are treated. The parametrization of extended generalized composition subsumes both phenomena.
(7.2) below holds - in a slightly impoverished and recoded format - one of the analyses DELILAH offers for the Dutch sentence

7.1
Wie zeg jij dat Henk de vrouw waarschijnlijk een pop had willen proberen te laten geven?
'Who say you that Henk the woman probably a puppet had want try to let give'
Whom do you think Henk had probably wanted to try to let give a puppet to the woman?

According to the analysis in (7.3), DELILAH derives, among others, the following semantic representation to this sentence, by abstraction, storage and conversion. Other representations exhibit alternative scopes.

7.2
?X.*person*(X) & ∃ Y.*puppet*(Y) & Pres.*say*(y, Prob.Past.Perf. *want*(h, *try*(h, *cause*(h, ι Z.*woman*(Z) & *give*(X, Y, Z)))))

This representation reflects the analysis that the dislocated wh-element binds the object of *laten* or, equivalently, the subject of *geven*. The sentence also illustrates obligatory verb clustering, optional clustering and *ipp*. In the analysis, top-down indentation marks composition; every composition is binary; equal levels of indentation mark composition to the less indented category immediately above; the relevant

argument is in italics. At crucial compositions the relevant instance of extended generalized composition is referred to. Every occurrence of the (sub)stack which contains the gap-argument is underlined. *NP* arguments are indexed afterwards, for transparency. Note that the dislocated wh-element *wie* is lexically equipped with a double category: one for the category it binds and the other a general operator on sentences. Note furthermore that the category for the adjunct *waarschijnlijk* is made transparent both for raising effects and for fronting by the same modality that accounts for the clustering properties of the verb proberen *proberen*.

7.3

(/o) **q**\ []/[] wie zeg jij dat henk de vrouw waarschijnlijk een pop had willen proberen te laten geven
(.) **q**\[]/[*s*^*o*] wie
(.) **s**\[]/[] wie zeg jij dat henk de vrouw waarschijnlijk een pop had willen proberen te laten geven
(.) **np**\[]/[] wie
(/d) **s**\[*np*^*d*]/[] zeg jij dat henk de vrouw waarschijnlijk een pop had willen proberen te laten geven
(/ni) **s**\[]/[*s_vn*^*ni*] zeg jij dat
(.) **s**\[]/[*s_sub*^*ni*] zeg jij
(.) **s**\[]/[*np*$_1$^*isl* **s_sub**^**ni**] zeg
(.) **np**\[]/[] jij
(.) **s_sub**\[]/[**s_vn**^**ni**] dat
(/isl) **s_vn**\[*np*^*d*]/[] henk de vrouw waarschijnlijk een pop had willen proberen te laten geven
(.) **np**\[]/[] henk
(.) **s_vn**\[*np*$_2$^*isl* *np*^*d*]/[] de vrouw waarschijnlijk een pop had willen proberen te laten geven
(.) **np**\[]/[] de vrouw
(/tc) **s_vn**\[*np*$_3$^*isl* **np**$_2$^**isl** *np*^*d*]/[] waarschijnlijk een pop had willen proberen te laten geven
(.) **s_vn**\[]/[*s_vn*^*tc*] waarschijnlijk
(.) **s_vn**\[**np**$_3$^**isl** **np**$_2$^**isl** *np*^*d*]/[] een pop had willen proberen te laten geven
(.) **np**\[]/[] een pop
(/ipp) **s_vn**\[**np**$_4$^**isl** **np**$_3$^**isl** **np**$_2$^**isl** *np*^*d*]/[] had willen proberen te laten geven
(.) **s_vn**\[**np**$_2$^**isl**]/[*vp*^*ipp*] had
(/vr) **vp**\[**np**$_4$^**isl** **np**$_3$^**isl** *np*^*d*]/[] willen proberen te laten geven
(.) **vp**\[]/[*vp*^*vr*] willen
(/tc) **vp**\[**np**$_4$^**isl** **np**$_3$^**isl** *np*^*d*]/[] proberen te laten geven
(.) **vp**\[]/[*vp_t*^*tc*] proberen
(/vr) **vp_t**\[**np**$_4$^**isl** **np**$_3$^**isl** *np*^*d*]/[] te laten geven
(.) **vp_t**\[]/[*vp*^*vr*] te
(/vr) **vp**\[**np**$_4$^**isl** **np**$_3$^**isl** *np*^*d*]/[] laten geven
(.) **vp**\[*np*^*d*]/[*vp*^*vr*] laten
(.) **vp**\[**np**$_4$^**isl** **np**$_3$^**isl**]/[] geven

7.4

$/_{\text{vr}}$: $p\backslash PL/[s\hat{}vr|PR] \quad s\backslash SL/SR \quad \Rightarrow \quad p\backslash PL/(SR+PR)$
iff PL and SL are unaffected

$/_{\text{tc}}$: $p\backslash PL/[s\hat{}tc|PR] \quad s\backslash SL/SR \quad \Rightarrow \quad p\backslash(SL+PL)/(SR+PR)$
iff PL is unaffected

$/_{\text{ipp}}$: $p\backslash PL/[s\hat{}ipp|PR] \quad s\backslash SL/SR \quad \Rightarrow \quad p\backslash(SL+PL)/(SR+PR)$
iff PL, SL unaffected, SR affected

$\backslash_{\text{d}}$: $s\backslash[]/[] \quad p\backslash[s\hat{}d]/[] \quad \Rightarrow \quad p\backslash[]/[]$

$/_{\text{o}}$: $p\backslash[]/[s\hat{}o] \quad s\backslash[]/[] \quad \Rightarrow \quad p\backslash[]/[]$

$/_{\text{isl}}$: $p\backslash PL/[s\hat{}isl|PR] \quad s\backslash[]/[] \quad \Rightarrow \quad p\backslash PL/PR$

8 Conclusion

In DELILAH's grammar, the construals for fronting and verb clustering are shaped by the same instruments: simple computable constraints on the modes of composition, of the sort stated in (5.1) to (5.4). These constraints have two components. One

component is the mode for appending argument lists. Choosing an option here is commendatory, and the option chosen reflects the way-of-being of the discontinuities involved. This component of the constraint is *post*-compositional. It is an output constraint. The other component relates to the status of the argument list *prior* to the composition; the constraints in this component can be viewed as input conditions. Leftward dislocation and all modes of verb clustering in Dutch are captured by small range variations in input and output constraints on linear Generalized Composition. Consequently, there is a point of view from which dislocation and clustering converge. This point of view is essentially derivational. It abstracts away from hypothetical reasoning in grammatical deduction.

References

[Ades and Steedman 1983] Ades, A.E. and M.J. Steedman, 'On the Order of Words', *Linguistics and Philosophy 4:4*, 1983, p. 517-558

[Barbiers 1998] Barbiers, S., 'English and Dutch as SOV-Languages and the Distribution of CP-Complements', in: R. van Bezooijen and R. Kager (eds), *Linguistics in the Netherlands 1998*, John Benjamins, 1998, p. 1 - 12

[Bar-Hillel *et al* 1960] Bar-Hillel, Y., C. Gaifman and E. Shamir, 'On categorial and phrase structure grammars', *The Bulletin of the Research Council of Israel 9F*, 1960, 16

[Buszkowski 1984] Buszkowski, W., 'Completeness Results for Lambek Syntactic Calculus', *Zeitschrift für mathematische Logik und Grundlagen der Mathematik' 32*, 1984, p. 13-28

[Buszkowski 1988] Buszkowski, W., 'Three theories of categorial grammar', in: W. Buszkowski *et. al.* (eds), *Categorial Grammar*, John Benjamins, 1988, p. 57 - 84

[Carpenter 1997] Carpenter, B., *Type-Logical Semantics*, MIT Press, 1997

[Cheng *t.a*] Cheng, L. L., 'Moving Just the Feature', in: U. Lutz, G. Müller and A. von Stechow(eds), *WH-Scope Marking*, John Benjamins, to appear

[Chomsky 1995] Chomsky, N., *The Minimalist Program*, MIT Press, 1995

[Chomsky 1998] Chomsky, N., *Minimalist Inquiries: The Framework*, MIT Occasional Papers in Linguistics 15, 1998

[Cremers 1993] Cremers, C., *On Parsing Coordination Categorially*, PhD diss, RU Leiden (HIL diss 5), 1993

[Den Dikken 1996] Den Dikken, M., 'The Minimal Links of Verb (Projection) Raising', in: W. Abraham *et al.* (eds), *Minimal Ideas. Syntactic studies in the minimalist framework*, John Benjamins, 1996

[Friedman *et al.* 1986] Friedman, J., D. Dai and W. Wang, 'The weak generative capacity of parenthesis-free categorial grammars', *Proceedings COLING 86*, Assoc for Comp Ling, 1986, p. 199-201

[Hepple 1990] Hepple, M, *The grammar and processing of order and dependency*, Ph. diss., University of Edinburgh, 1990

[Hepple 1996] Hepple, M., 'A Compilation-Chart Method for Linear Categorial Deduction', *Proceedings COLING 96*, Center for Sprokteknologi, Copenhagen, 1996, p. 537-542

[Hoeksema 1984] Hoeksema, J., *Categorial Morphology*, PhD diss, RU Groningen, 1984

[Houtman 1984] Houtman, J., 'Een kategoriale beschrijving van het Nederlands', *TABU 14:1*, 1984, p. 1-27

[Jacobson 1992] Jacobson, P. 'Comment on a paper by R.T. Oehrle', in: R. Levinson (ed), *Formal grammar: theory and implementation*, OUP, 1992, p. 129 - 167

[Joshi *et al.* 1991] Joshi, A.K., K. Vijay-Shanker and D. Weir, 'The Convergence of Mildly Context-Sensitive Grammar Formalisms', in: P. Sells, S.M. Shieber and T. Wasow (eds), *Foundational Issues in Natural Language Processing*, MIT Press, Cambridge, 1991, p. 31-82

[König 1990] König, E., *Der Lambek-Kalkül. Eine Logik für lexikalische Grammatiken*, Ph.diss University of Stuttgart, IBM IWBS Report 146, 1990

[Lambek 1958] Lambek, J., 'The Mathematics of Sentence Structure', *American Mathematical Monthly 65*, 1958, p. 154-170

[Mcgee Wood 1993] McGee Wood, M., *Categorial Grammars*, Routledge, London, 1993

[Montague 1973] Montague, R.M., 'The Proper treatment of Quantification in Ordinary English', in: J. Hintikkaa *et al. Approaches to Natural Language: Proceedings of the 1970 Workshop on Grammar and Semantics*, Reidel, 1973, p. 221-242

[Moortgat 1988] Moortgat, M., *Categorial Investigations*, Foris, Dordrecht, 1988

[Moortgat 1997] Moortgat, M., 'Categorial Type Logics', in: J. van Benthem and A. ter Meulen (eds), *Handbook of Logic and Language*, Elsevier, 1997

[Moortgat and Oehrle 1994] Moortgat, M. and R.T. Oehrle, 'Adjacency, dependency and order', in: P. Dekker and M. Stokhof (eds), *Proceedings Ninth Amsterdam Colloquium*, Univ of Amsterdam, ILLC, 1994, part II, p. 447-466

[Morrill 1994] Morrill, G., *Type Logical Grammar*, Kluwer, Dordrecht, 1994

[Pentus 1992] Pentus, M., 'Lambek grammars are context-free', ms. Yhubepecmet Mockbu, 1992

[Stabler 1996] Stabler, E., *Acquiring and parsing languages with movement*, ms UCLA, 1996

[Stabler 1999] Stabler, E., 'Remnant Movement and Structural Complexity', in: G. Bouma *et al.* (eds), *Constraints and Resources in Natural Language*, Syntax and Semantics,1999

[Steedman 1987] Steedman, M., 'Combinatory grammars and parasitic gaps', *Natural Language and Linguistic Theory 5*, 1987, p. 403-439

[Steedman 1990] Steedman, M., 'Gapping as constituent coordination', *Linguistics and Philosophy 13:2*, 1990, 207264

[Steedman 1996] Steedman, M., *Surface Structure and Interpretation*, MIT Press, 1996

[Steedman 2000] Steedman, M., *The Syntactic Process* , MIT Press, 2000

[Van Benthem 1991] Van Benthem, J., *Language in Action*, North Holland, Amsterdam, 1991

[Van de Woestijne 1999] Van de Woestijne, Chr., *A formal characterization of the Delilah system*, MSc thesis, Department of Information Sciences, Leiden University, 1999

[Vermaat 1999] Vermaat, W., *Controlling Movement. Minimalism in a deductive perspective*, CKI doctoraal scriptie, Utrecht University, 1999

[Vijay-Shanker and Weir 1993] Vijay-Shanker, K. and D.J. Weir, 'Parsing some Constrained grammar Formalisms', *Computational Linguistics 19:4*, 1993, p. 591 - 637

[Zinn 1993] Zinn, P., *Categoriale grammatica en de Chomsky hiërarchie*, MSc thesis, Universiteit Leiden, 1993

[Zwart 1993] Zwart, C.J.W., *Dutch Syntax*, PhD diss, RU Groningen, 1993

[Zwarts 1986] Zwarts, F., *Categoriale grammatica en algebraïsche semantiek*, PhD diss, RU Groningen, 1986

Received 31.03.00