FG 2007:
The 12th conference on Formal Grammar
Dublin, Ireland
August 4-5, 2007

Organizing Committee:

Laura Kallmeyer   Paola Monachesi
Gerald Penn   Giorgio Satta

CENTER FOR THE STUDY
OF LANGUAGE
AND INFORMATION

Contents

1

Multimodal categorial dependency grammars

Alexander Dikovsky

LINA CNRS FRE 2729, Université de Nantes

Abstract

Multimodal categorial dependency grammars introduced in this paper represent a uniform and efficient solution to the problem of compositional analysis of discontinuous dependencies (cross-serial dependencies included).

Ke ords Dependency Grammar, Categorial Grammar, Multimodal Architecture

1.1 Introduction

As far as it is a question of local domain syntactic relations, the analyses of categorial grammars (CG) and of dependency grammars (DG) are rather similar. The main differences are in the treatment of the determinants and of the adjuncts. The true challenge for both formalisms are the *discontinuous* syntactic relations and the *flexible word order*. The former need additional expressive means, the latter is fraught with explosion of types/rules system [1]. These grammars solve these problems in different ways. The logical type grammars extend the classical CG with the introduction rules making the choice of raisable types. This choice is unacceptable for the dependency syntax, in which the type of a word is uniquely determined by the dependencies through which it governs and is governed. To express discontinuous dependencies, the

[1] We don't consider this problem because of space reasons.

FG-2007.
Organizing Committee:, Laura Kallmeyer, Paola Monachesi, Gerald Penn, Giorgio Satta.
Copyright © 2007, CSLI Publications.

raisable types of logical type grammars are extended with modes [2] which determine the idiosyncratic compositionality of typed functors. In DG, where the order is naturally separated from dependency relations, a similar effect can be attained using stratification of dependency relations and of neighborhood topology and imposing constraints on both (cf. Duchier and Debusmann (2001), Duchier et al. (2004)). The two approaches are sufficiently expressive to explain the many complex syntactic phenomena, but lead to algorithmic intractability.

Recently introduced Generalized Categorial Dependency grammars (gCDG) (Dikovsky (2004), Dekhtyar and Dikovsky (2004)) define discontinuous dependencies using polarized valencies (left/right, positive/negative) and a simple valencies pairing principle **FA**: *"for every valency, the corresponding one is the closest dual valency in the indicated direction"*. This principle explains many well known discontinuous dependencies, but it does not apply to the cross-serial dependencies in Dutch. As simple as they are, the gCDG are very expressive: they generate all CF and various non-CF languages, in particular MIX. This means that probably, they are incomparable with the mildly CS languages. gCDG-languages form an AFL and have a polynomial time parsing algorithm. In this paper, we introduce into the gCDG a multimodal architecture considering valency pairing principles as modes and admitting different principles for individual discontinuous dependencies. In particular, we introduce for cross-serial dependencies a special valency pairing principle **FC**: *"first cross dual valency in the indicated direction"*. We prove that it is as efficient as **FA** and, in this way, obtain an efficient polynomial time parsing algorithm for the multimodal CDG under the two valency pairing principles.

1.2 Dependency types

Dependency type of a word (to be called *category*) represents its governor-subordinate valencies. The categories have the form α^P, where α is a *basic category* determining continuous (*projective*) dependencies and P is a *potential*, a string of *polarized valencies* determining discontinuous (*non-projective*) dependencies. The basic categories $\mathbf{B}(\mathbf{C})$ are FO types constructed from primitives $\mathbf{C}$: 1. $\mathbf{C} \subset \mathbf{B}(\mathbf{C})$. 2. If $\alpha \in \mathbf{C}$ and $\beta \in \mathbf{B}(\mathbf{C})$, then $[\alpha\backslash\beta]$, $[\alpha * \backslash\beta]$, $[\beta/\alpha*]$, $[\beta/\alpha] \in \mathbf{B}(\mathbf{C})$.$\Box$ The constructors $\backslash, /$ being associative, each basic category can be represented as $[a_{lm}\backslash ... \backslash a_{l1}\backslash f/a_{r1}/.../a_{rn}]$. Intuitively, f corresponds

[2] The multimodal architecture was proposed and developed by Oehrle, Morrill, Moortgat and Hepple (see Morrill (1994), Moortgat (1997)) and was applied to dependencies in an indirect way (see Moortgat and Morrill (1991), Kruijff (2001)).

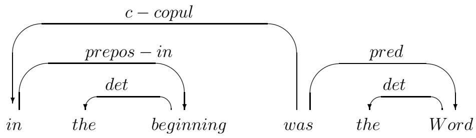

FIGURE 1 A projective DT

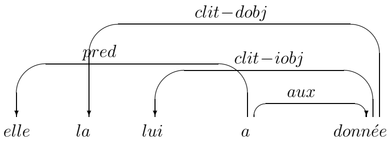

FIGURE 2 A non-projective DT (Fr.: **she it*$_{\text{FEM}}$ *to-him has given*)

to the incoming (governor) dependency and a_{li}, a_{rj} correspond to left and right outgoing (subordinate) dependencies. $d*$ corresponds to the iterated dependency d. A special type $S \in \mathbf{C}$ is reserved for sentences. E.g., the dependency tree (DT) in Fig. 1 gives the categories:

$in \mapsto [c{-}copul/prepos{-}in]$ $\quad beginning \mapsto [det\backslash prepos{-}in]$ $\quad the \mapsto [det]$
$was \mapsto [c{-}copul\backslash S/pred]$ $\quad Word \mapsto [det\backslash pred]$

The polarized valencies use four *polarities*: left and right positive $\nwarrow, \nearrow$ and left and right negative $\swarrow, \searrow$. For instance, $v = \nwarrow d$ requires a subordinate through dependency d situated *somewhere* on the left. The *dual* valency $\breve{v} = \swarrow d$ requires a governor through the same dependency d situated *somewhere* on the right. Together they describe the discontinuous dependency d. For negative valencies $\swarrow d, \searrow d$ are admitted special primitive *anchor* types $\#(\swarrow d)$, $\#(\searrow d)$. Elimination of such subtype creates no dependency. It only checks the adjacency of a distant subordinate to its host word. $gCat(\mathbf{C})$ will denote the set of all categories over $\mathbf{C}$. E.g., the DT in Fig. 2 is determined by the following categories (the clitics *la*, *lui* are anchored on the auxiliary *a*):

$elle \mapsto [pred]$ $\quad a \mapsto [\#(\swarrow clit{-}iobj)\backslash\#(\swarrow clit{-}dobj)\backslash pred\backslash S/aux]$
$la \mapsto [\#(\swarrow clit{-}dobj)]^{\swarrow clit{-}dobj}$ $\quad lui \mapsto [\#(\swarrow clit{-}iobj)]^{\swarrow clit{-}iobj}$
$donnée \mapsto [aux]^{\nwarrow clit{-}iobj \nwarrow clit{-}dobj}$

1.3 Generalized Categorial Dependency Grammar

A *Generalized Categorial Dependency grammar* (gCDG) $G = (W, \mathbf{C}, S, \delta)$, is defined by a function δ (its *lexicon*), assigning a finite set of categories in $gCAT(\mathbf{C})$ to each word in W. Correctness of category assignment $\Gamma \in \delta(x)$ to a string $x \in W^+$, denoted $\Gamma \vdash S$, is proved using the following grammar-non-specific **dependency calculus**[3].

[3]We show left-oriented rules. The right-oriented are symmetric.

$$\dfrac{\dfrac{\dfrac{[pred] \quad \dfrac{[\#^l(\alpha)]^\alpha \quad \dfrac{[\#^l(\beta)]^\beta[\#^l(\beta)\backslash\#^l(\alpha)\backslash pred\backslash S/aux]}{[\#^l(\alpha)\backslash pred\backslash S/aux]^\beta}\,(\mathbf{L}^l)}{[pred\backslash S/aux]^{\alpha\beta}}\,(\mathbf{L}^l)}{[S/aux]^{\alpha\beta}}\,(\mathbf{L^l}) \qquad [aux]^{\breve{\beta}\breve{\alpha}}}{[S]^{\alpha\beta\breve{\beta}\breve{\alpha}}}\,(\mathbf{L}^r)}{S}\,(\mathbf{D}^l\times 2)$$

$$(\alpha = \swarrow clit{-}dobj, \beta = \swarrow clit{-}iobj, \breve{\alpha} = \nwarrow clit{-}dobj, \breve{\beta} = \nwarrow clit{-}iobj)$$

FIGURE 3 A proof for the dependency tree in Fig. 2

$\mathbf{L^l}$. $C^{P_1}[C\backslash\beta]^{P_2} \vdash [\beta]^{P_1P_2}$
$\mathbf{I^l}$. $C^{P_1}[C^*\backslash\beta]^{P_2} \vdash [C^*\backslash\beta]^{P_1P_2}$
$\mathbf{\Omega^l}$. $[C^*\backslash\beta]^{P} \vdash [\beta]^{P}$
$\mathbf{D^l}$. $\alpha^{P_1(\swarrow C)P(\nwarrow C)P_2} \vdash \alpha^{P_1PP_2}$, if $(\swarrow C)P(\nwarrow C)$ satisfies the pairing principle
FA: P has no occurrences of $\swarrow C, \nwarrow C$.

$\mathbf{L^l}$ is the classical elimination rule. Eliminating the argument subtype $C \neq \#(\alpha)$ it constructs the (projective) dependency C and concatenates the potentials. $C = \#(\alpha)$ creates no dependency. $\mathbf{I^l}$ derives $k > 0$ instances of C. $\mathbf{\Omega^l}$ serves for the case $k = 0$. $\mathbf{D^l}$ derives discontinuous dependencies. It pairs and eliminates dual valencies satisfying the rule **FA** and creates the discontinuous dependency C.

For a dependency structure (DS) D and a string x, let $G(D, x)$ denote the relation: D is the graph constructed in a proof $\Gamma \vdash S$ for some $\Gamma \in \delta(x)$. Then the *language* (*DS language*) generated by G is the set $L(G)=_{df} \{w \mid \exists D\ G(D, w)\}$ ($\Delta(G)=_{df} \{D \mid \exists w\ G(D, w)\}$).

For instance, the dependency tree (DT) in Fig. 2 is constructed from the proof in the dependency calculus, shown in Fig. 3.

gCDG are very expressive (see the diagram in Fig. 4). Their subclass, the gCDG with bounded valency deficit[4], generate the CF languages. gCDG can also generate non-CF languages, e.g., the languages $\{a^nb^nc^n \mid n > 0\}$ and $L^{(m)} = \{a_1^na_2^n...a_m^n \parallel n \geq 1\}$ over $W = \{a_1, \dots, a_m\}$, $m \geq 2$ (Dikovsky (2004)). The languages $L^{(m)}$ are mildly CS and cannot be generated by basic TAGs starting from $m > 4$ (see Joshi et al. (1991)). Moreover, there is a gCDG generating the language MIX consisting of all permutations of the strings $a^nb^nc^n, n > 0 : MIX = \{w \in \{a, b, c\}^+ \parallel |w|_a = |w|_b = |w|_c\}$ (Béchet et al. (2005)). E. Bach conjectures that MIX is not a mildly CS

[4]The *valency deficit* $\sigma(n)$ is the maximal size of a potential used in the proofs of category assignments to strings of length n. $\mathcal{L}^{\sigma<c}(gCDG)$ is the class of languages generated by gCDG with bounded valency deficit. $\mathcal{L}^{\sigma<c}(gCDG) = \mathcal{L}(CF)$ (Dikovsky (2004)).

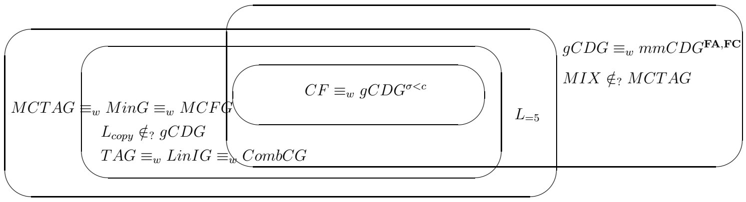

FIGURE 4 Comparison with other families

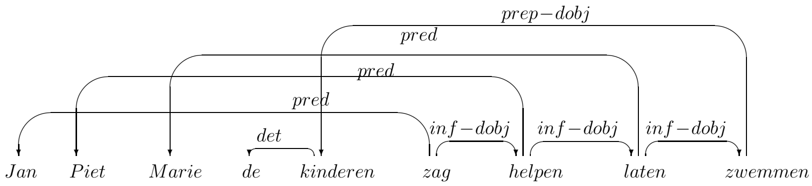

FIGURE 5 Dutch: * *Wim Jan Marie the children saw help learn swim*

language. Recently, it was shown that the family of gCDG languages $\mathcal{L}(gCDG)$ is an AFL (Dekhtyar and Dikovsky (2007)). So gCDG represent an interesting alternative to the mildly CS grammars. On the other hand, it is conjectured in (Dikovsky (2004), Dekhtyar and Dikovsky (2004)) that the copy language $L_{copy} = \{ww \parallel w \in \{a,b\}^+\}$, which is basic TAG, cannot be generated by a gCDG.

1.4 Multimodal architecture for gCDG

A new perspective on cross-serial dependencies in Dutch. L_{copy} is interesting because for a long time there is an opinion that it serves as an abstract model of the so called *unlimited cross-serial dependencies* in Dutch. This Dutch construction represents the sentences $n_1 n_2 \ldots n_m n_{m+1} v_1 v_{(inf)2} \ldots v_{(inf)m}$, where there is a predicative dependency $n_1 \overset{pred}{\longleftarrow} v_1$ between the finite verb v_1 and the noun n_1, predicative dependencies $n_i \overset{pred}{\longleftarrow} v_{(inf)i}$ between the verbs $v_{(inf)i}$ in the infinitive and the nouns n_i $(2 \le i \le m)$ and eventually also the dependency $n_{m+1} \overset{prep-dobj}{\longleftarrow} v_{(inf)m}$ if the verb $v_{(inf)m}$ is transitive and n_{m+1} is present (i.e., $n_{m+1} \neq \varepsilon$). In Fig. 5, we show an example of this construction, borrowed from (Bresnan et al. (1982)). So indeed, the polarized valencies (predicative in the example) should be paired not with the closest, but with the *most distant* dual valencies. Meanwhile, a more detailed analysis of this construction (Pulman and Ritchie (1985)) shows that an agreement exists only between the finite verb v_1 and its subject n_1 and the form of the object

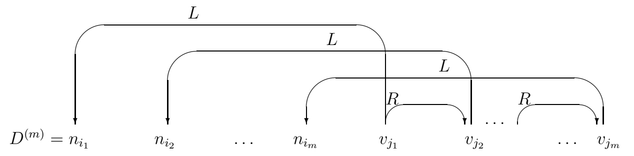

FIGURE 6 Dependency tree $D^{(m)}$

n_{m+1} (when it is present) is determined by the transitive verb $v_{(inf)m}$. From this follows that the true model of this construction is not L_{copy}, but the DS language $\Delta_{cross} = \{D^{(m)} \parallel m > 0\}$ over $W = N \cup V$, where $D^{(m)}$ is the DS shown in Fig.6 and $n_{i_l} \in N, v_{j_r} \in V$. The corresponding language is linear CF. gCDG cannot generate Δ_{cross}. So they are not adapted for Dutch.

Basic notions and facts. Our solution to the problem of cross-serial dependencies rests upon a property of independence of basic categories and polarized valencies in the proofs in the dependency calculus (Dekhtyar and Dikovsky (2004, 2007)). This property is expressed in terms of *projections* of categories and *"well-bracketing"* criteria for potentials.

Local and *valency* projections $\|\gamma\|_l$, $\|\gamma\|_v$ are defined as follows:
1. $\|\varepsilon\|_l = \|\varepsilon\|_v = \varepsilon$; $\|\alpha\gamma\|_l = \|\alpha\|_l\|\gamma\|_l$ and $\|\alpha\gamma\|_v = \|\alpha\|_v\|\gamma\|_v$ for $\alpha \in gCat(\mathbf{C})$ and $\gamma \in gCat(\mathbf{C})^*$.
2. $\|C^P\|_l = C$ and $\|C^P\|_v = P$ for $C^P \in gCAT(\mathbf{C})$.

To speak about "well-bracketing" of potentials, we interpret $\swarrow d$ and $\nearrow d$ as *left brackets* and $\nwarrow d$ and $\searrow d$ as *right brackets.* The sets of all left and right bracket valencies are denoted $V^l(\mathbf{C})$ and $V^r(\mathbf{C})$. $V(\mathbf{C}) =_{df} V^l(\mathbf{C}) \cup V^r(\mathbf{C})$. For a dependency d and a potential P, let $P \restriction d$ be the result of deleting the occurrences of all valencies but $\swarrow d$, $\nearrow d$, $\nwarrow d$ and $\searrow d$. Then P is *balanced* if $P \restriction d$ is well bracketed in the usual sense for every d. This property can be incrementally checked using the following values defined for every potential P and all valencies $\alpha \in V^l(\mathbf{C})$ and $\breve{\alpha} \in V^r(\mathbf{C})$:

$$\Delta_\alpha(P) = max\{|P'|_\alpha - |P'|_{\breve{\alpha}} \parallel P' \textit{ is a } \textsf{suffix} \textit{ of } P\},$$
$$\Delta_{\breve{\alpha}}(P) = max\{|P'|_{\breve{\alpha}} - |P'|_\alpha \parallel P' \textit{ is a } \textsf{prefix} \textit{ of } P\}$$

($|\gamma|_\alpha$ is the number of occurrences of α in γ). They express the deficits of right and left $\alpha-$brackets in P (i.e. the maximal number of right and left bracket α-valencies which need to be added to P on the right (left) so that it became balanced. It is not difficult to prove that a potential P is balanced iff $\sum\limits_{\alpha \in V(\mathbf{C})} \Delta_\alpha(P) = 0$ and that

$$\Delta_\alpha(P_1P_2) = \Delta_\alpha(P_2) + max\{\Delta_\alpha(P_1) - \Delta_{\breve{\alpha}}(P_2), 0\},$$
$$\Delta_{\breve{\alpha}}(P_1P_2) = \Delta_{\breve{\alpha}}(P_1) + max\{\Delta_{\breve{\alpha}}(P_2) - \Delta_\alpha(P_1), 0\}$$

for all potentials P_1, P_2, and every $\alpha \in V^l(\mathbf{C})$, $\breve{\alpha} \in V^r(\mathbf{C})$.

Let **c** be the projective core of the dependency calculus, consisting of the rules **L**, **I** and **Ω** and $\vdash_{\mathbf{c}}$ be the provability relation in this sub-calculus. One of the key results of (Dekhtyar and Dikovsky (2004, 2007)) is the following property of *projections independence* providing for gCDG a polynomial time parsing algorithm.

Theorem 1 *For a gCDG G with dictionary δ and for a string x, $x \in L(G)$ iff there is $\Gamma \in \delta(x)$ such that $\|\Gamma\|_l \vdash_{\mathbf{c}} S$ and $\|\Gamma\|_v$ is balanced.*

The dependency calculus of gCDG uses the rules **D** with the valency pairing principle **FA**. The only point in the proof of this theorem intrinsically related to this pairing principle is the following proposition evidently true for **FA**.

Lemma 1 *A potential P is balanced iff for every category α^P there is a proof $\alpha^P \vdash \alpha$ using only the rules $\mathbf{D}^l$ and $\mathbf{D}^r$.*

Multimodal CDG. Now, turning back to our solution to the problem of cross-serial dependencies, we can say more precisely that it consists in assigning to each polarized valency α its *mode*, i.e. the particular valency pairing principle $\mathbf{M}_\alpha$ for α, and respectively extending the dependency calculus to new discontinuous dependency rules $\mathbf{D}^l_{\mathbf{M}_\alpha}$ and $\mathbf{D}^r_{\mathbf{M}_\alpha}$. This needs a revision of the gCDG definition.

Definition 1 *$G = (W, \mathbf{C}, S, \delta, \mu)$ is a* multimodal CDG *(mmCDG) if $(W, \mathbf{C}, S, \delta)$ is a gCDG and μ is a function assigning to each polarized valency α a pairing principle $\mathbf{M}_\alpha$. There are rules $\mathbf{D}^l_{\mathbf{M}_\alpha}$ and $\mathbf{D}^r_{\mathbf{M}_\alpha}$ in the multimodal dependency calculus $\vdash_\mu$ for all α used in μ. The language (DS language) generated by G using a set of modes M is denoted $L^M(G)$ $(\Delta^M(G))$. mmCDG^M is the family of all such mmCDG.*

In particular, the new pairing principle (mode) $\mathbf{FC}^l$ defining the left cross dependencies is as follows:

$$\mathbf{D}_{\mathbf{FC}^l}.\quad \alpha^{P_1(\swarrow C)P(\nwarrow C)P_2} \vdash \alpha^{P_1 P P_2},$$

if $P_1(\swarrow C)P(\nwarrow C)$ satisfies the pairing principle
$\mathbf{FC}^l$: P_1 has no occurrences of $\swarrow C$ and P has no occurrences of $\nwarrow C$.
It is not difficult to show

Proposition 1 $\Delta_{cross} = \Delta^{\mathbf{FC}^l}(G_{cross})$, *where G_{cross} is the mmCDG:*
$n \mapsto \#(L)^{\swarrow L}, [\#(L)\backslash\#(L)]^{\swarrow L}$, $v \mapsto [\#(L)\backslash S/R]^{\nwarrow L}, [R/R]^{\nwarrow L}, R^{\nwarrow L}$,
for $n \in N$ and $v \in V$.

Example 1 *For instance, G_{cross} generates the DS $D^{(3)}$ in Fig. 6 if*
$n_{i_1} n_{i_2} n_{i_3} v_{j_1} v_{j_2} v_{j_3} \mapsto \#(L)^{\swarrow L}[\#(L)\backslash\#(L)]^{\swarrow L}[\#(L)\backslash\#(L)]^{\swarrow L}[\#(L)\backslash S/R]^{\nwarrow L}[R/R]^{\nwarrow L}R^{\nwarrow L}$.
A proof of $D^{(3)} \in \Delta_{cross}$ is shown in Fig. 7.

Extending the dependency calculus to a new rule $\mathbf{D}_{\mathbf{M}}$ for a pairing principle **M**, we should also extend Lemma 1 to this rule in order to

$$\dfrac{\dfrac{\dfrac{\dfrac{[\#(L)]^{\swarrow L}[\#(L)\backslash\#(L)]^{\swarrow L}}{[\#(L)]^{\swarrow L\swarrow L}}\,(\mathbf{L}^l)\quad [\#(L)\backslash\#(L)]^{\swarrow L}}{[\#(L)]^{\swarrow L\swarrow L\swarrow L}}\,(\mathbf{L}^l)\quad \dfrac{[\#(L)\backslash S/R]^{\nwarrow L}\quad \dfrac{[R/R]^{\nwarrow L}[R]^{\nwarrow L}}{[R]^{\nwarrow L\nwarrow L}}\,(\mathbf{L}^r)}{[\#(L)\backslash S]^{\nwarrow L\nwarrow L\nwarrow L}}\,(\mathbf{L}^r)}{[S]^{\swarrow L\swarrow L\swarrow L\nwarrow L\nwarrow L\nwarrow L}}\,(\mathbf{L}^l)}{[S]}\,(\mathbf{D}^l_{\mathbf{FC}}\times 3)$$

FIGURE 7 A proof of $D^{(3)}$

guarantee the projections' independence and so an efficient parsing. Let us do it for the principle $\mathbf{FC}^l$.

Lemma 2 *A potential P is balanced iff for every category α^P there is a proof $\alpha^P \vdash_{\mathbf{FC}^l} \alpha$ using only the rule $\mathbf{D}_{\mathbf{FC}^l}$.*

Proof scheme. Whatever is a pairing principle $\mathbf{M}$, the corresponding rule $\mathbf{D_M}$ concerns only one pair of dual valencies $\swarrow d$, $\nwarrow d$, for which $\mu(\swarrow d) = \mathbf{M}$. So, without loss of generality, we can presume that $P = P \upharpoonright d$. This implies

Statement 1. *P is balanced iff $|P'|_{\swarrow d} \geq |P'|_{\nwarrow d}$ for every prefix P' of P and $|P|_{\swarrow d} = |P|_{\nwarrow d}$.*

Lemma 2 follows from the following

Statement 2. *Let $\alpha^{P_1 \swarrow d P_2 \nwarrow d P_3} \vdash_{\mathbf{FC}^l} \alpha^{P_1 P_2 P_3}$. Then*

$P = P_1 \swarrow d P_2 \nwarrow d P_3$ is balanced iff $P' = P_1 P_2 P_3$ is so.

This statement is proved by induction on $n = |P|_{\swarrow d}$.

1. For $n = 1$ it is trivial ($P = \swarrow d \nwarrow d$).

2. $|P|_{\swarrow d} = n + 1$. Then $P_1 = \varepsilon$, $P_2 = (\swarrow d)^m$ for some $m \geq 0$ and $P = \swarrow d(\swarrow d)^m \nwarrow d P_3$. By Statement 1, P is balanced iff $P' = (\swarrow d)^m P_3$ is so. Hence, Statement 2 follows from the fact $|P'|_{\swarrow d} = n$. $\square$

Corollary 1 *For a mmCDG $G = (W, \mathbf{C}, S, \delta, \mu)$ with modes $M = \{\mathbf{FA}, \mathbf{FC}^l\}$ and for $x \in W^+$, $x \in L(G)$ iff there is $\Gamma \in \delta(x)$ such that $\|\Gamma\|_l \vdash_{\mathbf{c}} S$ and $\|\Gamma\|_v$ is balanced.*

Corollary 2 $\mathcal{L}(gCDG) = \mathcal{L}(mmCDG^{\mathbf{FC}}) = \mathcal{L}(mmCDG^{\mathbf{FA},\mathbf{FC}})$.

1.5 Parsing Complexity

In (Dekhtyar and Dikovsky (2004)) was described a polynomial time parsing algorithm for a subclass of gCDG [5]. Corollary 1 lets extend it to $mmCDG^{\mathbf{FA},\mathbf{FC}^l}$. Below, we present the extended algorithm.

Preliminaries. Let us fix for the rest a mmCDG $G = (W, \mathbf{C}, S, \delta, \mu)$ with left polarized valencies $V^l(\mathbf{C}) = \{v_1, \ldots, v_p\}$ and dual right valencies $V^r(\mathbf{C}) = \{\breve{v}_1, \ldots, \breve{v}_p\}$. We start with two failure functions used in the algorithm. Let $w = w_1 w_2 ... w_n \in W^+$, $\alpha \in V^l(\mathbf{C})$ and $1 \leq i \leq n$. Then

$$\pi^L(\alpha, i) = max\{\Delta_\alpha(\|\Gamma\|_v) \mid \Gamma \in \delta(w_1 ... w_i)\}$$

[5] It is implemented in Lisp by D. and H. Todorov and in $C\#$ by I. Zaytsev.

is the *left failure function*. Similarly,

$$\pi^R(\alpha,i) = max\{\Delta_\alpha(\|\Gamma\|_v) \parallel \Gamma \in \delta(w_{n-i+1}...w_n)\}$$

is the *right failure function* for $\alpha \in V^r(\mathbf{C})$. We set $\pi^L(\alpha,0) = \pi^R(\alpha,0) = 0$. It is not difficult to prove the following properties of these functions.

Lemma 3 *(i) Let* $1 \le i \le n-1$. *Then*
$\pi^L(\alpha, i+1) =$
$max\{\Delta_\alpha(P) + max\{\pi^L(\alpha,i) - \Delta_{\breve{\alpha}}(P), 0\} \parallel P = \|\gamma\|_v, \gamma \in \delta(w_{i+1})\}$,
$\pi^R(\alpha, i+1) =$
$max\{\Delta_{\breve{\alpha}}(P) + max\{\pi^R(\alpha,i) - \Delta_\alpha(P), 0\} \parallel P = \|\gamma\|_v, \gamma \in \delta(w_{n-i+1})\}$.
(ii) If $\Gamma \vdash S$ *for some* $\Gamma = \gamma_1...\gamma_n \in \delta(w)$, *then*

$$\Delta_\alpha(\|\gamma_i...\gamma_j\|_v) \le \pi^R(\breve{\alpha}, n-j),\ \Delta_{\breve{\alpha}}(\|\gamma_i...\gamma_j\|_v) \le \pi^L(\alpha, i-1)$$

for all $1 \le i \le j \le n$, $\alpha \in V^l(\mathbf{C})$, $\breve{\alpha} \in V^r(\mathbf{C})$.

mmCdgParser applies to the mmCDG G and to a string $w = w_1w_2...w_n \in W^+$ and fills up a $n \times n$ triangle matrix M with *items*. Each cell $M[i,j]$, $i \le j$, corresponds to the string interval $w_i...w_j$ and contains a finite set of items. Each item codes a category C^P and has the form $\langle C, \Delta^L, \Delta^R, I^l, I^r\rangle$, where: C is a basic category $C \in \mathbf{B}(\mathbf{C})$, $\Delta^L = (\Delta_{v_1}, \ldots, \Delta_{v_p})$ and $\Delta^R = (\Delta_{\breve{v}_1}, \ldots, \Delta_{\breve{v}_p})$ are integer vectors whose component i contains the corresponding deficits of right (left) non-paired v-brackets in the potential P, I^l, I^r are left and right angle items from which I is calculated (for I in diagonal $M[i,i]$, $I^l = I^r = \emptyset$).

Correctness. Correctness of CalcFailFuncL() and CalcFailFuncR() follows from Lemma 3.

Theorem 2 *Let* $G = (W, \mathbf{C}, S, \delta, \mu)$ *be a mmCDG and* $w = w_1w_2...w_n \in W^+$. *Then for any* $1 \le i \le k \le j \le n$, *an item* $I = \langle\theta, \Delta^L, \Delta^R, I^l, I^r\rangle$ *falls to* $M[i,j]$ *iff there is* $\Gamma = \Gamma_1\gamma_i \ldots \gamma_j\Gamma_2 \in \delta(w)$ *such that* $\gamma_i \ldots \gamma_j \in \delta(w_i \ldots w_j)$ *and*
(i) $\gamma_i \ldots \gamma_j \vdash \theta$,
(ii) $\Delta^L[\alpha] = \Delta_\alpha(\|\gamma_i \ldots \gamma_j\|_v)$, $\Delta^R[\alpha] = \Delta_{\breve{\alpha}}(\|\gamma_i \ldots \gamma_j\|_v)$
for all $\alpha \in V^l(\mathbf{C}), \breve{\alpha} \in V^r(\mathbf{C})$,
(iii) $\gamma_i \ldots \gamma_j$ *satisfies point (ii) of Lemma 3.*

Complexity. For a mmCDG $G = (W, \mathbf{C}, S, \delta, \mu)$, let $l_G = |\delta|$ be the number of category assignments in the lexicon, $a_G = max\{k \parallel \exists x \in W\ ([\alpha_k\backslash...\backslash\alpha_1\backslash C/\beta]^P \in \delta(x)\ \vee\ [\beta\backslash C/\alpha_1/.../\alpha_k]^P \in \delta(x))\}$ be the maximal number of argument subtypes in assigned categories, $p_G = |V^l(\mathbf{C})| = |V^r(\mathbf{C})|$ be the number of polarized valencies and $\Delta_G = max\{\Delta_\alpha(P) \parallel \exists x \in W(C^P \in \delta(x)\ \vee\ \alpha \in V(\mathbf{C}))\}$ be the maximal valency deficit in assigned categories. In the complexity bound below, n will denote the length of the input string $n = |w|$.

Parsing algorithm mmCdgParser.

```
Algorithm mmCdgParser
//Input: mmCDG G, string w = w_1...w_n
//Output: ⟨"yes", DSD⟩ iff w ∈ L(G)
{
  CalcFailFuncL();
  CalcFailFuncR();
  for (k = 1,...,n)
  {
      Propose( k )
  }
  for (l = 2,...,n)
  {
      for (i = 1,...,n − l)
      {
          j := i + l − 1;
          for (k = i,...,j − l)
          {
              SubordinateL(i, k, j);
              SubordinateR(i, k, j);
          }
      }
  }
  if (I = ⟨S, (0, 0,...,0), (0, 0,...,0), I^l, I^r⟩ ∈ M[1, n])
      return ⟨"yes", Expand(I)⟩;
  //procedure Expand( I ) calculates the output DS
  //It is the only one sensitive to the valency pairing
  //principles FA,FC. It is tedious and is not shown here
  else
      return ⟨"no", ∅⟩;
}
```

```
//For 1 ≤ i ≤ n
Propose( i )
{
  (loop) foreach (C^P ∈ δ(w_i))
  {
      foreach (v ∈ V^l(C))
      {
          Δ^L[v] := Δ_v(P);
          if (Δ^L[v] > π^R[v̆, n − j]) next (loop);
          Δ^R[v̆] := Δ_v̆(P);
          if (Δ^R[v̆] > π^L[v, i − 1]) next (loop);
      }
      AddItem( M[i, i], ⟨C, Δ^L, Δ^R, ∅, ∅⟩ );
  }
}

AddItem( M[i, j], ⟨C, Δ^L, Δ^R, I^l, I^r⟩ )
{
  M[i, j] := M[i, j] ∪ {⟨C, Δ^L, Δ^R, I^l, I^r⟩};
  if (C = [C' ∗ \β])
  {
      AddItem( M[i, j], ⟨[β], Δ^L, Δ^R, I^l, I^r⟩ );
  }
  if (C = [β/C'∗])
  {
      AddItem( M[i, j], ⟨[β], Δ^L, Δ^R, I^l, I^r⟩ );
  }
}
```

```
CalcFailFuncL()
{
  foreach (v ∈ V^l(C))
  {
      π^L[v, 0] := 0;
      for (i = 1,...,n)
      {
          π_max := 0;
          foreach (C^P ∈ δ(w_i))
          {
              π_max := max{π_max, Δ_v(P)+
                  max{π^L[v, i − 1] − Δ_v̆(P), 0}};
          }
          π^L[v, i] := π_max;
      }
  }
}
```

CalcFailFuncR() is similar.

```
//For 1 ≤ i ≤ k ≤ j ≤ n
SubordinateL( i, k, j )
{
  (loop)
  foreach (I_1 = ⟨α_1, Δ^L_1, Δ^R_1, I^l_1, I^r_1⟩ ∈ M[i, k],
           I_2 = ⟨α_2, Δ^L_2, Δ^R_2, I^l_2, I^r_2⟩ ∈ M[k + 1, j])
  {
      foreach (v ∈ V^l(C))
      {
          Δ^L[v] := Δ^L_2(v) + max{Δ^L_1(v) − Δ^R_2(v), 0};
          if (Δ^L[v] > π^R[v̆, n − j]) next (loop);
          Δ^R[v̆] := Δ^R_1(v̆) + max{Δ^R_2(v̆) − Δ^L_1(v̆), 0};
          if (Δ^R[v̆] > π^L[v, i − 1]) next (loop);
      }
      if ( α_1 = C and α_2 = [C\β] )
      {
          AddItem( M[i, j], ⟨[β], Δ^L, Δ^R, I_1, I_2⟩ );
      }
      elseif ( (α_1 = C and α_2 = [C ∗ \β]) or α_1 = [ε] )
      {
          AddItem( M[i, j], ⟨α_2, Δ^L, Δ^R, I_1, I_2⟩ );
      }
  }
}
```

SubordinateR(i, k, j) is similar.

Theorem 3 *Algorithm* **mmCdgParser** *has time complexity*

$$\mathbf{O}(\mathsf{l_G} \cdot \mathsf{a}^2_\mathsf{G} \cdot (\Delta_\mathsf{G} \cdot \mathsf{n})^{2\mathsf{p_G}} \cdot \mathsf{n}^3).$$

Proof. A category $\gamma \in \delta(x)$ may be reduced to no more than a_G^2 different categories. So the maximal number of matrix cell elements is $l_G \cdot a_G^2$. The valency deficits are bounded by the maximal value of the failure functions. So the maximal deficit of a polarized valency is $\Delta_G \cdot n$. Therefore, the number of different valency deficit vectors is bounded by $(\Delta_G \cdot n)^{2p_G}$. Filling one matrix cell needs visiting n cells. There are $\frac{n^2}{2}$ cells in M. This proves the time bound. □

Remark 1 *1. For a fixed grammar G, the values l_G, a_G, p_G and Δ_G are constant. If G may vary, the membership problem becomes NP-complete (Dekhtyar and Dikovsky (2004)).*
2. When no polarized valencies, the parsing time is evidently $\mathbf{O}(\mathsf{n}^3)$.
3. Every mmCDG G with bounded valency deficit $\sigma < c$ can be translated into an equivalent mmCDG G_c without polarized valencies (so with parsing time $\mathbf{O}(\mathsf{n}^3)$*). Of course, the size of G_c is exponential:* $|G_c| = \mathbf{O}(|\mathsf{G}| \cdot \mathsf{c}^{\mathsf{p_G}})$. *However, if the number of discontinuous dependencies in sentences is uniformly bounded by a constant, then the parsing time is* $\mathbf{O}(\mathsf{n}^3)$ *without exponential explosion of the grammar size.*

1.6 Concluding remarks

The categorial dependency grammars express discontinuous dependencies by joining the paradigme of logical type grammars with the cognitivistic paradigme. From the first paradigme they inherit the core first order types and constructor elimination rules. From the second one they introduce valency pairing principles modeling fast dynamic memory structures. E.g., the pairing principle **FA** corresponds to the hypothesis that the discontinous dependencies are controlled using non-communicating stacks memory: one stack per valency type. The principle **FC** corresponds to another kind of fast memory: the queue. Due to this join, the linguistic definitions of dependency types in the lexicon become completely local. The projective dependencies of a word are those of its local domain as a governor. A discontinuous dependency is defined in two local domains: that of the governor assigns it the positive valency and that of the subordinate assigns it the negative valency. The dependency itself results from pairing the two valencies. Another important consequence is that our multimodal extension of CDG expands their strong expressive power without changing the complexity and the weak expressive power.

Acknowledgements

We are grateful to R. Moot for his helpful comments on the cross-serial dependencies in Dutch.

References

Béchet, Denis, Alexander Dikovsky, and Annie Foret. 2005. Dependency structure grammars. In P. Blache and E. Stabler, eds., *Proc. of the 5th Int. Conf. "Logical Aspects of Computational Linguistics" (LACL'2005)*, LNAI 3492, pages 18–34.

Bresnan, J., R.M. Kaplan, S. Peters, and A. Zaenen. 1982. Cross-serial dependencies in dutch. *Linguistic Inquiry* 13(4):613–635.

Dekhtyar, Michael and Alexander Dikovsky. 2004. Categorial dependency grammars. In M. Moortgat and V. Prince, eds., *Proc. of Intern. Conf. on Categorial Grammars*, pages 76–91. Montpellier.

Dekhtyar, Michael and Alexander Dikovsky. 2007. Generalized categorial dependency grammars. Forthcoming in LNCS. Available from http://www.sciences.univ-nantes.fr/info/perso/permanents/dikovsky/.

Dikovsky, Alexander. 2004. Dependencies as categories. In *"Recent Advances in Dependency Grammars". COLING'04 Workshop*, pages 90–97.

Duchier, D. and R. Debusmann. 2001. Topological dependency trees : A constraint-based account of linear precedence. In *Proc. of the Intern. Conf. ACL'2001*, pages 180–187. Toulouse, France: ACL & Morgan Kaufman.

Duchier, D., R. Debusmann, and G.-J. M. Kruijff. 2004. Extensible dependency grammar : A new methodology. In *COLING'04 Workshop*, pages 78–84. Geneva.

Joshi, Aravind K., Vijay K. Shanker, and David J. Weir. 1991. The convergence of mildly context-sensitive grammar formalisms. In P. Sells, S. Shieber, and T. Wasow, eds., *Foundational issues in natural language processing*, pages 31–81. Cambridge, MA: MIT Press.

Kruijff, G-J.M. 2001. *Dependency Grammar Logic and Information Structure*. Ph.D. thesis, Charles University, Prague.

Moortgat, Michael. 1997. Categorial type logics. In J. van Benthem and A. ter Meulen, eds., *Handbook of Logic and Language*, chap. 2, pages 93–177. Amsterdam, Cambridge: North-Holland Elsevier, The MIT Press.

Moortgat, Michael and Glin V. Morrill. 1991. Heads and phrases. Type calculus for dependency and constituent structure. Ms OTS, Utrecht.

Morrill, Glin V. 1994. *Type Logical Grammar. Categorial Logic of Signs*. Dordrecht: Kluwer Academic Publishers.

Pulman, S.G. and G.D. Ritchie. 1985. Indexed grammars and interesting dependencies. *UEA Papers in Linguistics* 23:21–38.