Multimodal Categorial Dependency Grammars

Alexander Dikovsky

LINA CNRS FRE 2729, Université de Nantes

Objective and Plan

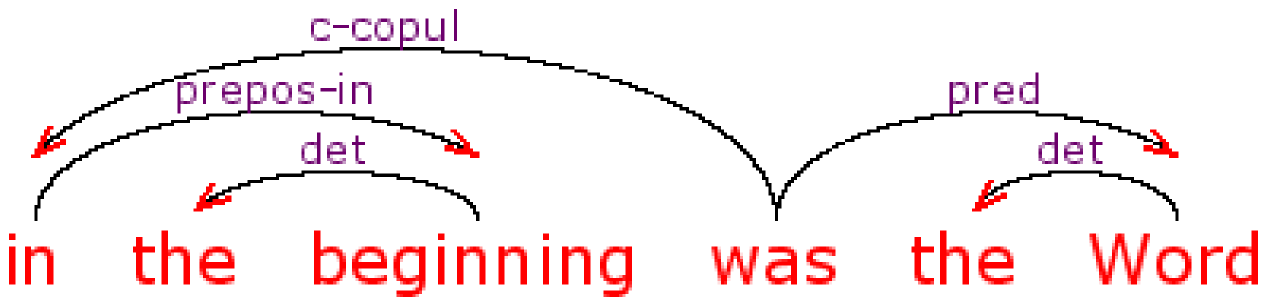
OBJECTIVE: To express unlimited discontinuous dependencies using Categorial Dependency Grammars:

- with 1st order types
- without "dictionary explosion" when the order is flexible
- parsable in a low polynomial time

PLAN:

- Dependencies: projective vs. discontinuous
- Categorial Dependency Grammars (CDG)
- Multimodal CDG (mmCDG)
- mmCDG for cross-serial dependencies and for scrambling
- Expressivity / Complexity
- Conclusion

Projective dependencies



Dependency : Governor $\stackrel{d}{\longrightarrow}$ Subordinate

Government = Dependency$^{(refl,trans)}$ $(\longrightarrow^*)$

Precedence : (total) WO $(<)$

Projection of a word w :

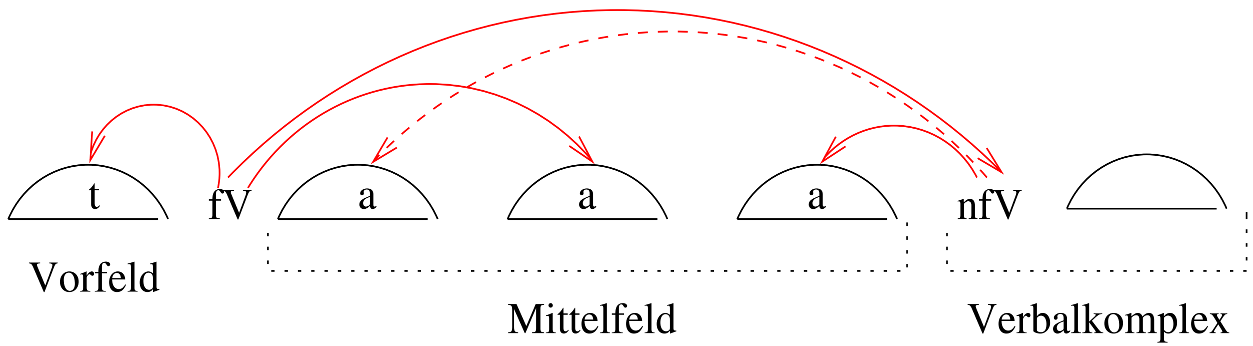
$$d(w) = \{w' \parallel w \longrightarrow^* w'\} \text{ ordered by the WO}$$

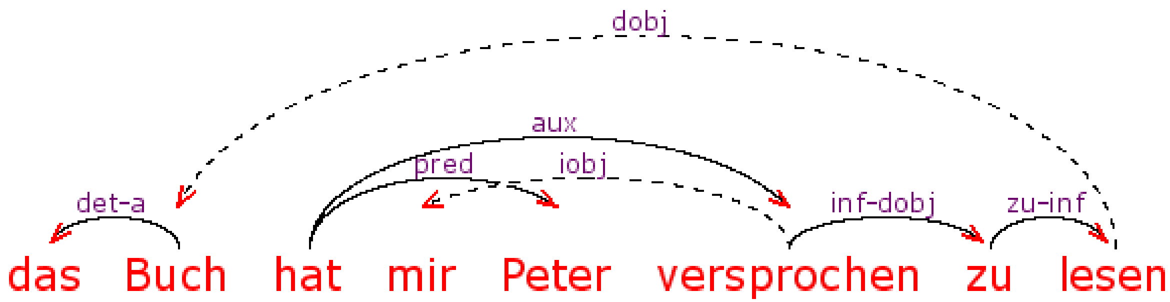
Projective dependency tree D : $\{d(w) \parallel w \in D\}$ is a CS

Non-projective dependencies

Clitics in French :

Scrambling in German :

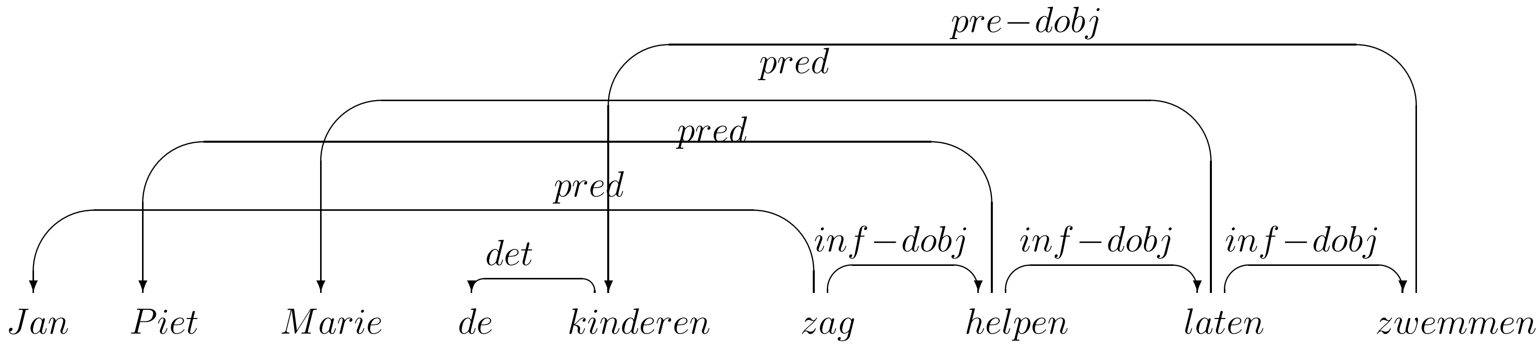




Bulgarian, Russian : quasi free WO in the matrix clause

Cross-serial dependencies

Dutch (Bresnan,Kaplan,Peters,Zaenen'1982) :



Jan Piet Marie children saw help make swim

Sentences : $n_1 n_2 \ldots n_m n_{m+1} v_1 v_{(inf)2} \ldots v_{(inf)m}$, where :

$1.$ a predicative dependency $n_1 \overset{pred}{\longleftarrow} v_1$ from a finite verb v_1 to a noun n_1,

$2.$ predicative dependencies $n_i \overset{pred}{\longleftarrow} v_{(inf)i}$ from verbs $v_{(inf)i}$ in infinitive to nouns n_i, $2 \le i \le m$ (no accord)

$3.$ possibly, a direct object dependency $n_{m+1} \overset{dobj}{\longleftarrow} v_{(inf)m}$ if the verb $v_{(inf)m}$ is transitive and the name n_{m+1} is present.

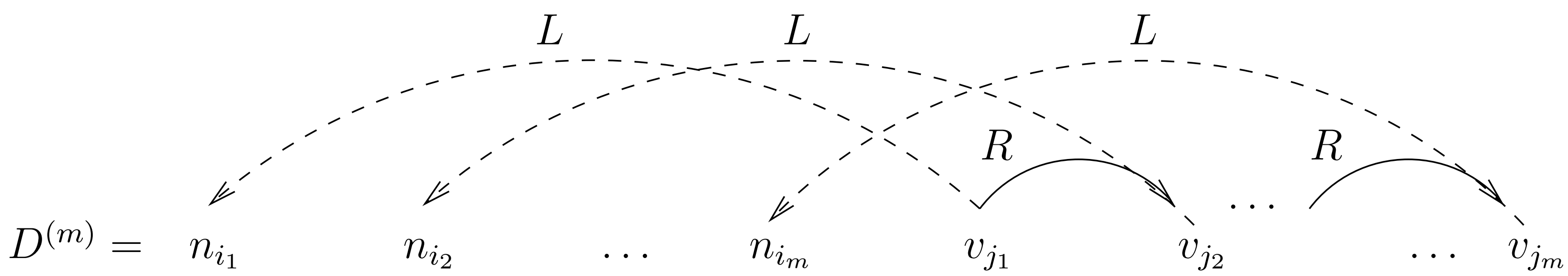
Cross-serial dependencies continued

REMARK : So an adequate model of this construction is not the "copy language"

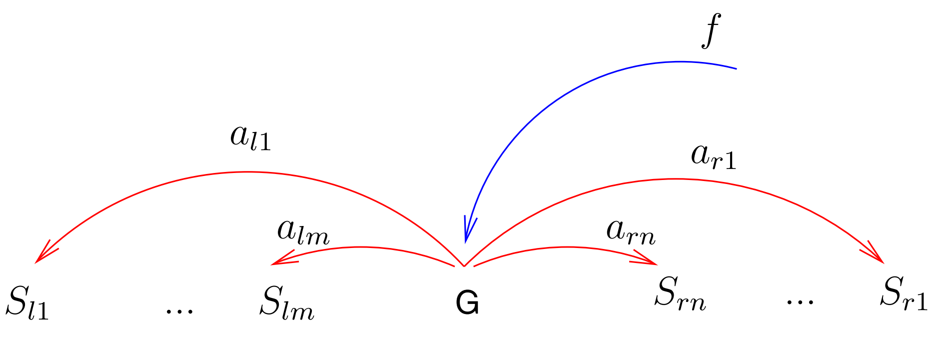
$$L_{copy} = \{xx \parallel x \in W^*\},$$

but the tree language $\Delta_{cross} = \{D^{(m)} \parallel m > 0\}$,

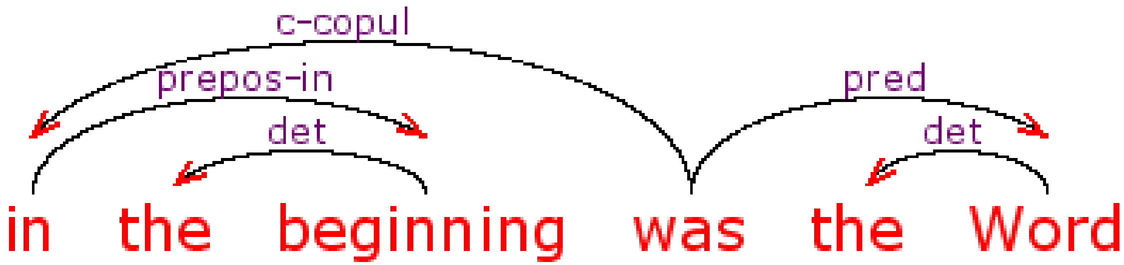
where $D^{(m)}$ is the following tree ($n_{i_l} \in N, v_{j_r} \in V$) :



Types of projective dependencies



$$G \mapsto [a_{lm}\backslash \ldots \backslash a_{l1}\backslash f / a_{r1} / \ldots / a_{rn}]$$



in $\mapsto [c\!-\!copul/prepos\!-\!in]$

the $\mapsto [det]$

beginning $\mapsto [det\backslash prepos\!-\!in]$

was $\mapsto [c\!-\!copul\backslash S/pred]$

Word $\mapsto [det\backslash pred]$

Types of non-projective dependencies

Polarized valencies :

$Gov_l \mapsto [\alpha]^{\nearrow d}$ $\qquad$ $Gov_r \mapsto [\alpha]^{\nwarrow d}$

$Sub_l \mapsto [\beta]^{\searrow d}$ $\qquad$ $Sub_r \mapsto [\beta]^{\swarrow d}$

EX :

la $\overset{clit-dobj}{\longleftarrow}$ donnée

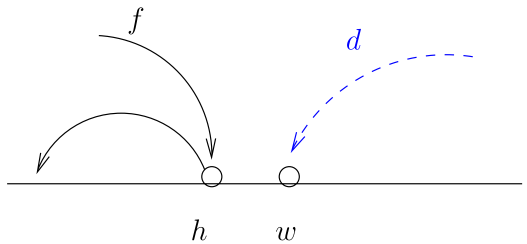
la $\mapsto [\varepsilon]^{\swarrow clit\text{–}dobj}$

donnée $\mapsto [aux]^{\nwarrow clit\text{–}iobj \nwarrow clit\text{–}dobj}$

Anchored valencies (to place distant subordinates) :

h anchors w : $h \mapsto [\beta\backslash f / \#^r(\swarrow d)]$

w anchored on h : $w \mapsto [\#^r(\swarrow d)/\alpha]^{\swarrow d}$



Clitics in French

elle $\mapsto [pred]$

la $\mapsto [\#^l(\swarrow clit-dobj)]^{\swarrow clit-dobj}$

lui $\mapsto [\#^l(\swarrow clit-iobj)]^{\swarrow clit-iobj}$

a $\mapsto [\#^l(\swarrow clit-iobj)\backslash\#^l(\swarrow clit-dobj)\backslash pred\backslash S/aux]$

donnée $\mapsto [aux]^{\nwarrow clit-iobj \nwarrow clit-dobj}$

Categorial dependency grammar

$$\mathbf{L^l}.\quad C^{P_1}[C\backslash\beta]^{P_2} \vdash [\beta]^{P_1P_2}$$

$$\mathbf{I^l}.\quad C^{P_1}[C^*\backslash\beta]^{P_2} \vdash [C^*\backslash\beta]^{P_1P_2}$$

$$\mathbf{\Omega^l}.\quad [C^*\backslash\beta]^{P} \vdash [\beta]^{P}$$

$$\mathbf{D^l}.\quad \alpha^{P_1(\swarrow C)P(\nwarrow C)P_2} \vdash \alpha^{P_1PP_2},$$

if $(\swarrow C)P(\nwarrow C)$ satisfies the following principle of valency pairing :

> Principle **FA** : v is saturated in $P_1 v P \breve{v} P_2$ with the closest available valency $\breve{v}$ (i.e. P has no occurrences of $v, \breve{v}$)

Elimination of dual valencies $v = (\swarrow C)$, $\breve{v} = (\nwarrow C)$ with the rule $\mathbf{D^l}$ creates the discontinuous dependency C.

Elimination of anchored valency $\#(v)$ with the rules $\mathbf{L^l}, \mathbf{I^l}$ creates no dependency.

Typing correctness proofs

EX : *elle la lui a donnée*

Without anchoring :

$$[pred]\ [\swarrow clit-dobj]\ [\swarrow clit-iobj]\ [pred\backslash S/aux]\ [\nwarrow clit-iobj\backslash \nwarrow clit-dobj\backslash aux]$$

$$\dfrac{[\varepsilon]^{\swarrow clit-iobj}\,[pred\backslash S/aux]}{[pred\backslash S/aux]^{\swarrow clit-iobj}}\ (\mathbf{L}^{l})$$

$$\dfrac{[\varepsilon]^{\swarrow clit-dobj}\quad [pred\backslash S/aux]^{\swarrow clit-iobj}}{[pred\backslash S/aux]^{\swarrow clit-dobj\swarrow clit-iobj}}\ (\mathbf{L}^{l})$$

$$\dfrac{[pred]\quad [pred\backslash S/aux]^{\swarrow clit-dobj\swarrow clit-iobj}}{[S/aux]^{\swarrow clit-dobj\swarrow clit-iobj}}\ (\mathbf{L}^{\mathbf{l}})$$

$$\dfrac{[S/aux]^{\swarrow clit-dobj\swarrow clit-iobj}\quad [aux]^{\nwarrow clit-iobj\nwarrow clit-dobj}}{[S]^{\swarrow clit-dobj\swarrow clit-iobj\nwarrow clit-iobj\nwarrow clit-dobj}}\ (\mathbf{L}^{r})$$

$$\dfrac{[S]^{\swarrow clit-dobj\swarrow clit-iobj\nwarrow clit-iobj\nwarrow clit-dobj}}{S}\ (\mathbf{D}^{l}\times 2)$$

With anchoring :

$$[pred]\ [\#^{l}(\swarrow clit-dobj)]\ [\#^{l}(\swarrow clit-iobj)]\ [\#^{l}(\swarrow clit-iobj)\backslash\#^{l}(\swarrow clit-dobj)\backslash pred\backslash S/aux]\ [\nwarrow clit-iobj\backslash \nwarrow clit-dobj\backslash aux]$$

$$\dfrac{[\#^{l}(\swarrow clit-iobj)]^{\swarrow clit-iobj}\,[\#^{l}(\swarrow clit-iobj)\backslash\#^{l}(\swarrow clit-dobj)\backslash pred\backslash S/aux]}{[\#^{l}(\swarrow clit-dobj)\backslash pred\backslash S/aux]^{\swarrow clit-iobj}}\ (\mathbf{L}^{l})$$

$$\dfrac{[\#^{l}(\swarrow clit-dobj)]^{\swarrow clit-dobj}\quad [\#^{l}(\swarrow clit-dobj)\backslash pred\backslash S/aux]^{\swarrow clit-iobj}}{[pred\backslash S/aux]^{\swarrow clit-dobj\swarrow clit-iobj}}\ (\mathbf{L}^{l})$$

$$\dfrac{[pred]\quad [pred\backslash S/aux]^{\swarrow clit-dobj\swarrow clit-iobj}}{[S/aux]^{\swarrow clit-dobj\swarrow clit-iobj}}\ (\mathbf{L}^{\mathbf{l}})$$

$$\dfrac{[S/aux]^{\swarrow clit-dobj\swarrow clit-iobj}\quad [aux]^{\nwarrow clit-iobj\nwarrow clit-dobj}}{[S]^{\swarrow clit-dobj\swarrow clit-iobj\nwarrow clit-iobj\nwarrow clit-dobj}}\ (\mathbf{L}^{r})$$

$$\dfrac{[S]^{\swarrow clit-dobj\swarrow clit-iobj\nwarrow clit-iobj\nwarrow clit-dobj}}{S}\ (\mathbf{D}^{l}\times 2)$$

Multimodal architecture in CDG

PB : FA is not adequate for the cross-serial dependencies

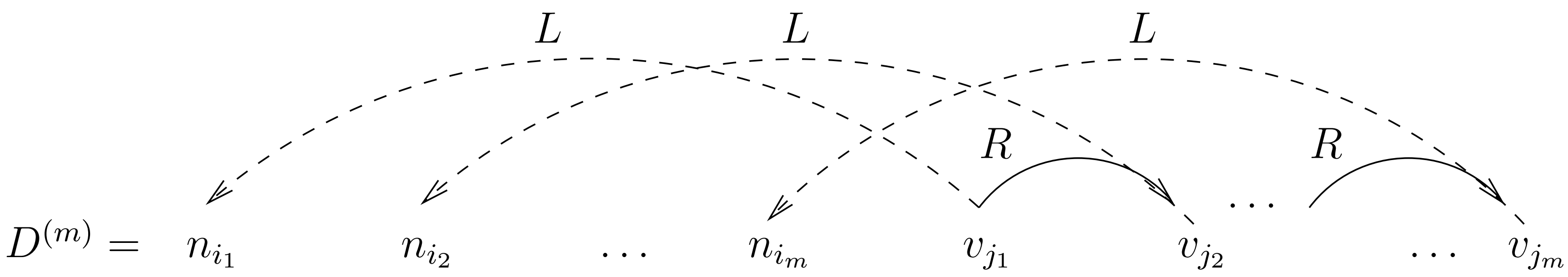
Multi-modal CDG (mmCDG) : every dependency may have its own compositionality mode :

- argument position variations for projective dependencies,
- valency pairing principles for non-projective dependencies

For the Dutch predicative dependency of verbs in the infinitive, the pairing principle is :

Principle **FC** : in $P_1(\swarrow C)P(\nwarrow C)P_2$, P_1 has no occurrences of $\swarrow C$ and P has no occurrences of $\nwarrow C$ (i.e. v is saturated with the first available cross dual valency $\breve{v}$)

Dutch cross-serial dependencies : a solution



Δ_{cross} is generated by the following $mmCDG^{\mathbf{FC}}$:

$$G_{cross} = \begin{cases} n \mapsto [\#(\swarrow L)]^{\swarrow L}, [\#(\swarrow L)\backslash\#(\swarrow L)]^{\swarrow L}, & \text{for } n \in N \\ v \mapsto [\#(\swarrow L)\backslash S/R]^{\nwarrow L}, [R/R]^{\nwarrow L}, [R]^{\nwarrow L}, & \text{for } v \in V \end{cases}$$

E.g., $D^{(3)} \in \Delta_{cross}$ due to the proof :

$$\frac{[\#(\swarrow L)]^{\swarrow L}[\#(\swarrow L)\backslash\#(\swarrow L)]^{\swarrow L}}{[\#(\swarrow L)]^{\swarrow L \swarrow L}}\ (\mathbf{L}^l)$$

$$\frac{[\#(\swarrow L)]^{\swarrow L \swarrow L} \quad [\#(\swarrow L)\backslash\#(\swarrow L)]^{\swarrow L}}{[\#(\swarrow L)]^{\swarrow L \swarrow L \swarrow L}}\ (\mathbf{L}^l)$$

$$\frac{[R/R]^{\nwarrow L}[R]^{\nwarrow L}}{[R]^{\nwarrow L \nwarrow L}}\ (\mathbf{L}^r)$$

$$\frac{[\#(\swarrow L)\backslash S/R]^{\nwarrow L} \quad [R]^{\nwarrow L \nwarrow L}}{[\#(\swarrow L)\backslash S]^{\nwarrow L \nwarrow L \nwarrow L}}\ (\mathbf{L}^r)$$

$$\frac{[\#(\swarrow L)]^{\swarrow L \swarrow L \swarrow L} \quad [\#(\swarrow L)\backslash S]^{\nwarrow L \nwarrow L \nwarrow L}}{[S]^{\swarrow L \swarrow L \swarrow L \nwarrow L \nwarrow L \nwarrow L}}\ (\mathbf{L}^l)$$

$$\frac{[S]^{\swarrow L \swarrow L \swarrow L \nwarrow L \nwarrow L \nwarrow L}}{[S]}\ (\mathbf{D}^l_{\mathbf{FC}} \times 3)$$

Scrambling : projective case

Partially ordered subtypes for flexible order: $[\{C_1, \ldots, C_k\}\backslash\alpha]$

PO subtypes elimination rule :

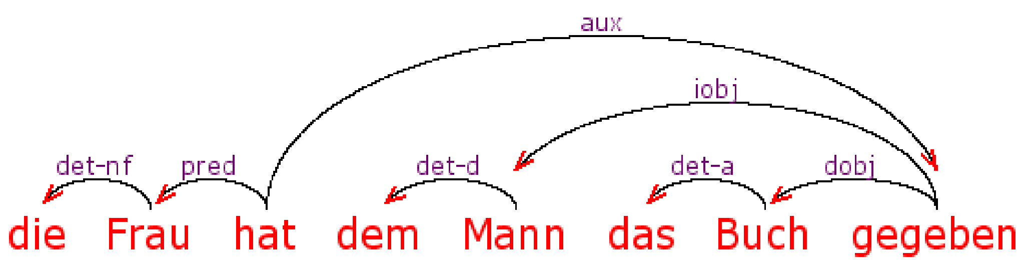
F$^{\mathbf{l}}$. $C^{P_1}[\{\alpha_1, C, \alpha_2\}\backslash\beta]^{P_2} \vdash [\{\alpha_1, \alpha_2\}\backslash\beta]^{P_1 P_2}$

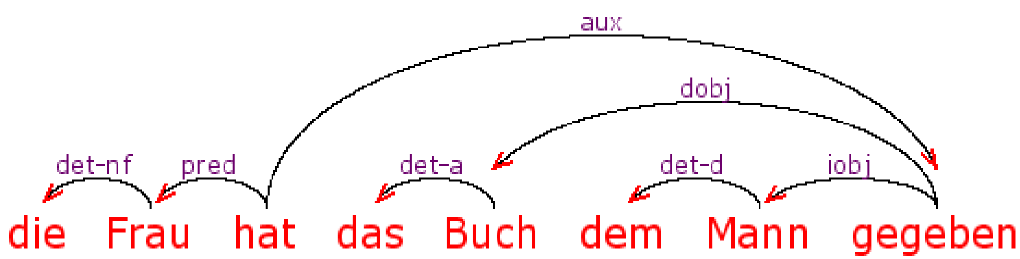
EX Mittelfeld arguments permutations:

hat : $\mapsto [pred\backslash S/aux]$ gegeben : $\mapsto [\{dobj, iobj\}\backslash aux]$

die : $\mapsto [det_{nf}]$ dem : $\mapsto [det_d]$ das : $\mapsto [det_a]$

Frau : $\mapsto [det_{nf}\backslash pred]$ Mann : $\mapsto [det_d\backslash iobj]$ Buch : $\mapsto [det_a\backslash dobj]$





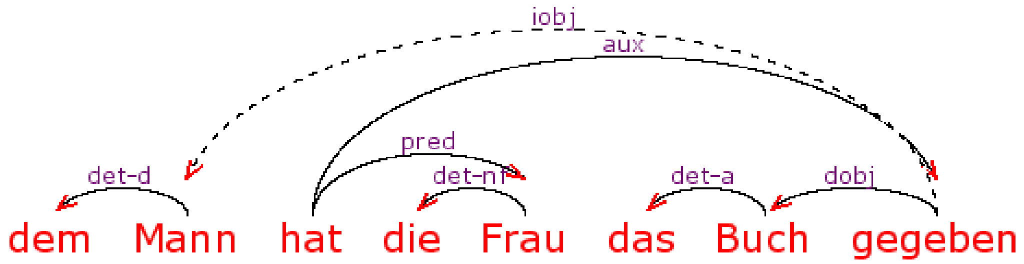
Scrambling : non-projective case

EX Vorfeld-Mittelfeld arguments permutations:

hat : $\mapsto [\#(\swarrow iobj)\backslash S/aux/pred]$ gegeben : $\mapsto [dobj\backslash aux]^{\nwarrow iobj}$

Mann : $\mapsto [det_d\backslash\#(\swarrow iobj)]^{\swarrow iobj}$ Buch : $\mapsto [det_a\backslash dobj]$

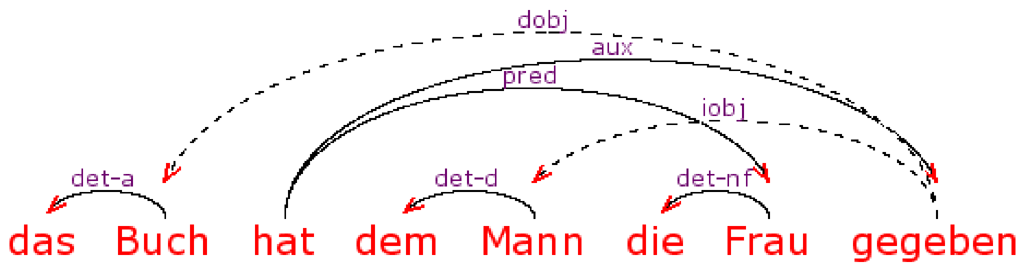
Frau : $\mapsto [det_{nf}\backslash pred]$



Mann : $\mapsto [det_d\backslash\#(\swarrow iobj)]^{\swarrow iobj}$ Buch : $\mapsto [det_a\backslash\#(\swarrow dobj)]^{\swarrow dobj}$

hat : $\mapsto [\#(\swarrow dobj)\backslash S/aux/pred/\#(\swarrow iobj)]$

gegeben : $\mapsto [aux]^{\nwarrow dobj \nwarrow iobj}$ Frau : $\mapsto [det_{nf}\backslash pred]$



Independence of projections

Projections. *local* : $\|C^P\|_l = C$ and *valency* : $\|C^P\|_v = P$

Projective dependecy calculus $\vdash_{\mathbf{c}}$: rules $\mathbf{L^l}$, $\mathbf{I^l}$, $\mathbf{\Omega^l}$.

TH For a gCDG $G = (W, \mathbf{C}, S, \delta)$ and $x \in W^+$,

$x \in L(G)$ iff there is $\Gamma \in \delta(x)$ such that:

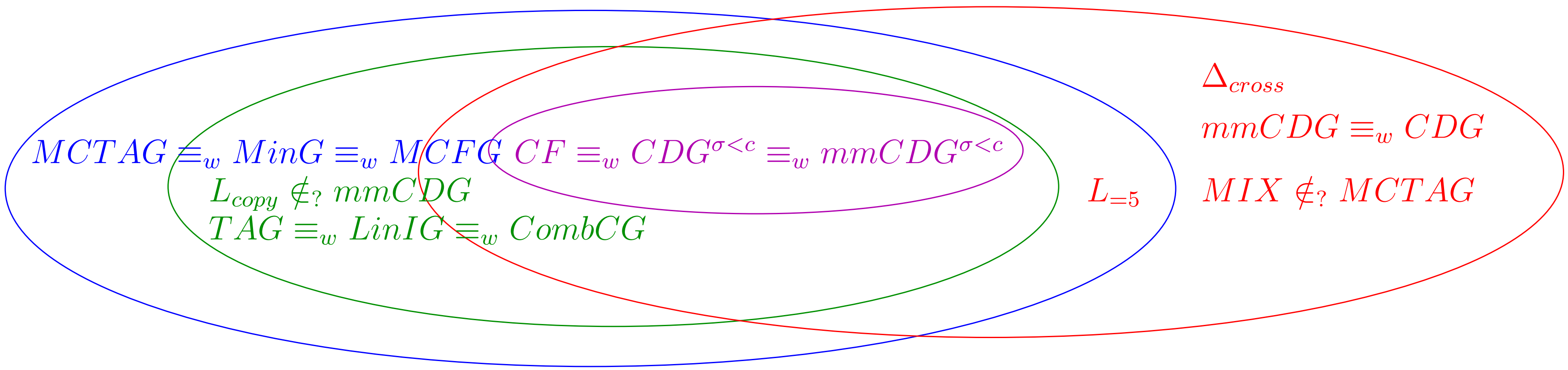
1. $\|\Gamma\|_l \vdash_{\mathbf{c}} S$,
2. $\|\Gamma\|_v$ is well-bracketed for every valency.

This criterion holds for both principles **FA** and **FC**

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

Expressivity / complexity

Valency deficit $\sigma(G)$: maximal potential size in proofs of G.



$$L_{copy} = \{ww \parallel w \in W^+\}, \quad L_{=i} = \{a_1^n \ldots a_i^n \parallel n > 0\}, \quad MIX = \{w \in \{a, b, c\}^+ \parallel |w|_a = |w|_b = |w|_c\}$$

Parsing complexity (Dekhtyar,Dikovsky'2004) :

1. theoretical : $O(n^{3+2p})$ (for p polarized valencies)
2. in practice ($\sigma(G) < const$) : $O(n^3)$.

Conclusion

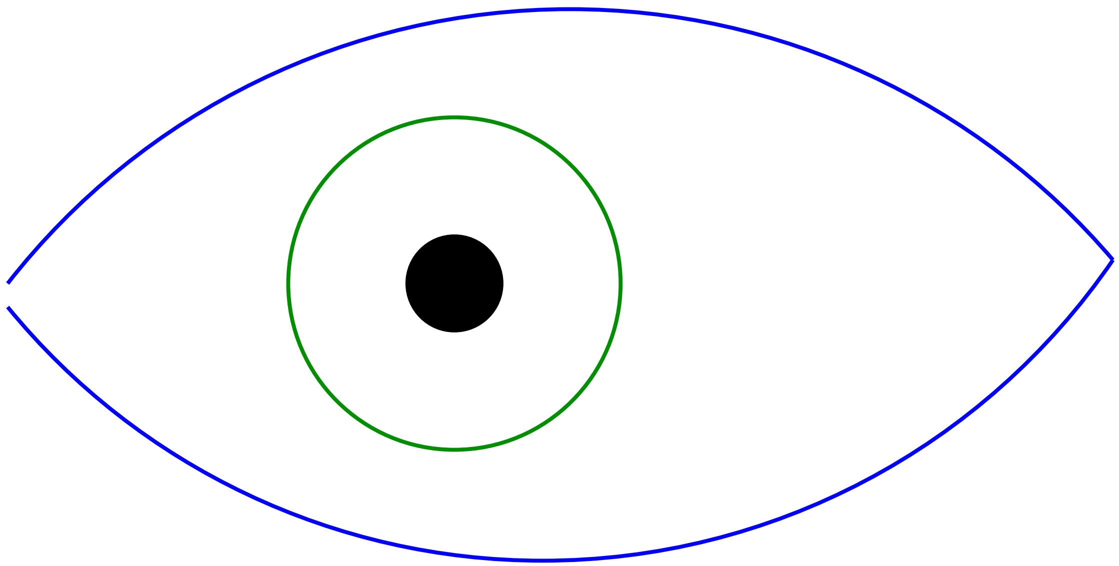
Multimodal CDG :

1. express unlimited discontinuous dependencies and flexible WO,
2. are parsed in polynomial time ($O(n^3)$ for real grammars),
3. are simple to use in practice

Work in progress :

1. tools for real size grammars (grammars of large corpora, morphology, lexical DB),
2. incremental parser trained over corpora,
3. compositional logical semantics.

Parser demonstration



A gCDG for a non-CF language

$$G_{abc}: \begin{cases} a & \mapsto \quad A^{\swarrow A}, [A\backslash A]^{\swarrow A} \\ b & \mapsto \quad [B/C]^{\nwarrow A}, [A\backslash S/C]^{\nwarrow A} \\ c & \mapsto \quad C, [B\backslash C] \end{cases}$$

TH: $\mathcal{L}(G_{abc}) = \{a^n b^n c^n \mid n > 0\}.$

A derivation of $a^3b^3c^3 \in L(G_{abc})$:

Types assignment:

$$a^3b^3c^3 \mapsto A^{\swarrow A}[A\backslash A]^{\swarrow A}[A\backslash A]^{\swarrow A}[A\backslash S/C]^{\nwarrow A}[B/C]^{\nwarrow A}[B/C]^{\nwarrow A}C[B\backslash C][B\backslash C]$$

$$\frac{[A]^{\swarrow A}[A\backslash A]^{\swarrow A}}{[A]^{\swarrow A\swarrow A}}\ (\mathbf{L}^l)$$

$$\frac{[A]^{\swarrow A\swarrow A}\quad [A\backslash A]^{\swarrow A}}{[A]^{\swarrow A\swarrow A\swarrow A}}\ (\mathbf{L}^l)$$

$$\frac{[B/C]^{\nwarrow A}C}{B^{\nwarrow A}}\ (\mathbf{L}^r)$$

$$\frac{B^{\nwarrow A}\quad [B\backslash C]}{C^{\nwarrow A}}\ (\mathbf{L}^l)$$

$$\frac{[B/C]^{\nwarrow A}\quad C^{\nwarrow A}}{B^{\nwarrow A\nwarrow A}}\ (\mathbf{L}^r)$$

$$\frac{B^{\nwarrow A\nwarrow A}\quad [B\backslash C]}{C^{\nwarrow A\nwarrow A}}\ (\mathbf{L}^l)$$

$$\frac{[A\backslash S/C]^{\nwarrow A}\quad C^{\nwarrow A\nwarrow A}}{[A\backslash S]^{\nwarrow A\nwarrow A\nwarrow A}}\ (\mathbf{L}^l)$$

$$\frac{[A]^{\swarrow A\swarrow A\swarrow A}\quad [A\backslash S]^{\nwarrow A\nwarrow A\nwarrow A}}{[S]^{\swarrow A\swarrow A\swarrow A\nwarrow A\nwarrow A\nwarrow A}}\ (\mathbf{L}^l)$$

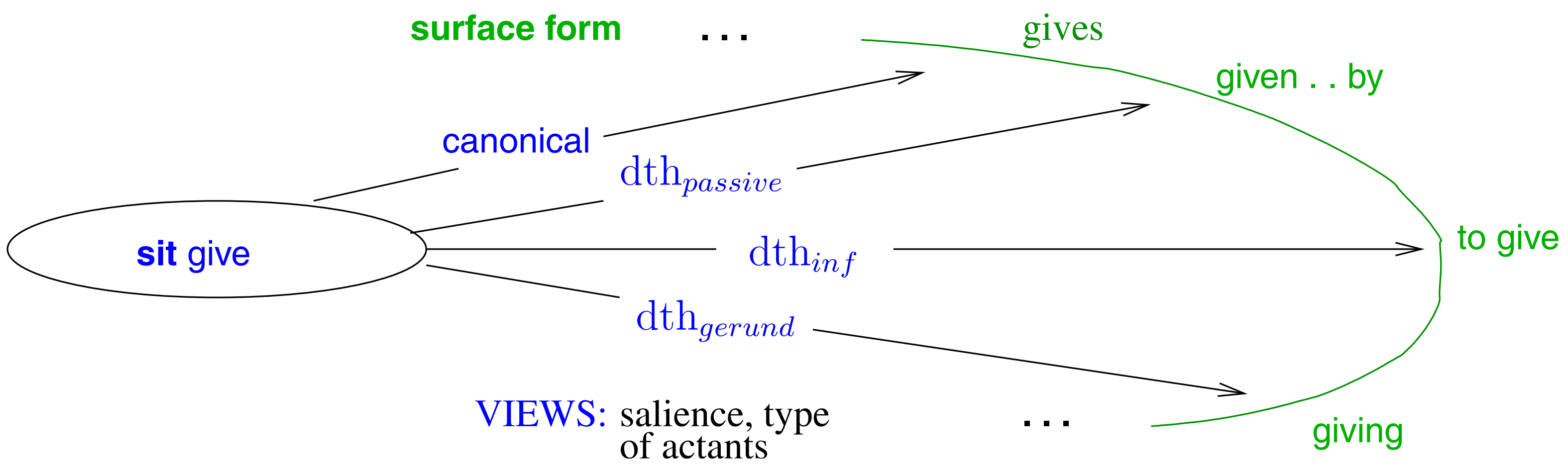
$$\frac{[S]^{\swarrow A\swarrow A\swarrow A\nwarrow A\nwarrow A\nwarrow A}}{S}\ (\mathbf{D}^l \times 3)$$

Semantic sources of dependencies

- Surface dependencies originate from semantic argument-function relations represented with underspecified semantic structures that we call discourse plans (DP). The DP determine compositions of *situations* or their derivatives, through *semantic diatheses*
- Types of dependencies can be calculated from DP using finite-state tree transducers

Situations, diatheses

Situations : invariants of communicative views (= semantic diatheses)



EX(G. Frege, "Begriffsschrift", 1879) :

Bei Platae siegten die Griechen über die Perser

Bei Platae wurden die Perser von den Griechen besiegt

common situation : siegen(SBJ, OBJ)

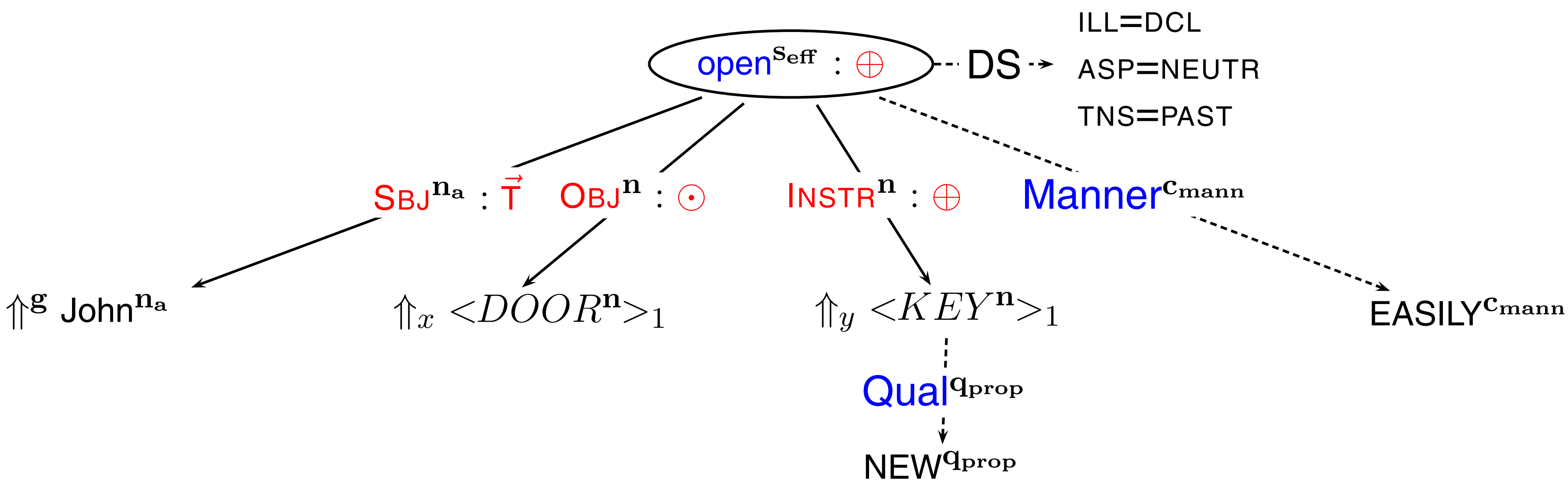
- Arguments of situations :
  - actants, identified with their thematic roles, vary from one diathesis derivative to another according to the intended communicative ranks;
  - circumstantials, identified with other (optional) attributes,
- communicative ranks : $\vec{T}$ (topic $\simeq$, O (implied topic 0), $\odot$ (rhematic focus), $\oplus$ (background), $\ominus$ (periphery)

⟪open⟫ : canonical profile

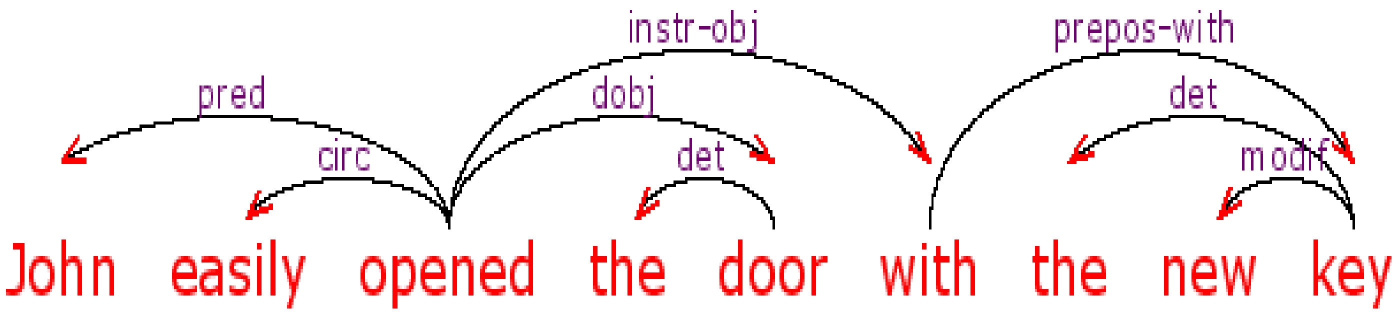
EX: *John easily opened the door with the new key*

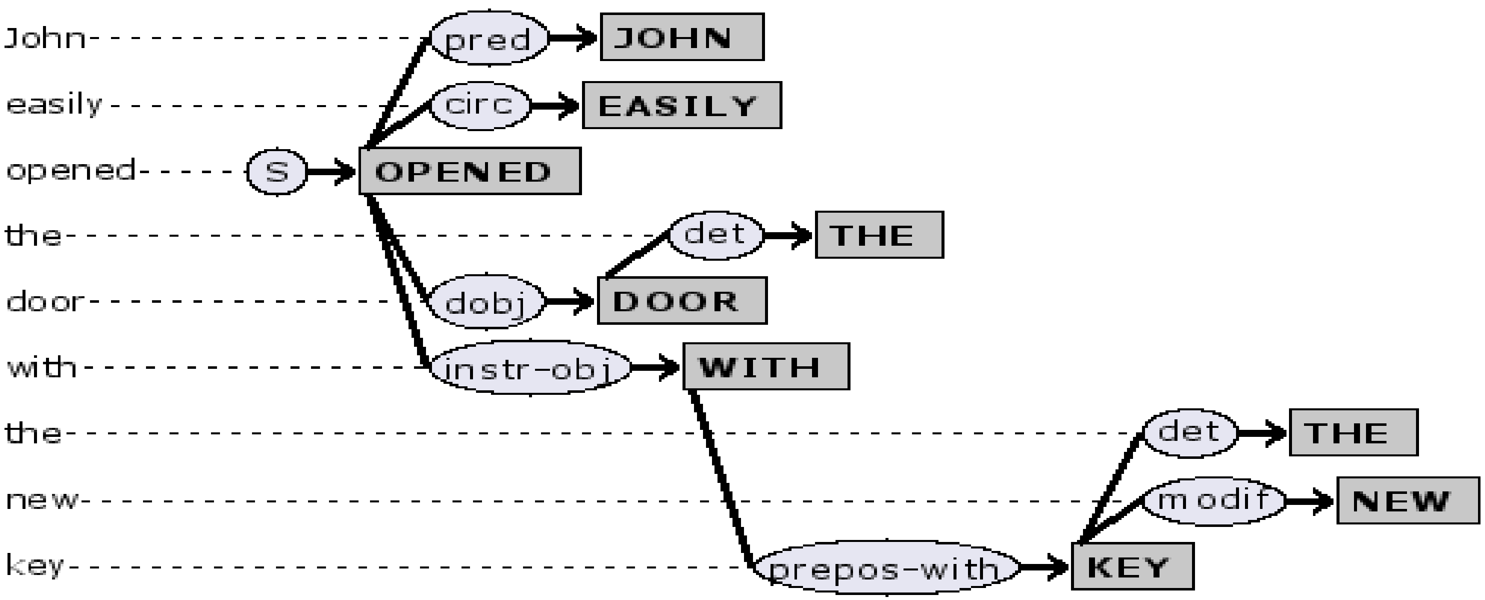
Situation ⟪open⟫. Canonical profile :

$\langle\!\langle \text{open}\left(\text{SBJ}^{\mathbf{n_a}}, \text{OBJ}^{\mathbf{n}}, \text{INSTR}^{\mathbf{n}}\right)^{\mathbf{s_{eff}}} \rangle\!\rangle$



DT for canonical profile

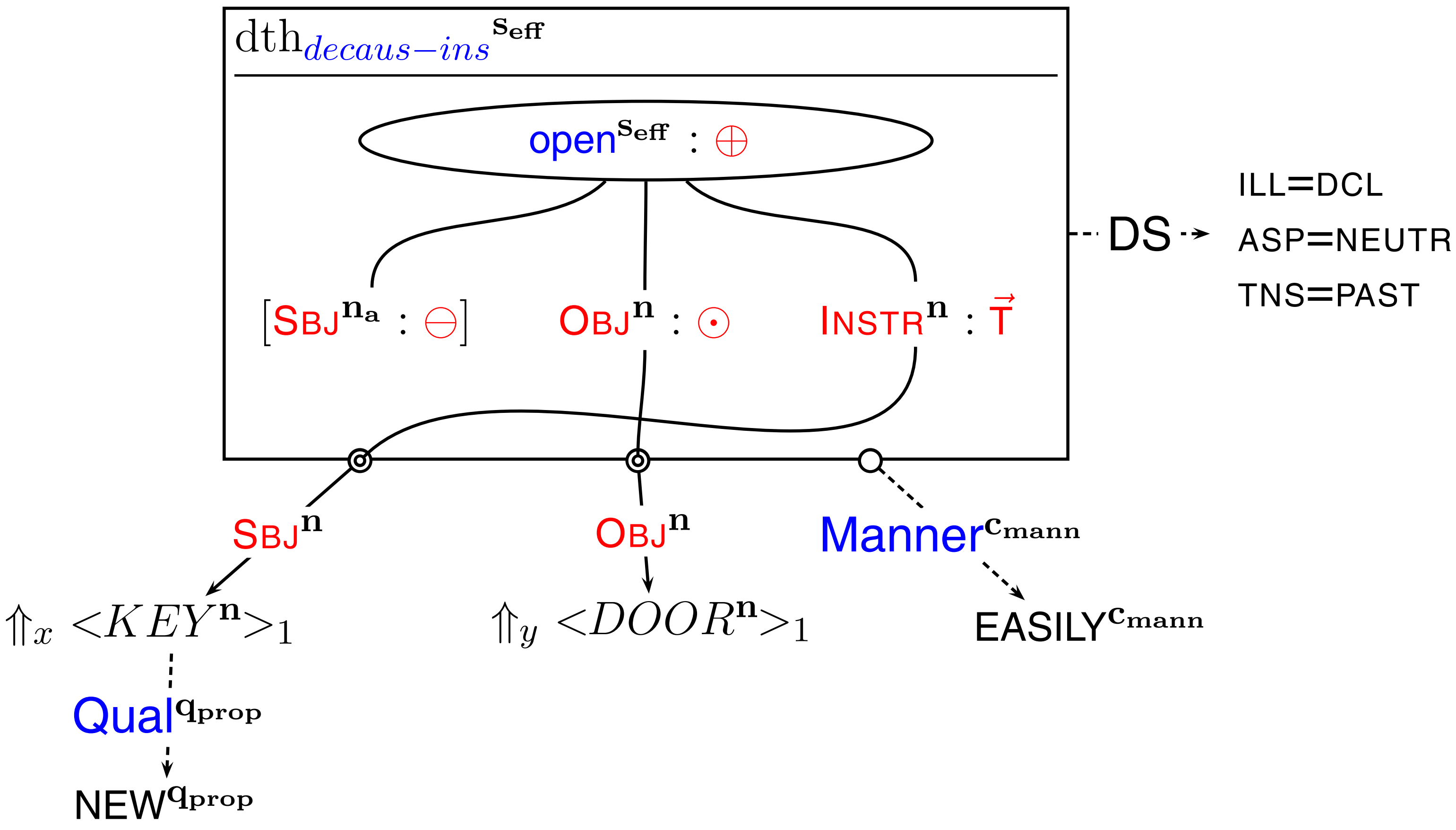




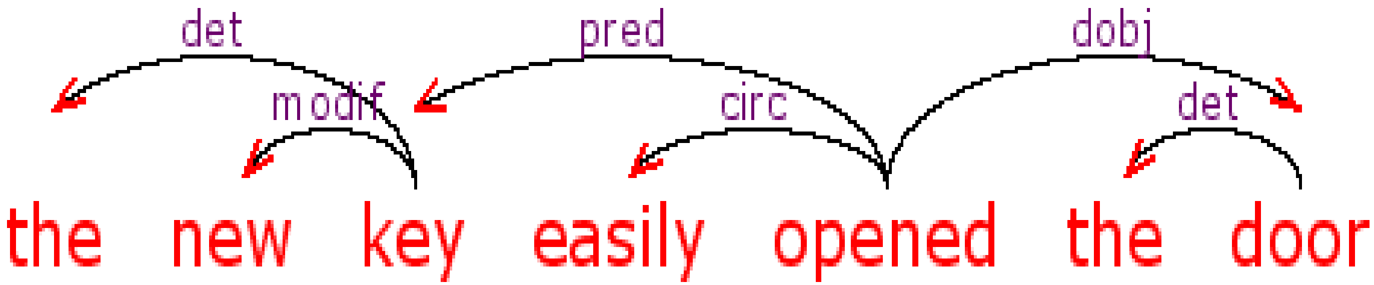
⟪open⟫ : 1^{st} object alternation

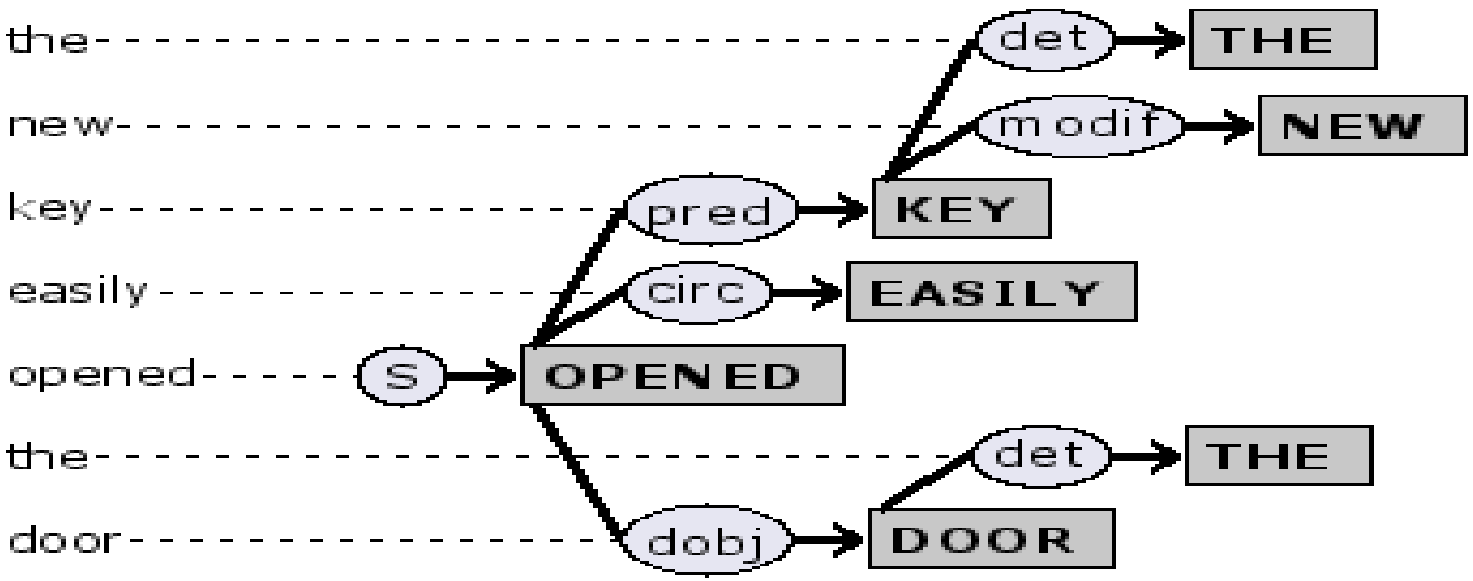
EX: *The new key easily opened the door*

⟪open⟫: $\langle\!\langle \mathrm{dth}_{decaus-ins}\,(\bigstar_{\oplus}, \emptyset \hookleftarrow \textsc{Sbj}_{\ominus}, \textsc{Sbj} \hookleftarrow \textsc{Instr}_{\vec{\top}}, \textsc{Obj} \hookleftarrow \textsc{Obj}_{\odot})^{\mathbf{s_{eff}}} \rangle\!\rangle$



Dependency tree

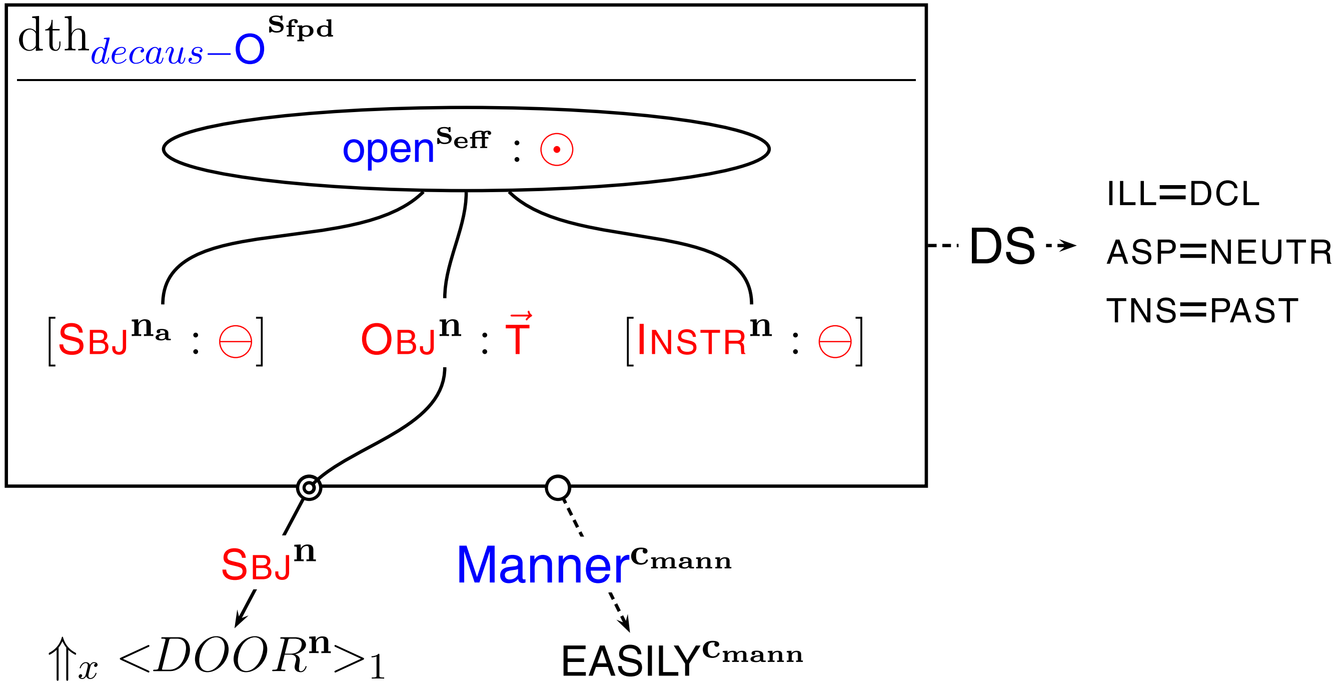




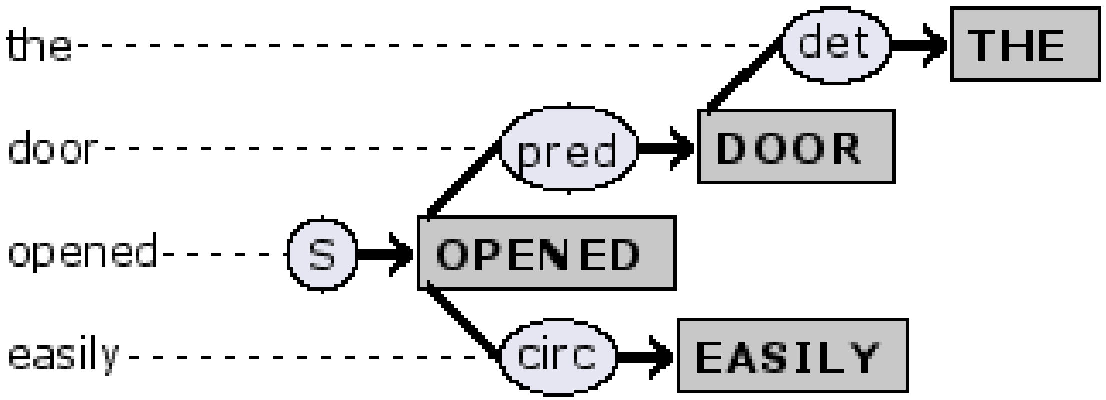
⟪open⟫ : 2^d object alternation

EX: *The door opened easily*

$\langle\!\langle$open$\rangle\!\rangle$: $\langle\!\langle \mathrm{dth}_{decaus-\mathsf{O}}\,(\bigstar_{\odot}, \emptyset \hookleftarrow \textsc{Sbj}_{\ominus}, \textsc{Sbj} \hookleftarrow \textsc{Obj}_{\vec{\mathsf{T}}}, \emptyset \hookleftarrow \textsc{Instr}_{\ominus})^{\mathbf{S_{fpd}}} \rangle\!\rangle$



Dependency tree

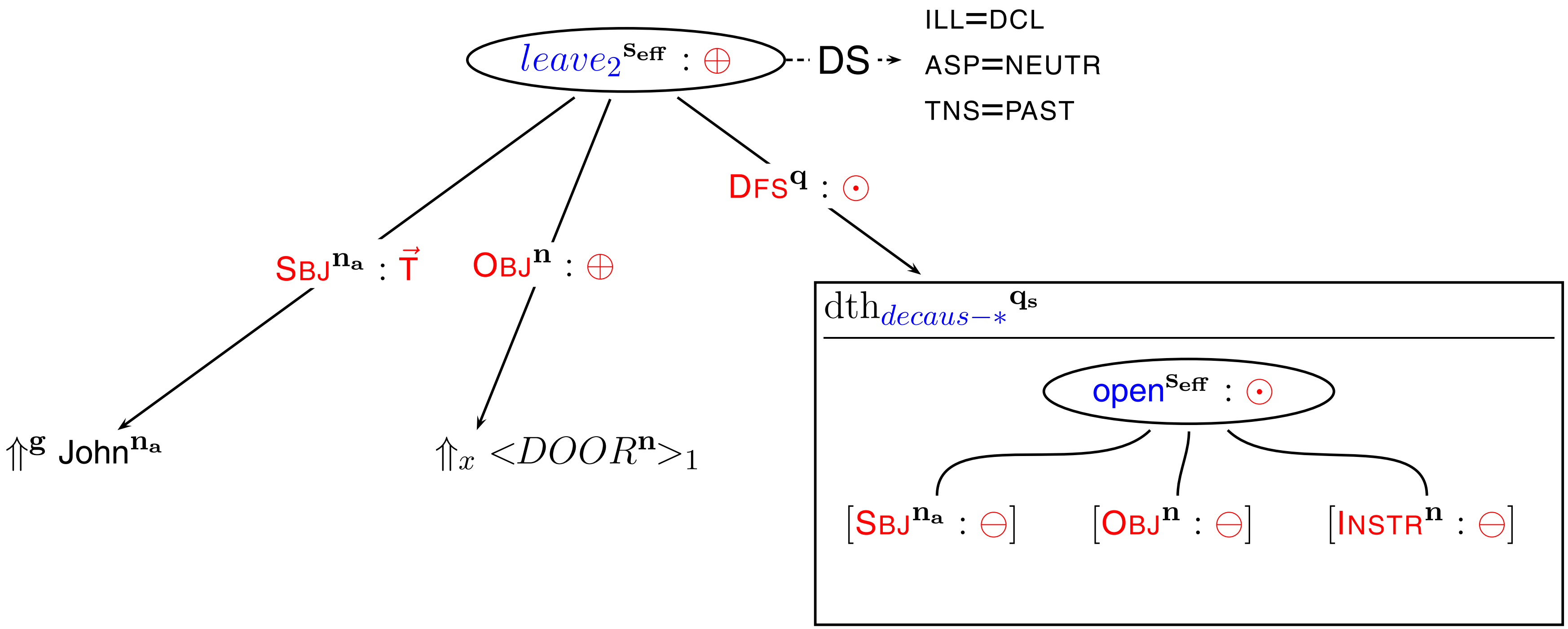


⟪open⟫ : objects elimination

EX: *John left the door opened*

⟪open⟫: $\langle\!\langle \mathrm{dth}_{decaus-*}\,(\bigstar_{\odot}, \emptyset \hookleftarrow \textsc{Sbj}_{\ominus}, \emptyset \hookleftarrow \textsc{Obj}_{\ominus}, \emptyset \hookleftarrow \textsc{Instr}_{\ominus})^{\mathbf{q_s}} \rangle\!\rangle$

$\langle\!\langle leave_2 \rangle\!\rangle$ canonical profile: $\langle\!\langle leave_2(\textsc{Sbj}^{\mathbf{n_a}}, \textsc{Obj}^{\mathbf{n}}, \textsc{Dfs}^{\mathbf{q}})^{\mathbf{s_{eff}}} \rangle\!\rangle$



Dependency tree