A Finite-State Functional Grammar Architecture

Alexander Dikovsky

LINA CNRS FRE 2729, Université de Nantes

PLAN

- New proposal and its context
- FS-Functional grammar
  - Output: dependency types
  - Input: Discourse Plans
  - FS-conversion
- Expressiveness/Complexity
- Conclusion

Generativistic architectures

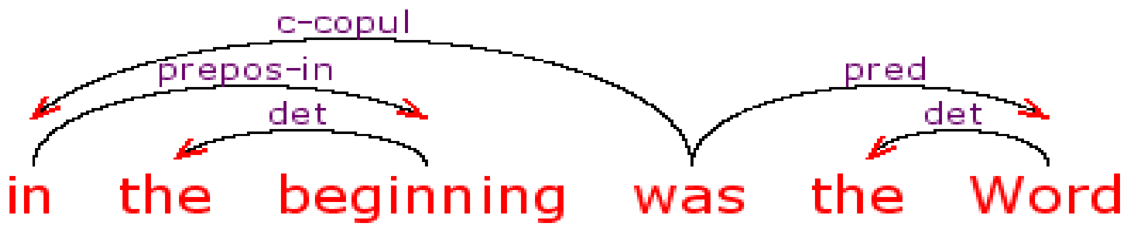
- Chomskian (enumerative and autonomous wrt semantics)
  - Generative Semantics [G.Lakoff, J.Ross, P.Postal, J.McCawley] (deep-structure $\Rightarrow$ c-structure)
- Meaning-Text Theory [I.Mel'čuk] (multi-level, rule-based, meaning graphs $\Rightarrow$ d-structure)
- Prague Functional Generative Description [P.Sgall, E.Hajičova et al., G-J.Kruijff] (tectogrammatical structure = abstract syntax + topic/focus articulation) $\Rightarrow$ surface structure)

FS Functional Grammar

- generates surface forms, their dependency types and the word order from underspecified meaning structures: Discourse Plans (DP) expressing speaker's intended view on predications,
- metaphor: "Generate-While-Planning-the-Meaning" : DP feasibility is determined through correctness of generated types and checked using an efficient type calculus

Output: dependency types

Continuous dependencies (projective):



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

Discontinuous dependencies (non-projective):

EX: *Elle le lui a donné*

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

Dependency types are 1st order signature types (like in pregroups) i.e. not raisable ! (unlike in Lambek grammars and CCG [M.Steedman])

Categorial dependency grammar

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

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

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

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

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.

$$\frac{[\#(\swarrow clit{-}iobj)]^{\swarrow clit\text{–}iobj}\;[\#(\swarrow clit{-}iobj)\backslash\#(\swarrow clit{-}dobj)\backslash pred\backslash S/aux]}{[\#(\swarrow clit{-}dobj)\backslash pred\backslash S/aux]^{\swarrow clit\text{–}iobj}}\;(\mathbf{L}^l)$$

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

$$\frac{[pred]\;[pred\backslash S/aux]^{\swarrow clit\text{–}dobj\,\swarrow clit\text{–}iobj}}{[S/aux]^{\swarrow clit\text{–}dobj\,\swarrow clit\text{–}iobj}}\;(\mathbf{L^l})$$

$$\frac{[S/aux]^{\swarrow clit\text{–}dobj\,\swarrow clit\text{–}iobj}\;[aux]^{\nwarrow clit\text{–}iobj\,\nwarrow clit\text{–}dobj}}{[S]^{\swarrow clit\text{–}dobj\,\swarrow clit\text{–}iobj\,\nwarrow clit\text{–}iobj\,\nwarrow clit\text{–}dobj}}\;(\mathbf{L}^r)$$

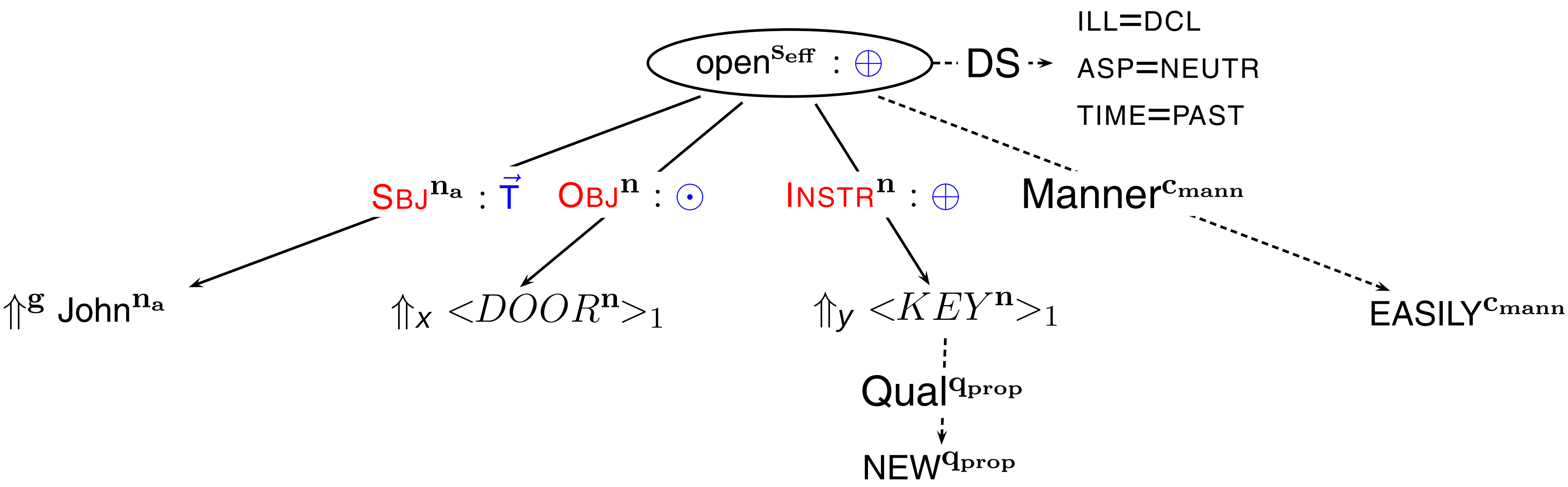
$$\frac{[S]^{\swarrow clit\text{–}dobj\,\swarrow clit\text{–}iobj\,\nwarrow clit\text{–}iobj\,\nwarrow clit\text{–}dobj}}{S}\;(\mathbf{D}^l \times 2)$$

Input: Discourse Plans

Influenced by: Functional Generative Description (functionality, communicative views), Frame semantics (multiple predications of one situation) and DRT (dynamic reference)

Ditransitive construction: *John easily opened the door with the new key*

Input DP π_1:



Output WO, forms and typing:

$John \mapsto [pred]$ $easily \mapsto [circ]$ $opened \mapsto [circ * \backslash pred \backslash S / instr-obj / dobj]$

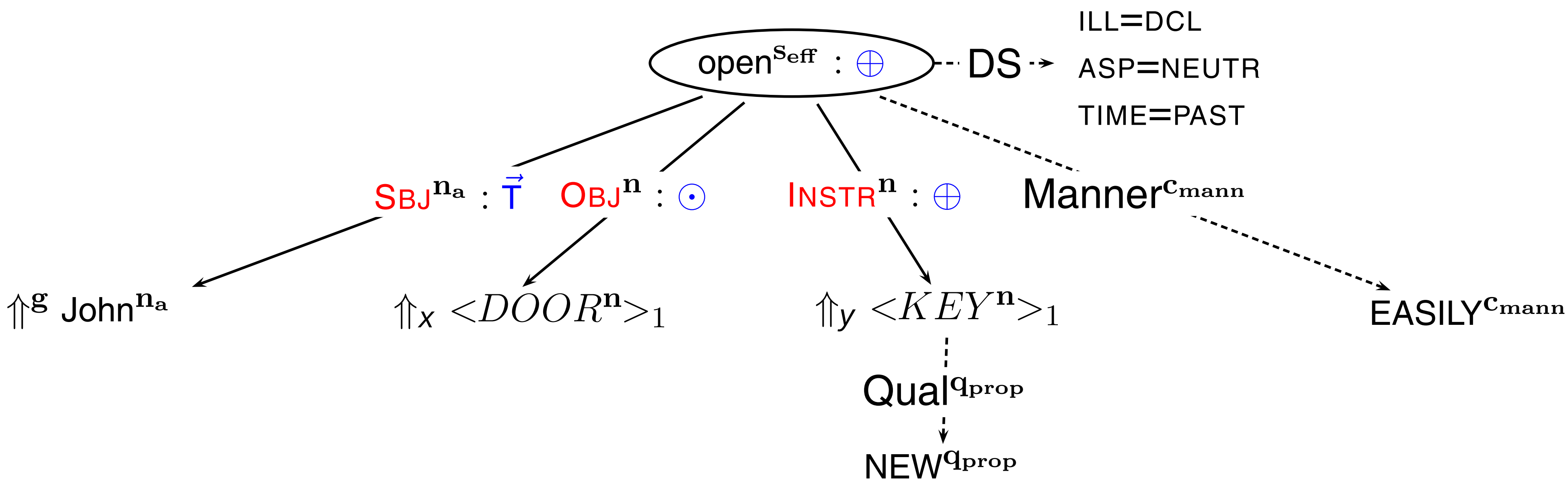
$the \mapsto [det]$ $door \mapsto [det \backslash dobj]$ $with \mapsto [instr-obj / prepos-with]$

$the \mapsto [det]$ $new \mapsto [modif]$ $key \mapsto [modif * \backslash det \backslash prepos-with]$

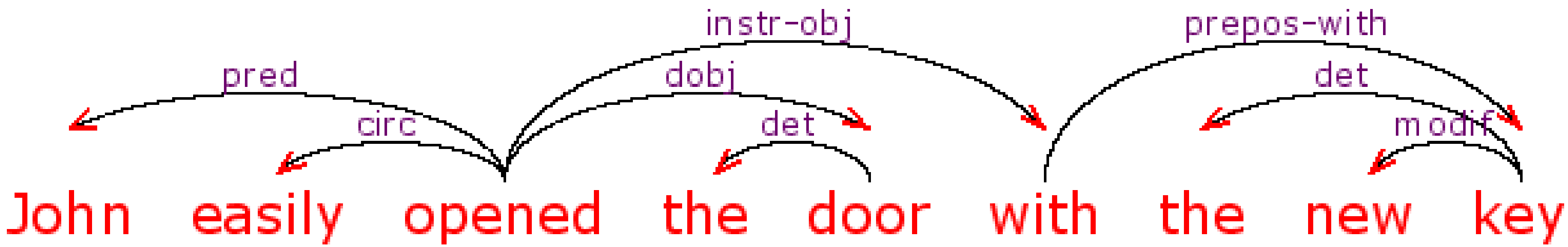
π_1 is feasible (the output typing is correct in this WO: $T(\pi_1) \vdash S$).

Generated dependency tree

Input DP π_1:



Output DT:

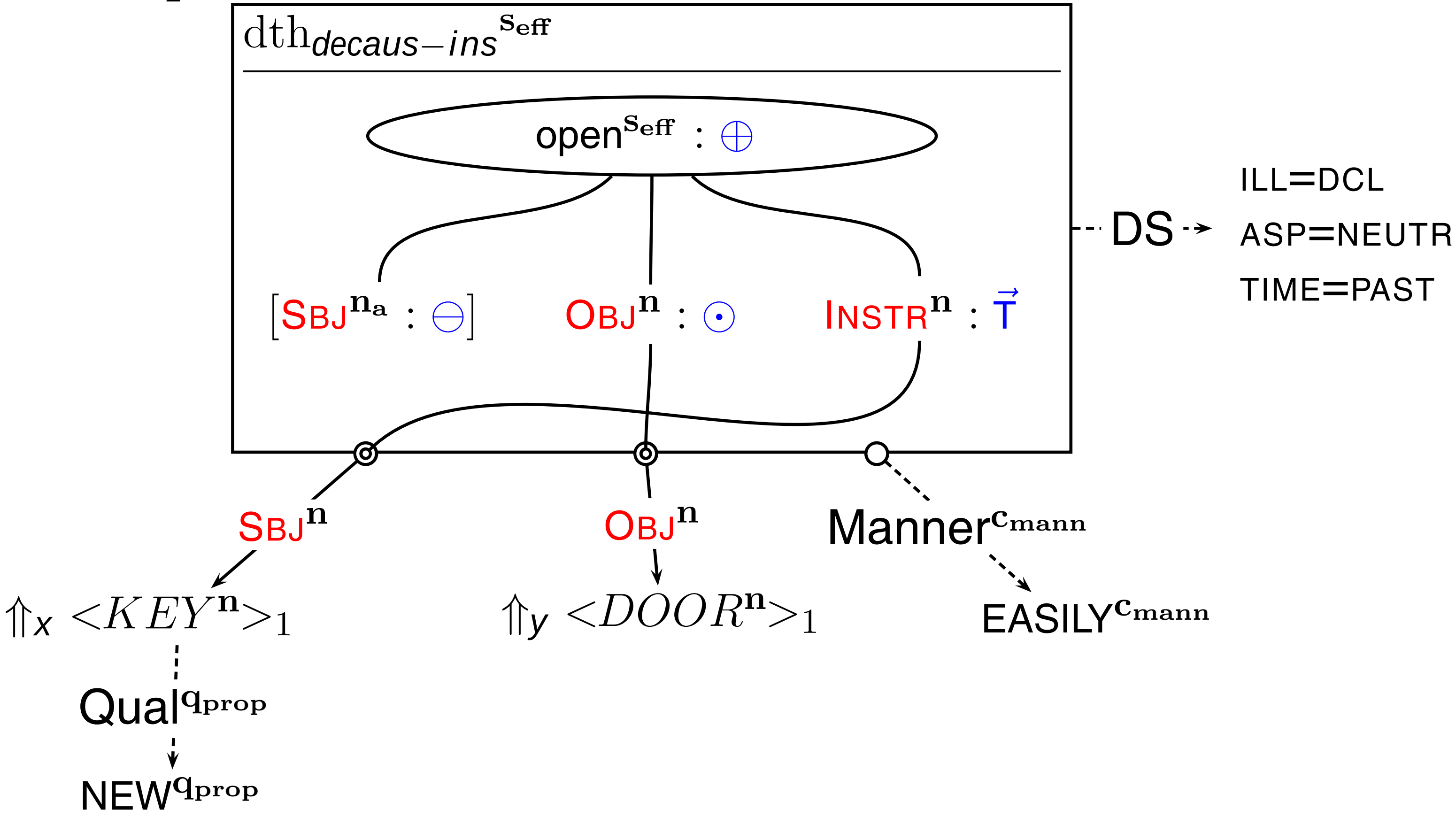


Decausative object alternation

EX: *The new key easily opened the door*

Diathesis: $\langle\langle \mathrm{dth}_{decaus-ins}(\bigstar_{\oplus}, \emptyset \longleftrightarrow \mathrm{SBJ}_{\ominus}, \mathrm{SBJ} \longleftrightarrow \mathrm{INSTR}_{\vec{T}}, \mathrm{OBJ} \longleftrightarrow \mathrm{OBJ}_{\odot})^{\mathbf{s_{eff}}} \rangle\rangle$

Input DP π_2:



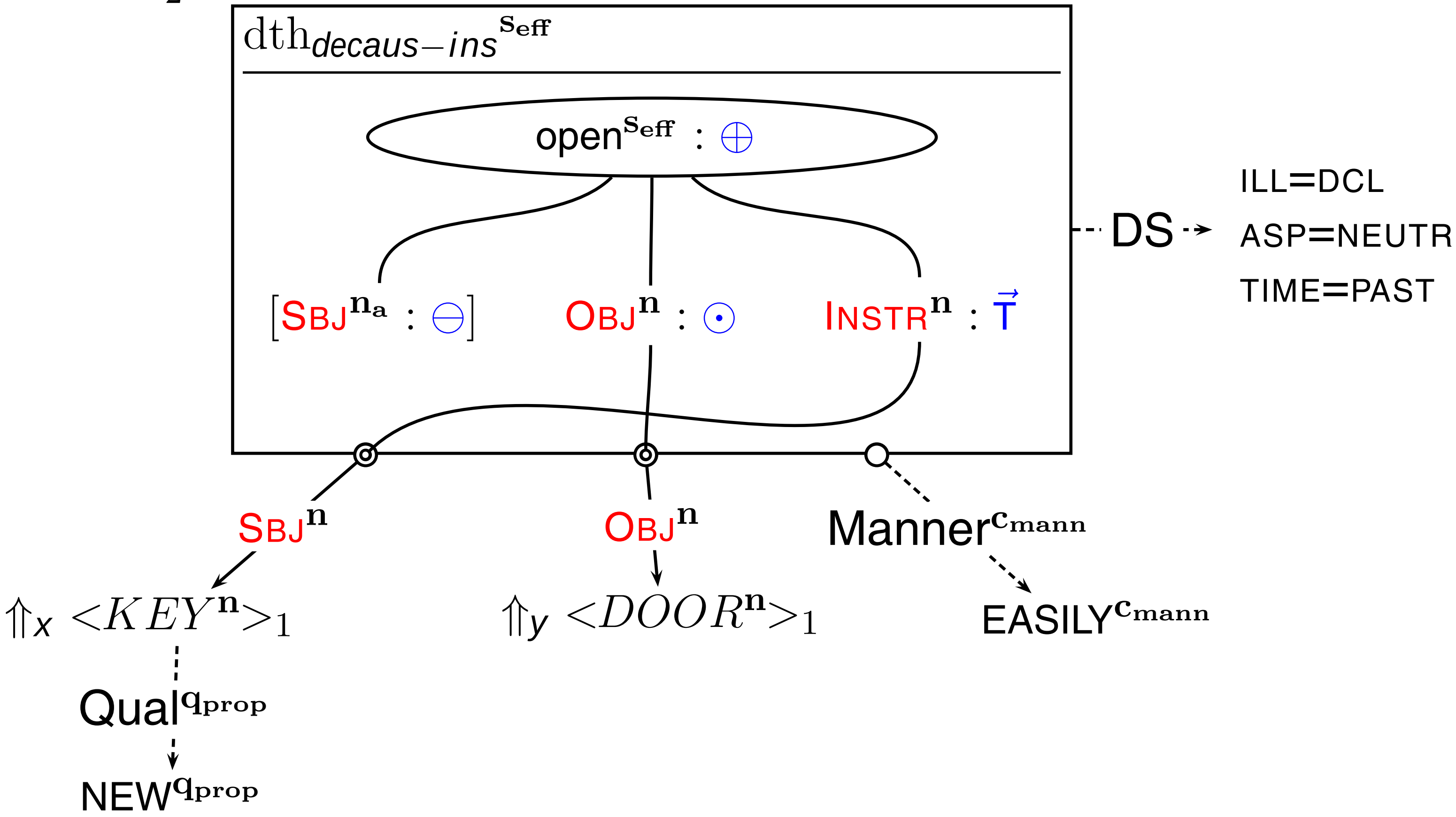
Output WO, forms and typing:

$\mathit{The} \mapsto [det]\ \mathit{new} \mapsto [modif]\ \mathit{key} \mapsto [modif * \backslash det \backslash pred]\ \mathit{easily} \mapsto [circ]$
$\mathit{open\,ed} \mapsto [circ * \backslash pred \backslash S/dobj]\ \mathit{the} \mapsto [det]\ \mathit{door} \mapsto [det \backslash dobj]$

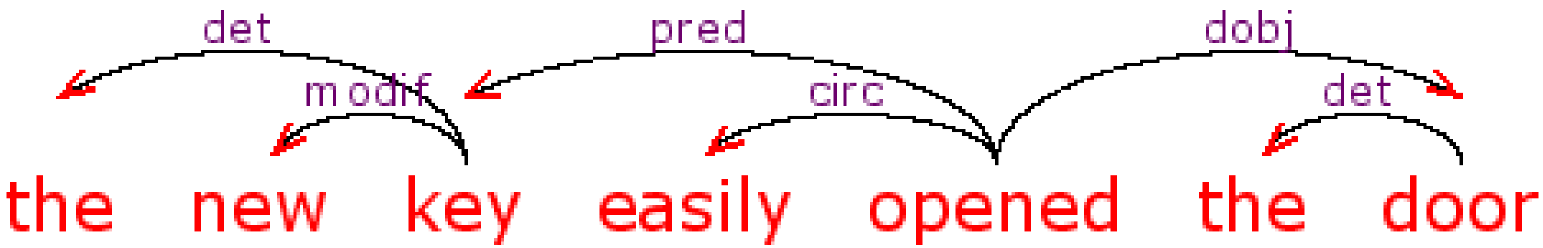
π_2 is also feasible (the output typing and WO are correct: $T(\pi_2) \vdash S$).

Generated dependency tree

Input DP π_2:



Output DT:



A mystery to unravel

QU: How can the WO be calculated from the DP? It often depends on arguments' scopes!

RE - At the extra-situational level, the DP are hierarchic (as the DRS).
- At the intra-situational level, the argument scopes (and the WO) are uniquely identified through the fundamental semantic prominence hierarchy of thematic roles (e.g., SBJ > OBJ > OBL) and through the intended perspective, i.e. the arguments' communicative ranks.

Scope principle [a]: More prominent args overscope less prominent args

The communicative ranks are ordered by the topicality order: $\vec{T}/O > \odot > \oplus$ Their canonical assignment to arguments matches their semantic prominence: SBJ:($\vec{T}$ / O), OBJ:$\odot$, OBL:$\oplus$.

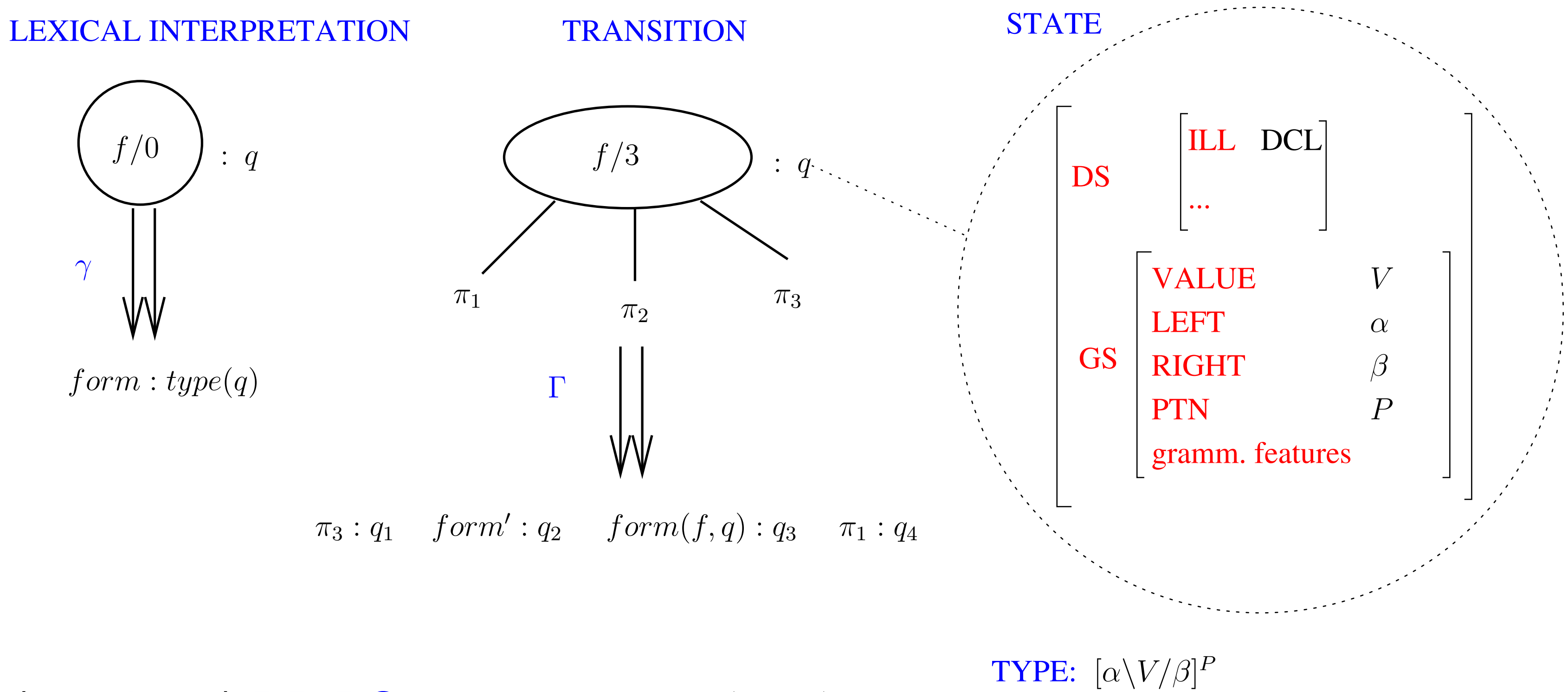
Different assignments of ranks may change (through the corresponding diathesis) the args' scopes and the WO

EX: The sentence *(Every linguist)*$_{SBJ}$ *knows (two languages)*$_{OBJ}$ may have only the $\forall\exists$ reading.

---

[a] The Scope Principle distinguishes the DP underspecification from that of the scope-abstracted underspecified semantics like Minimal Recursion Sem. [A.Copestake e al.], Hole semantics [J.Bos], etc.

DFS-transducers applied to DP



(top-down) DP-DS-converter $T = (\Gamma, G)$:
- G is a CDG; its lexicon λ assigns to words finite sets of types
- Γ is a DFS-transducer on DP coordinated with λ: the type $type(q) = \gamma(f, q)$ assigned by Γ is one of those assigned by G: $type(q) \in \lambda(f)$

Feasible DP π : that converted by T into a correctly typed sentence

$T(\pi) = f_1 \ldots f_n : t_1 \ldots t_n$ and $t_1 \ldots t_n \vdash_G S$. So $x = f_1 \ldots f_n \in L(G)$

and there is a DS D extracted from the proof (denoted $T(\pi, x, D)$)

Language of T : $L(T) = \{x \parallel \exists \pi, D\ (T(\pi, x, D))\}$

This is how it works:

For a DP π in a state q the unique applicable transition is determined through consecutive refinements of partial transitions unified with π and q

Consecutive transition refinements:

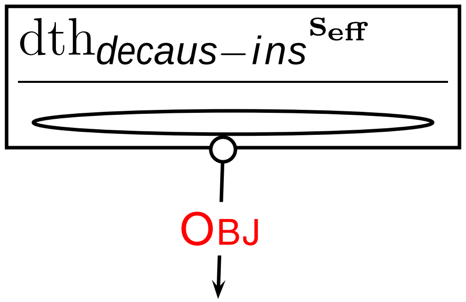
Partial transition $t_{decaus-ins}$:

$$\frac{\mathrm{dth}_{decaus-ins}{}^{\mathbf{s}_{\mathbf{eff}}}}{\qquad}$$

$\Rightarrow$

First transition refinement

Partial transition $t_{decaus-ins} \circ t_{(dcl,obj)}$:

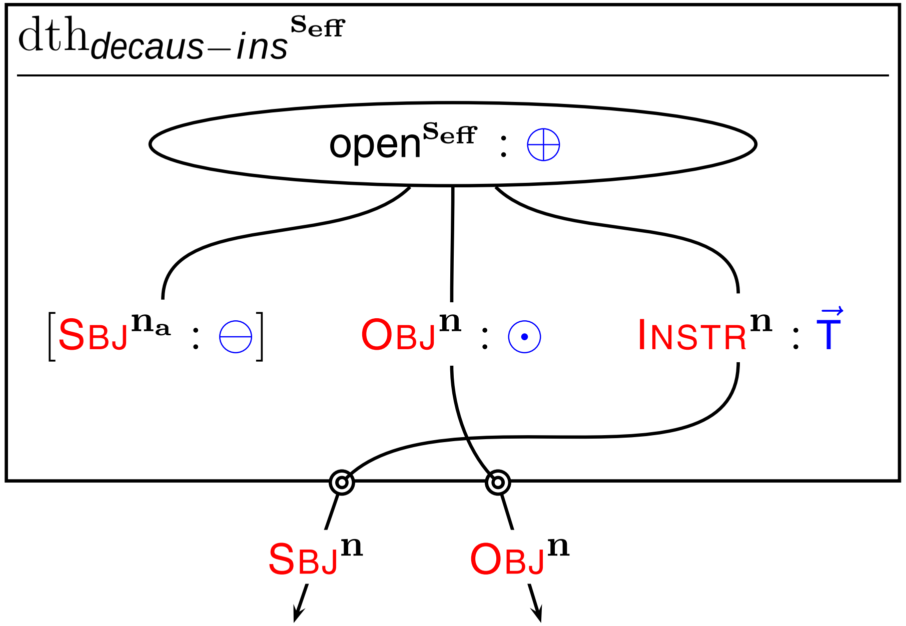


$[\text{DS} \; [\text{ILL} \; Dcl]]$

$\Rightarrow$ HEAD $[\text{GS} \; [\text{VALUE} \; S]]$

Second transition refinement

Partial transition $t_{decaus-ins} \circ t_{(dcl,obj)} \circ t_{PN}$:



$$\left[\text{DS} \left[\begin{array}{ll} \text{ILL} & Dcl \\ \text{ASP} & Neutr \\ \text{TIME} & Past \end{array} \right] \right]$$

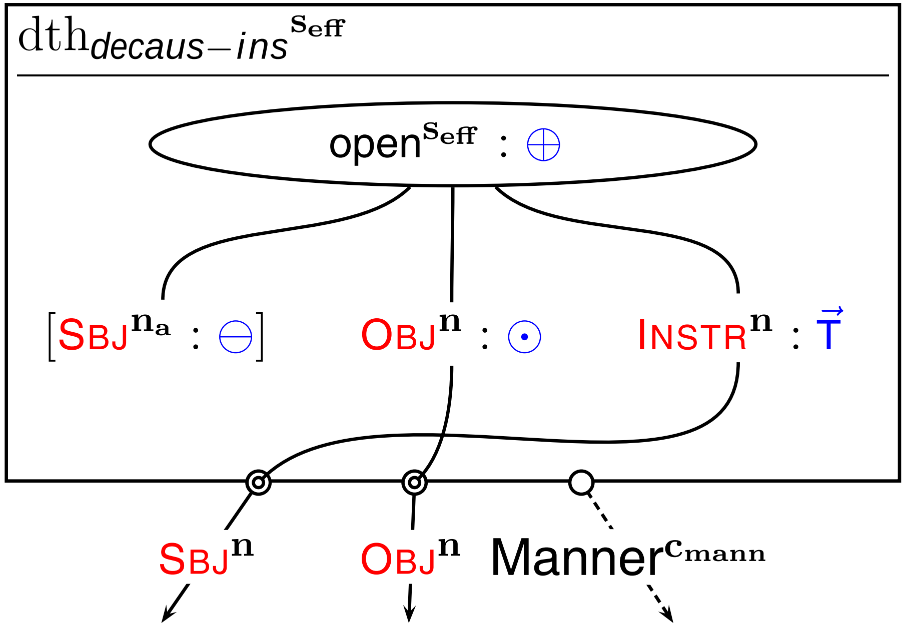
⇒ SBJ [DS [VALUE *pred*]] $<$

$$(form(\text{OPEN}, f) = \textit{OPENED})\ f : \left[\begin{array}{ll} \text{LEFT} & circ * \backslash pred \backslash \\ \text{RIGHT} & /circ * /dobj \\ \text{VALUE} & S \\ \text{TNS} & PastSimple \\ \text{GRCL} & V_t \end{array} \right] <$$

OBJ [GS [VALUE *dobj*]]

Last refined transition t_1

Transition $t_1 = t_{decaus-ins} \circ t_{(dcl,obj)} \circ t_{PN} \circ t_{(PN,\mathbf{c_{mann}})}$:



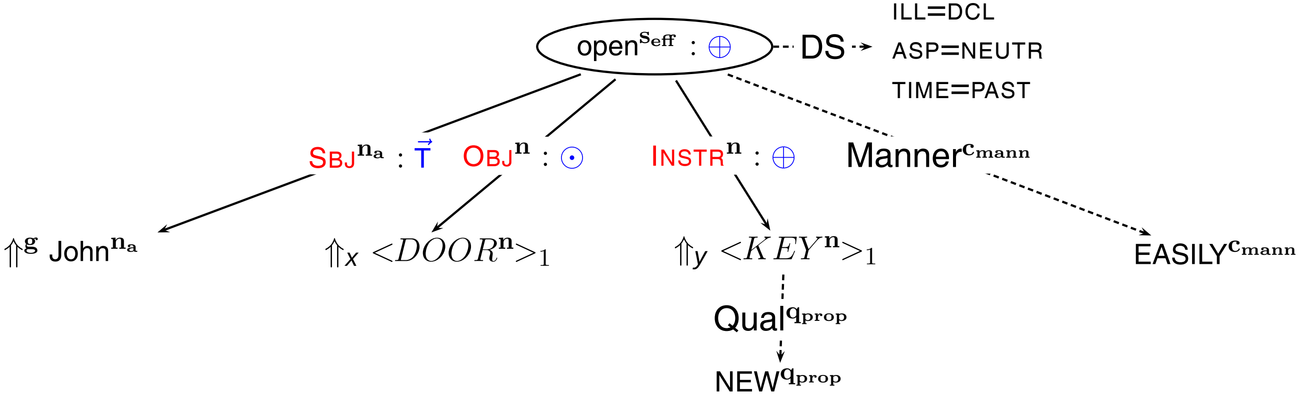
$$\left[\text{DS} \begin{bmatrix} \text{ILL} & Dcl \\ \text{ASP} & Neutr \\ \text{TIME} & Past \end{bmatrix} \right]$$

$\Rightarrow$

SBJ [DS [VALUE $pred$]] $<$ MANNER [GS [VALUE $circ$]] $<$

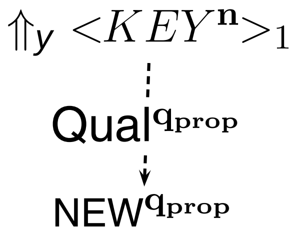
OPENED : $[circ * \backslash pred \backslash S / circ * / dobj]$ $<$ OBJ [GS [VALUE $dobj$]]

Transition application



$\left[\text{DS}\left[\begin{array}{ll}\text{ILL} & Dcl\\ \text{ASP} & Neutr\\ \text{TIME} & Past\end{array}\right]\right]$

$\overset{t_1}{\Longrightarrow}$



$[\text{DS}\ [\text{VALUE}\ pred]] < \textit{EASILY} : [\text{GS}\ [\text{VALUE}\ dobj]] <$

$\textit{OPENED} : [circ{*}\backslash pred\backslash S/circ{*}/dobj] < \Uparrow_x <DOOR^{\mathbf{n}}>_1 [\text{GS}\ [\text{VALUE}\ dobj]]$

Expressiveness, complexity

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

Projective dependency 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.

CR For every sequence of types γ with distinguished constituents (i.e. without elimination ambiguity), the provability $\gamma \vdash S$ is resolved in real time $|\gamma|$ using one stack and v counters, where v is the number of polarized valency names.

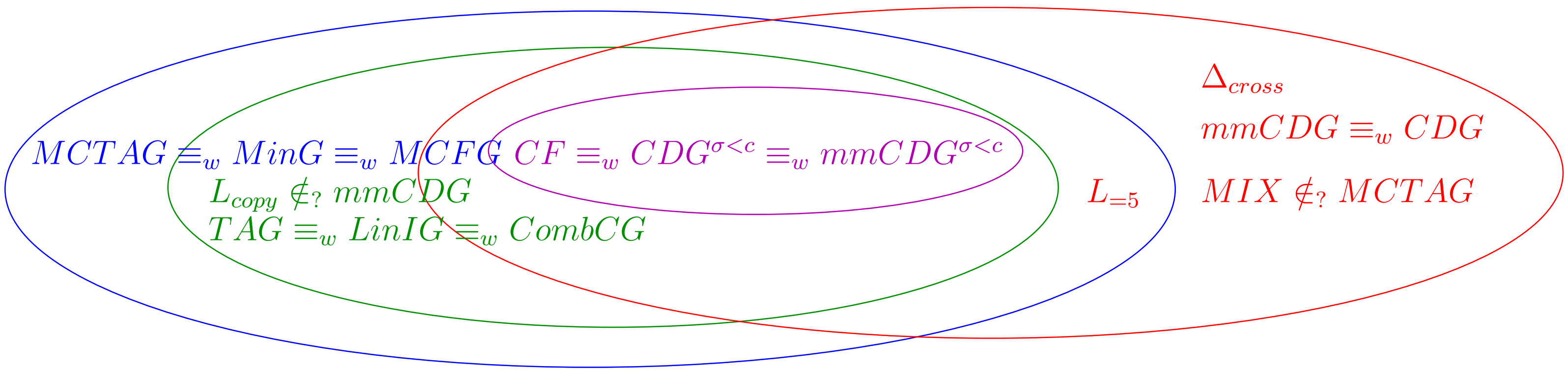
Expressiveness, complexity continued

TH The family of languages $T(L)$, where L is a regular f-tree language and T is an FT-DS converter, coincides with $\mathcal{L}(gCDG)$.

RM The f-tree language of DP is regular

Parsing complexity (Dekhtyar,Dikovsky'2004) :

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



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

Conclusion

- FS-Functional Grammar calculates a lexicalization, d-types and a WO using a deterministic top-down FS-transducer applied to feasible discourse plans expressing the speaker's view on predications
- In the place of a general feasibility proof, the FS-FG uses a real-time proof scheme check procedure
- In principle, the DP composition, its FS-conversion and the proof checking can be synchronized implementing in this way the metaphor

  Generate-While-Planning-the-Meaning

A gCDG for a non-CF language

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

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:

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

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

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

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

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

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

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

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

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

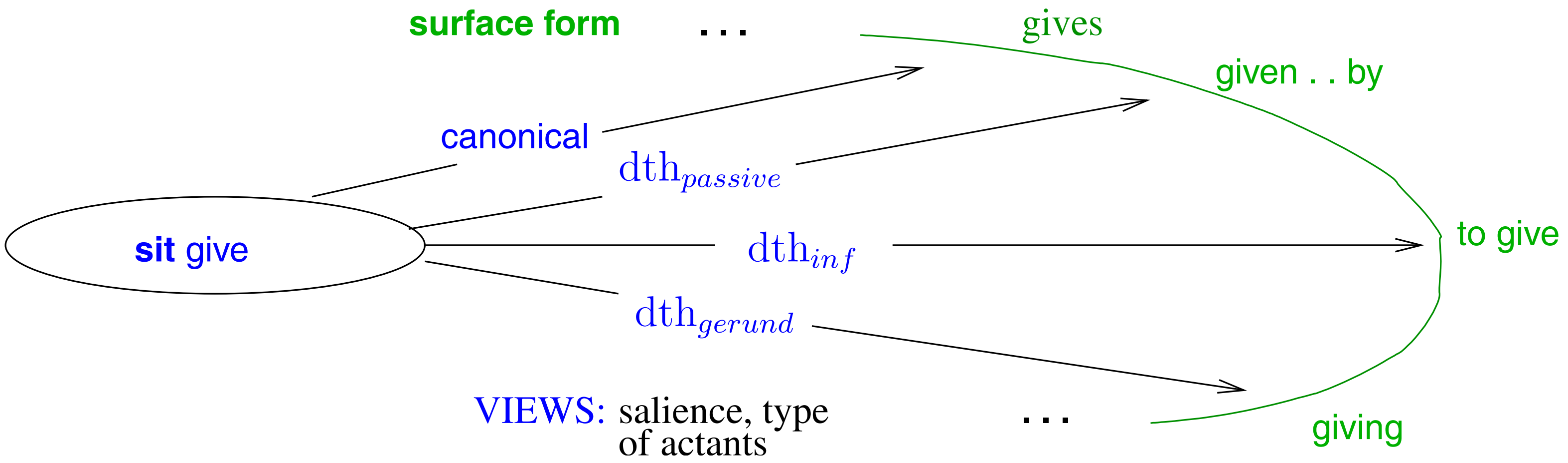
$$\frac{[S]^{\swarrow A\swarrow A\swarrow A\nwarrow A\nwarrow A\nwarrow A}}{S}\ (\mathbf{D}^l \times 3)$$

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)

A case of clitisation in French

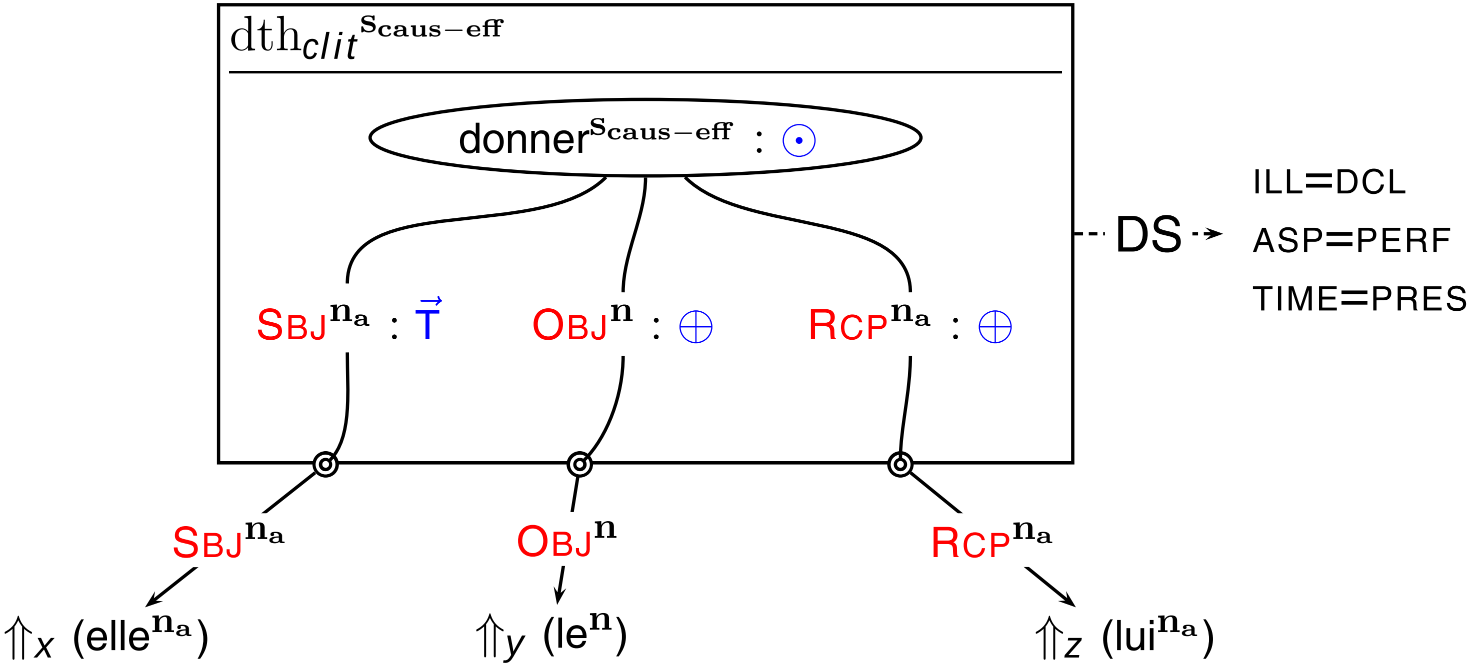
EX: *Elle le lui a donné*

Situation $\langle\!\langle donner \rangle\!\rangle$.

Canonical profile: $\langle\!\langle donner\left(\text{SBJ}^{\mathbf{n_a}}, \text{OBJ}^{\mathbf{n}}, \text{RCP}^{\mathbf{n_a}}\right)^{\mathbf{S_{caus-eff}}} \rangle\!\rangle$

Diathesis: $\langle\!\langle \mathrm{dth}_{clit}\left(\bigstar_{\odot}, \text{OBJ} \hookleftarrow \text{OBJ}_{\oplus}\text{RCP} \hookleftarrow \text{RCP}_{\oplus},\right)^{\mathbf{S_{caus-eff}}} \rangle\!\rangle$

Input DP π_3:



Generated dependency tree

Output WO, forms and typing:

$$elle \mapsto [pred]$$

$$le \mapsto [\#(\swarrow clit-dobj)]^{\swarrow \textsf{clit–dobj}}$$

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

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

$$donn\acute{e} \mapsto [aux]^{\nwarrow \textsf{clit–iobj} \nwarrow \textsf{clit–dobj}}$$

π_3 is feasible: $T(\pi_3) \vdash S$).

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

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

$$\frac{[pred]\ \ [pred\backslash S/aux]^{\swarrow \textsf{clit–dobj} \swarrow \textsf{clit–iobj}}}{[S/aux]^{\swarrow \textsf{clit–dobj} \swarrow \textsf{clit–iobj}}}\ (\mathbf{L}^l)$$

$$\frac{[S/aux]^{\swarrow \textsf{clit–dobj} \swarrow \textsf{clit–iobj}}\ \ [aux]^{\nwarrow \textsf{clit–iobj} \nwarrow \textsf{clit–dobj}}}{[S]^{\swarrow \textsf{clit–dobj} \swarrow \textsf{clit–iobj} \nwarrow \textsf{clit–iobj} \nwarrow \textsf{clit–dobj}}}\ (\mathbf{L}^r)$$

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

One more example

EX: *la commission ne la lui a pas refusée*

Situation $\langle\langle refuser \rangle\rangle$.

Canonical profile: $\langle\langle refuser(\text{SBJ}^{\mathbf{n_a}}, \text{OBJ}^{\mathbf{s}}, \text{RCP}^{\mathbf{n_a}})^{\mathbf{s_{eff}}} \rangle\rangle$

Diathesis: $\langle\langle \mathrm{dth}_{clit}\,(\bigstar_{\odot}, \text{OBJ} \hookleftarrow \text{OBJ}_{\oplus}\text{RCP} \hookleftarrow \text{RCP}_{\oplus},)^{\mathbf{s_{caus-eff}}} \rangle\rangle$

Input DP π_3: