Architecture compositionnelle pour les dépendances croisées

Alexandre Dikovsky

LINA CNRS FRE 2729, Université de Nantes

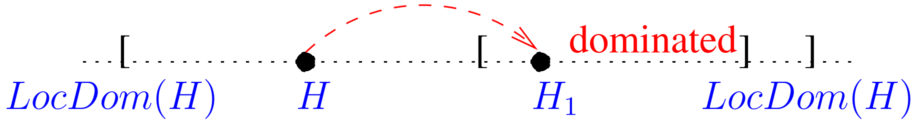
PLAN

- 1. Positionnement
- 2. Dépendances et l'ordre des mots
- 3. Grammaires catégorielles de dépendances
- 4. CDG multimodales
- 5. Expressivité / Complexité
- 6. Conclusion

Positionnement

Syntaxe de surface : **Dominance** / **Centralisation** / **Précédence**

Ordre canonique



garantit une sémantique compositionnelle

Déformations de l'OM canonique sont :

omniprésentes, signifiantes, complexes

Grammaires qui les expriment (e.g., GUnif, HPSG, GTL multimodales etc.)

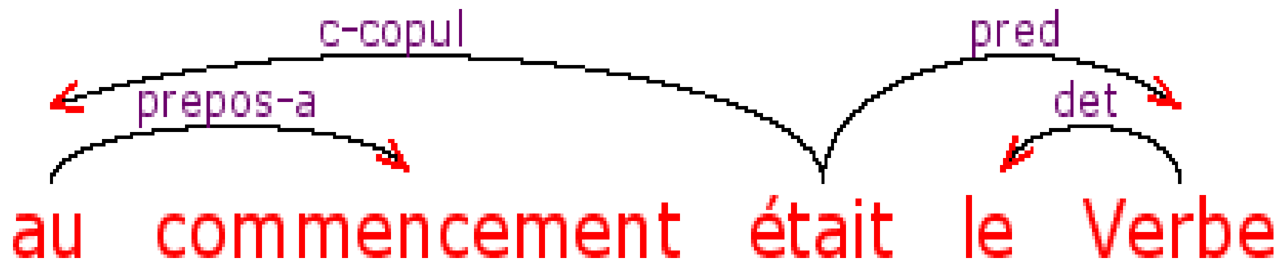
- soit ignorent la sémantique,
- soit sont excessivement complexes.

Objectif : calcul élémentaire de valences syntaxiques avec des principes généraux de saturation des valences à la place de règles :

- *suffisant pour exprimer les déformations de l'OM,*
- *avec une sémantique compositionnelle,*
- *avec une analyse efficace.*

Choix de syntaxe : structures et grammaires de dépendances

Structures de dépendances (de surface)



Dépendance ($\longrightarrow$) : Gouverneur $\xrightarrow{d}$ Subordonné

EX : était $\xrightarrow{pred}$ Verbe

Précédance ($<$) : OM (total)

Gouvernement ($\longrightarrow^*$) : $\longrightarrow^* = \longrightarrow^{(refl,trans)}$

Projection d'un mot w :

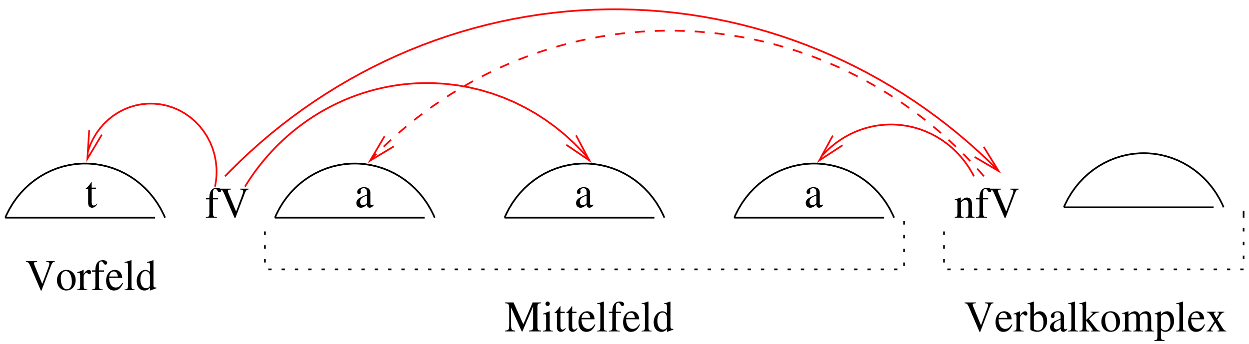
$$d(w) = \{\{w' \parallel w \longrightarrow^* w'\}\} \text{ ordonné par l'OM}$$

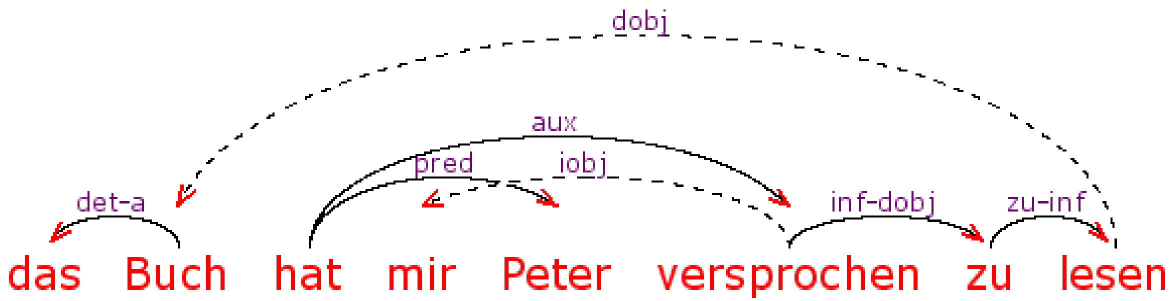
AD D projectif : $\{d(w) \parallel w \in D\}$ est une SC

Non-projectivités de l'OM

français (clitiques) :

allemand (scrambling) :

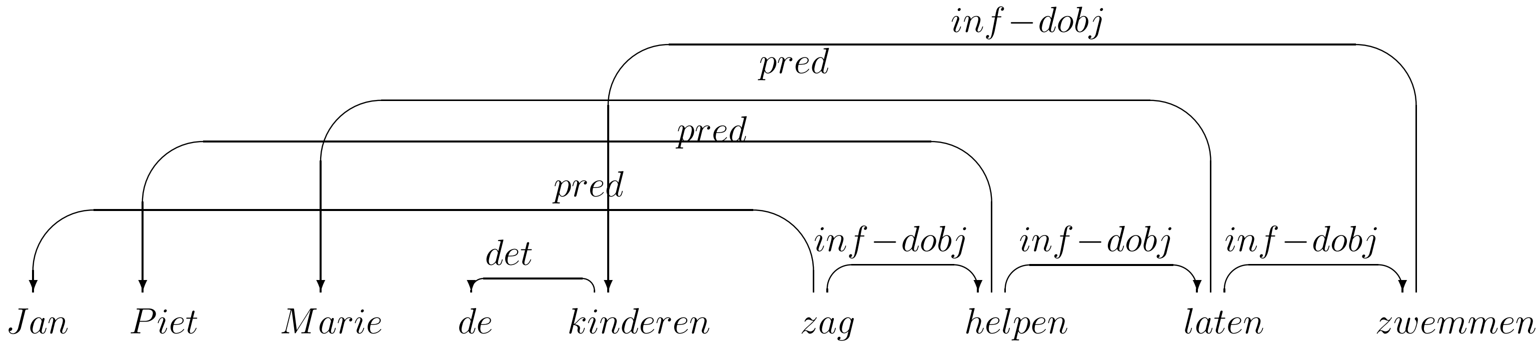




bulgare, russe : OM quasi libre dans la clause principale

Dépendances croisées non limitées

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



**Jan Piet Marie les enfants a vu aider faire nager*

Les phrases : $n_1 n_2 \ldots n_m n_{m+1} v_1 v_{(inf)2} \ldots v_{(inf)m}$, où il y a :

1. une dépendance prédicative $n_1 \xleftarrow{pred} v_1$ entre le verbe v_1 en forme finie et le nom n_1,

2. les dépendances prédicatives $n_i \xleftarrow{pred} v_{(inf)i}$ entre les verbes $v_{(inf)i}$ à l'infinitif et les noms n_i, ($2 \leq i \leq m$),

3. éventuellement, une dépendance d'objet direct

$$n_{m+1} \xleftarrow{dobj} v_{(inf)m}$$

si le verbe $v_{(inf)m}$ est transitif et le nom n_{m+1} est présent.

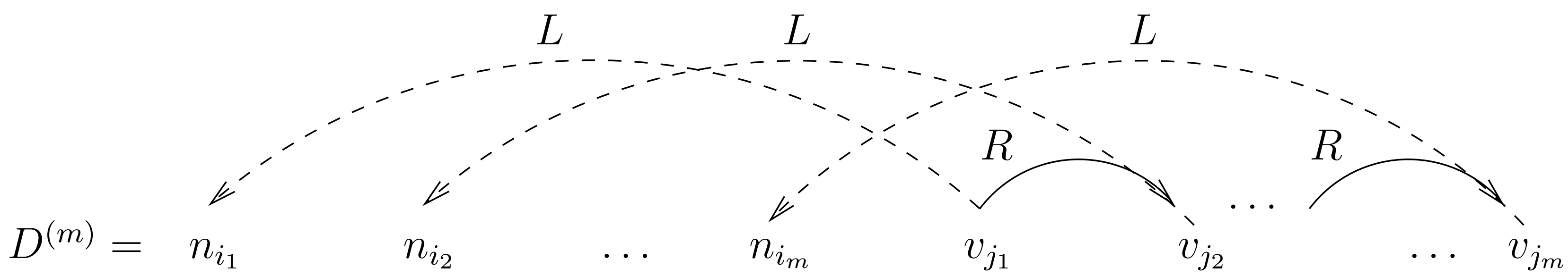
Dépendances croisées, suite

REMARQUE : Le modèle de cette construction n'est point le langage

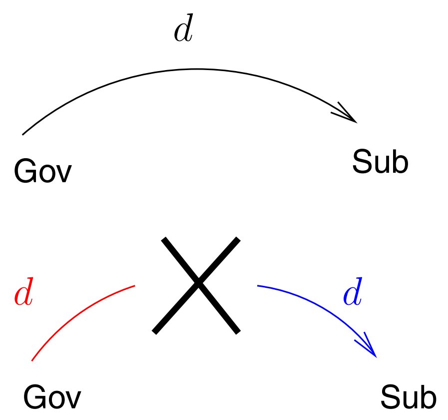
$$L_{copy} = \{xx \parallel x \in W^*\},$$

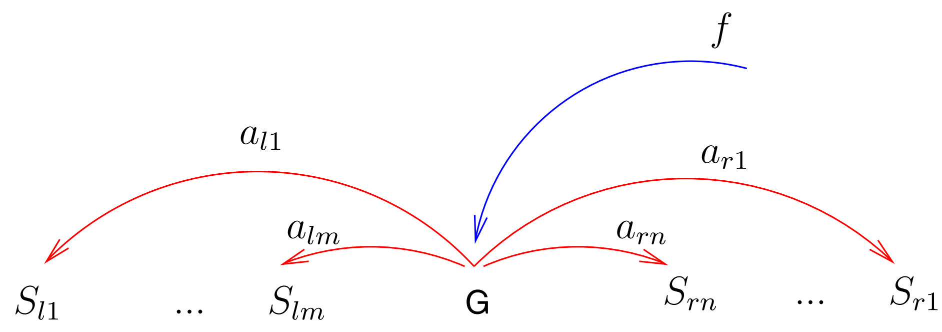
mais le langage d'AD $\Delta_{cross} = \{D^{(m)} \parallel m > 0\}$,

où $D^{(m)}$ est l'AD suivant ($n_{i_l} \in N, v_{j_r} \in V$) :

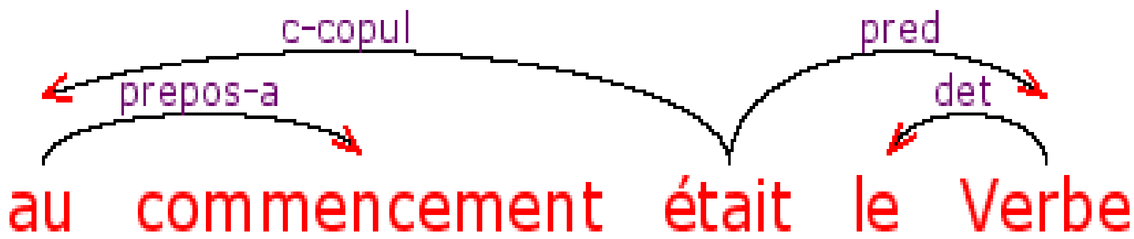


Types de dépendances projectives





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



au $\mapsto [c{-}copul/prepos{-}a]$

commencement $\mapsto [prepos{-}a]$

était $\mapsto [c{-}copul\backslash S/pred]$

le $\mapsto [det]$

Verbe $\mapsto [det\backslash pred]$

Types de dépendances non-projectives

Valences polarisées :

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

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

EX :

la $\xleftarrow{clit-dobj}$ donnée

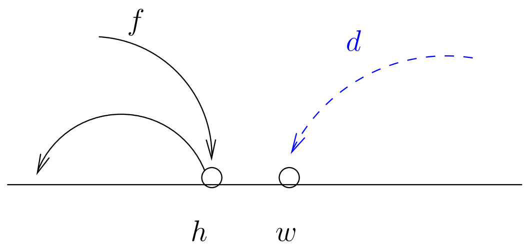
la $\mapsto [\varepsilon]^{\swarrow clit\text{–}dobj}$

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

Valences ancrées (pour placer les subordonnés distants) :

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

w s'ancre sur h : $w \mapsto [\#^r(\swarrow d)/\alpha]^{\swarrow d}$



Clitiques en français

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

Grammaire catégorielle de dépendances

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

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

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

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

si $(\swarrow C)P(\nwarrow C)$ satisfait le principe suivant d'appariement de valences :

Principe **FA** : v est saturée dans $P_1vP\breve{v}P_2$ par la plus proche valence disponible $\breve{v}$ (i.e. P n'a pas d'occurrence de $v, \breve{v}$)

Elimination des valences duales $v = (\swarrow C)$, $\breve{v} = (\nwarrow C)$ avec la règle $\mathbf{D^l}$ crée la dépendance discontinue C.

Elimination d'une valence ancrée $\#(v)$ avec les règles $\mathbf{L^l}, \mathbf{I^l}$ ne crée pas de dépendance.

Preuves de correction d'affectation des types

EX : *elle la lui a donnée*

Sans ancrage :

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

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

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

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

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

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

Avec ancrage :

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

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

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

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

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

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

Architecture multimodale pour les CDG

PB : FA n'est pas suffisant pour les dépendances croisées

CDG multi-modales (mmCDG) : toute dépendance peut avoir son propre mode de compositionnalité :

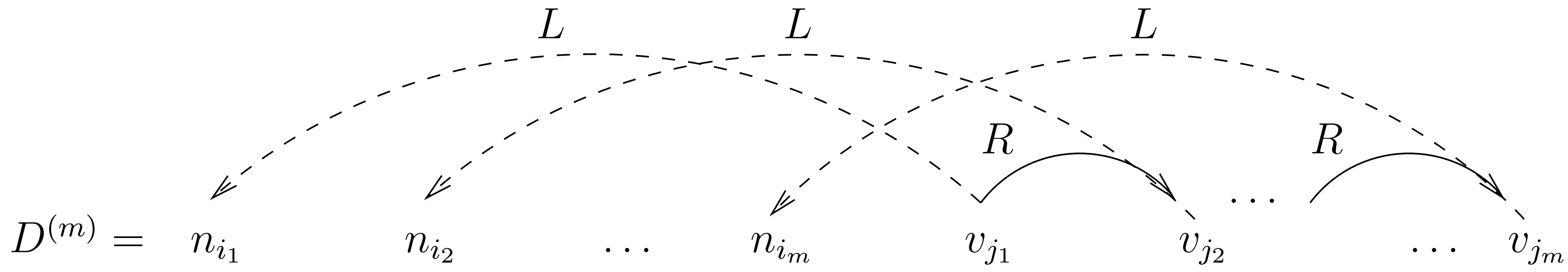
- variations des positions des arguments pour les dépendances projectives,
- principes d'appariement de valences pour les dépendances non-projectives

Pour le néerlandais, le principe d'appariement pour la dépendance prédicative des verbes à l'infinitif est :

Principe **FC** : dans $P_1(\swarrow C)P(\nwarrow C)P_2$, P_1 n'a pas d'occurrence de $\swarrow C$ et P n'a pas d'occurrence de $\nwarrow C$ (i.e. v est saturée par la première valence croisée disponible $\breve{v}$)

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

Dépendances croisées : une solution



Δ_{cross} est généré par $mmCDG^{\mathbf{FC}}$ suivante :

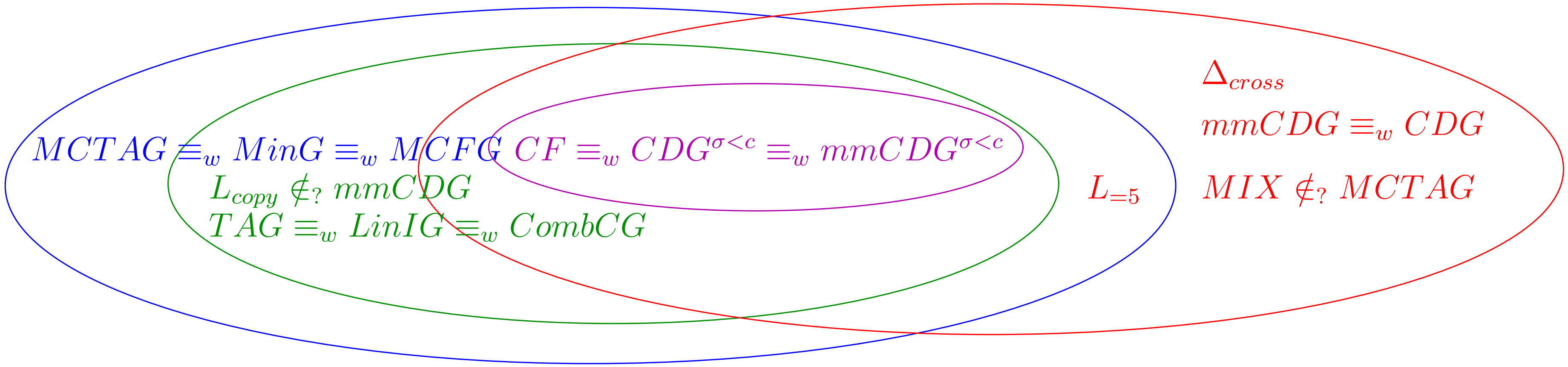
$$G_{cross} = \begin{cases} n \mapsto [\#(\swarrow L)]^{\swarrow L}, [\#(\swarrow L)\backslash\#(\swarrow L)]^{\swarrow L}, & \text{pour } n \in N \\ v \mapsto [\#(\swarrow L)\backslash S/R]^{\nwarrow L}, [R/R]^{\nwarrow L}, [R]^{\nwarrow L}, & \text{pour } v \in V \end{cases}$$

E.g., $D^{(3)} \in \Delta_{cross}$ grâce à la preuve :

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

Expressivité / complexité

Déficit de valence $\sigma(G)$: taille maximum des potentiels dans les preuves de G.



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

Complexité d'analyse syntaxique (Dekhtyar,Dikovsky'2004) :

1. théorique : $\mathbf{O}(\mathsf{n}^{3+2\mathsf{p}})$ (pour p valences polarisées)
2. en pratique ($\sigma(G) < const$) : $\mathbf{O}(\mathsf{n}^3)$.

Conclusion

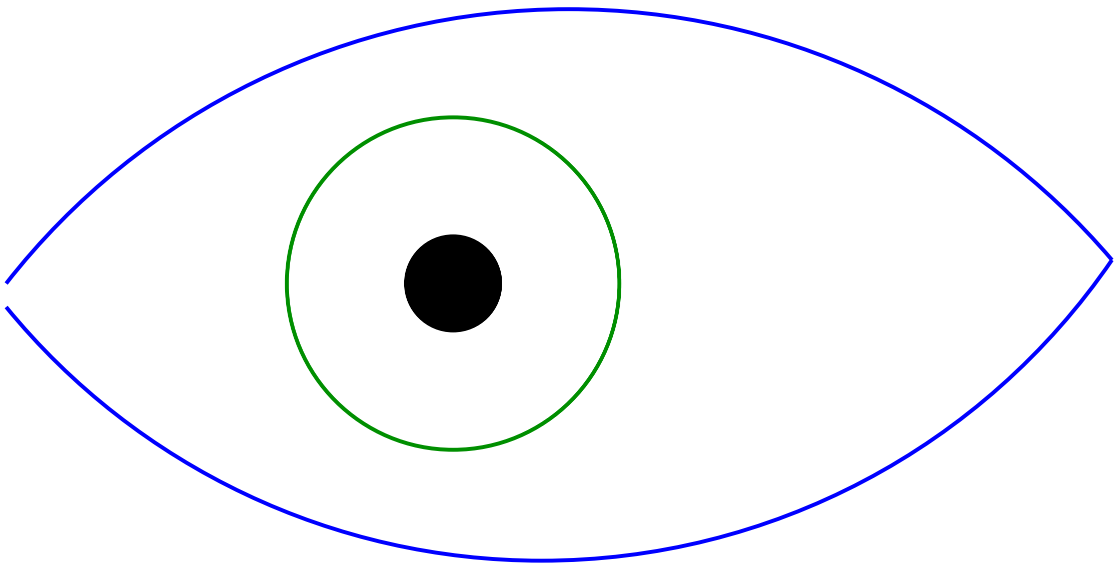
Les CDG multimodales :

1. expriment les dépendances non bornées et l'ordre flexible
2. disposent d'un algorithme universel d'analyse en temps $O(n^3)$ en pratique,
3. supportent les sémantiques compositionnelles,
4. sont simples à employer

Travail en cours :

1. passage à l'échelle (lexiques des corpus, morphologie, BD lexicale),
2. analyseur incrémental entraîné sur les corpus,
3. sémantique logique dynamique.

Démonstration de l'analyseur

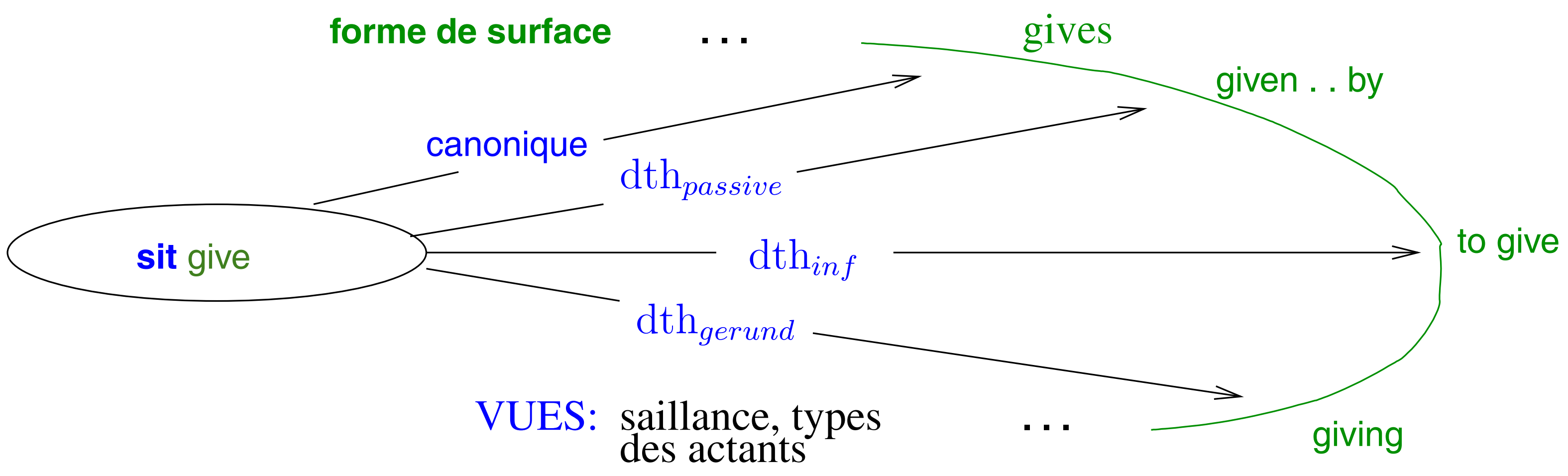


Sources sémantiques des dépendances

- Dépendances proviennent des structures sémantiques sous spécifiées (applées plans discoursifs (DP)) qui sont les arbres de composition des situations ou des diathèses sémantiques des situations
- Types des dépendances peuvent être calculés à partir de DP par un transducteur fini descendant

Situations, diathèses

Situations : invariants des vues communicatives (= diathèses sémantiques)



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

Bei Platae siegten die Griechen über die Perser

Bei Platae wurden die Perser von den Griechen besiegt

situation commune : siegen(SBJ, OBJ)

- Arguments des situations :

  - actants, identifiés par ses rôles thématiques, varient d'une diathèse à l'autre selon leurs rangs communicatifs envisagés;

  - circonstants, identifiés par des attributs, sont optionnels

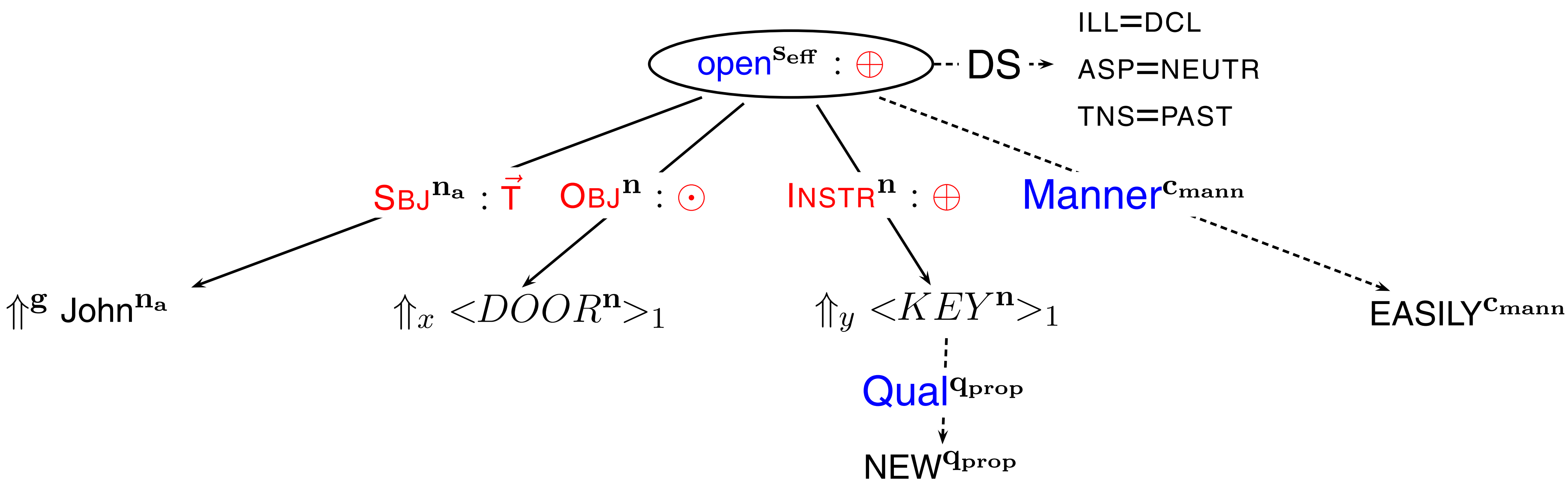
- rangs : $\vec{T}$ (topique $\simeq$, O (topique impliqué 0), $\odot$ (focus rhématique), $\oplus$ (fond), $\ominus$ (périphérie)

⟪open⟫ : profil canonique

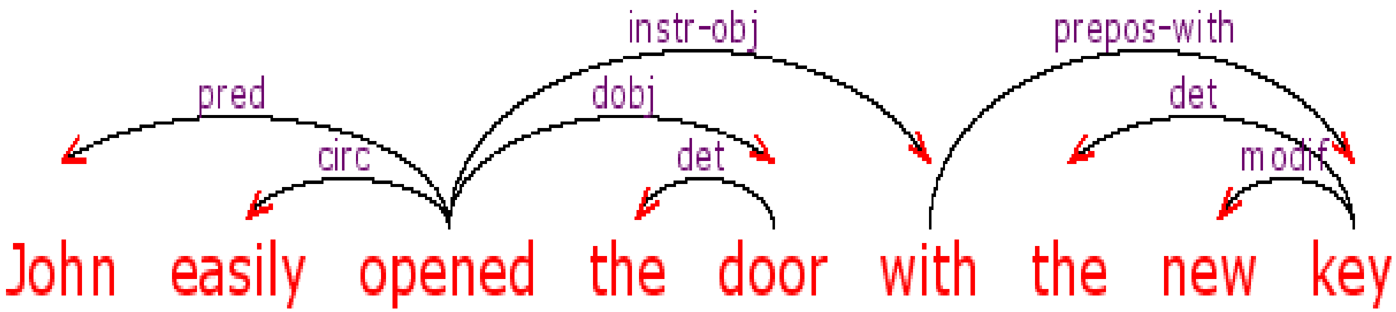
EX: *John easily opened the door with the new key*

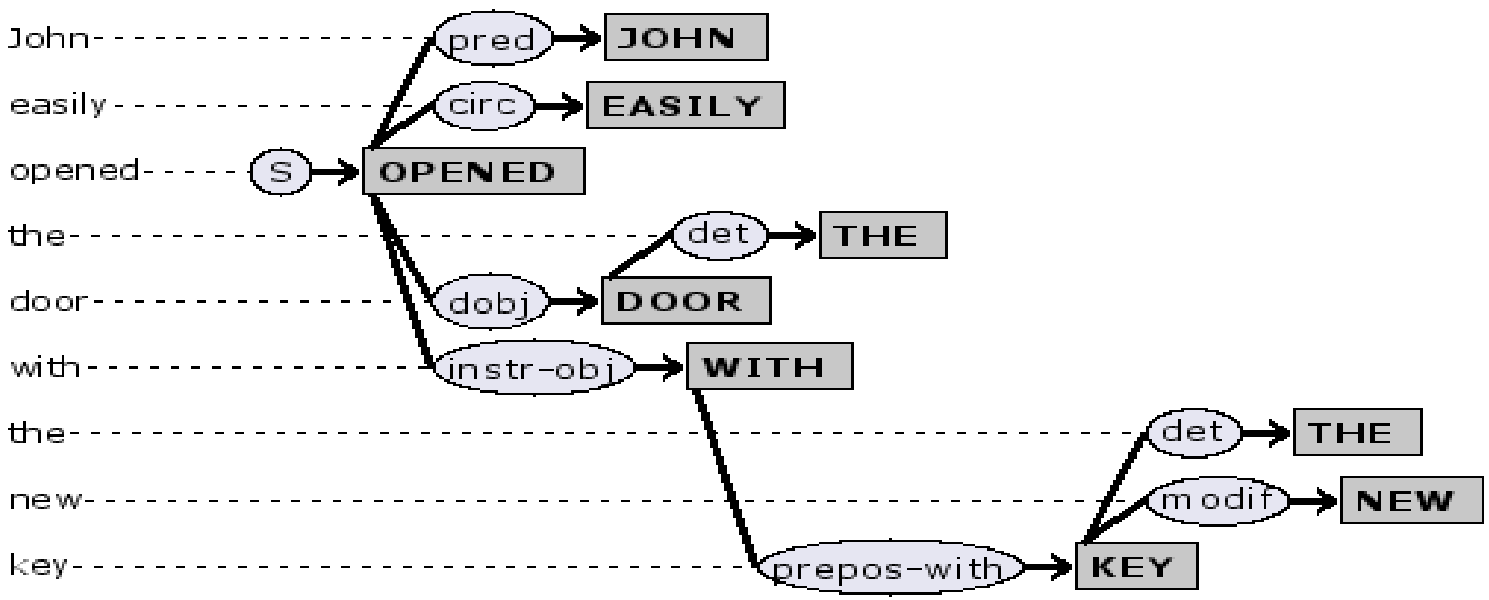
Situation ⟪open⟫. Profil canonique :

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



AD : profil canonique

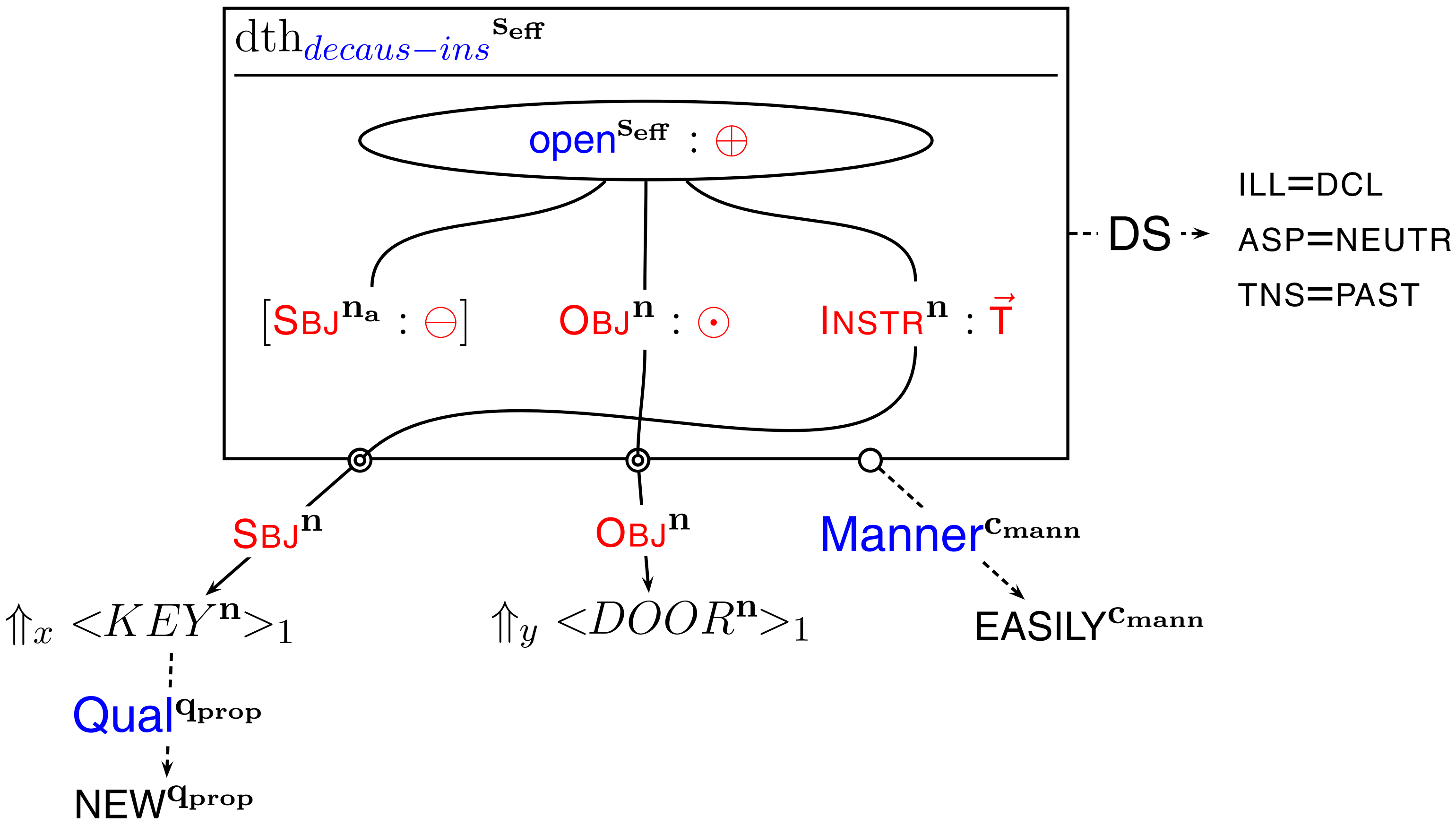




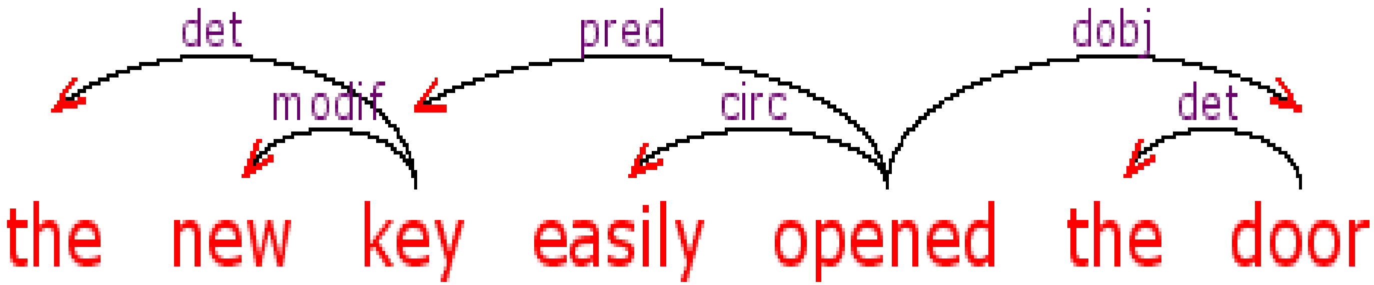
⟪open⟫ : $1^{ère}$ diathèse de décausation

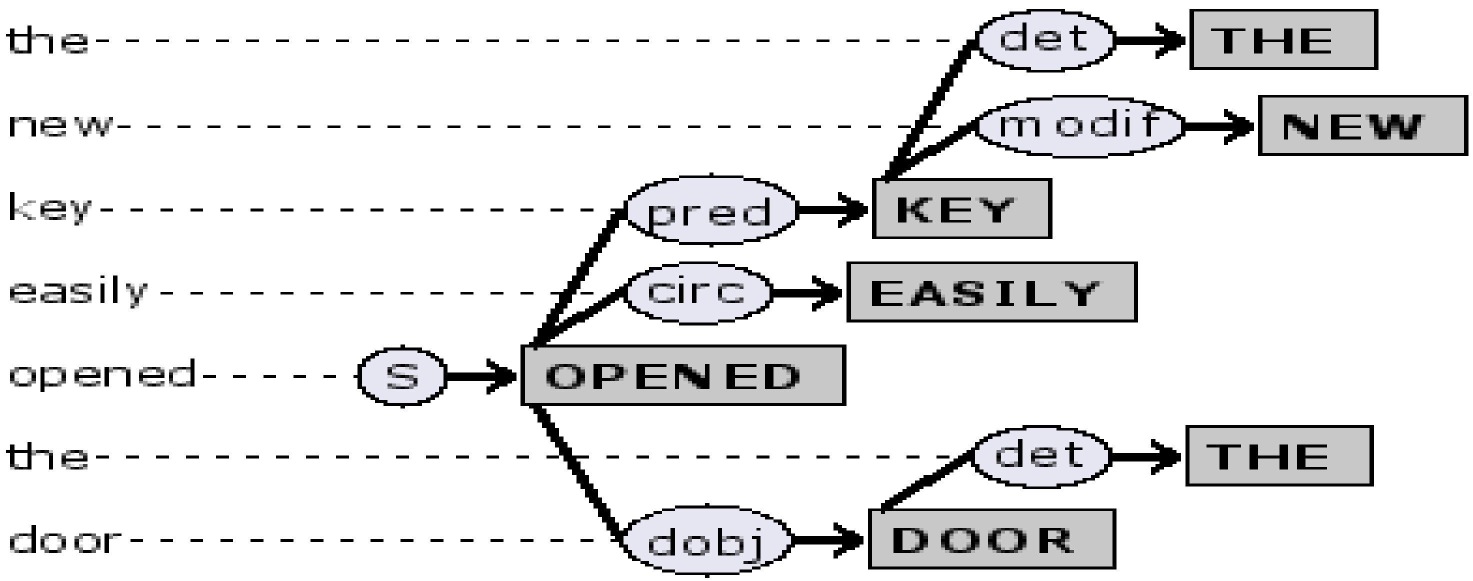
EX: *The new key easily opened the door*

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



AD : $1^{ère}$ diathèse

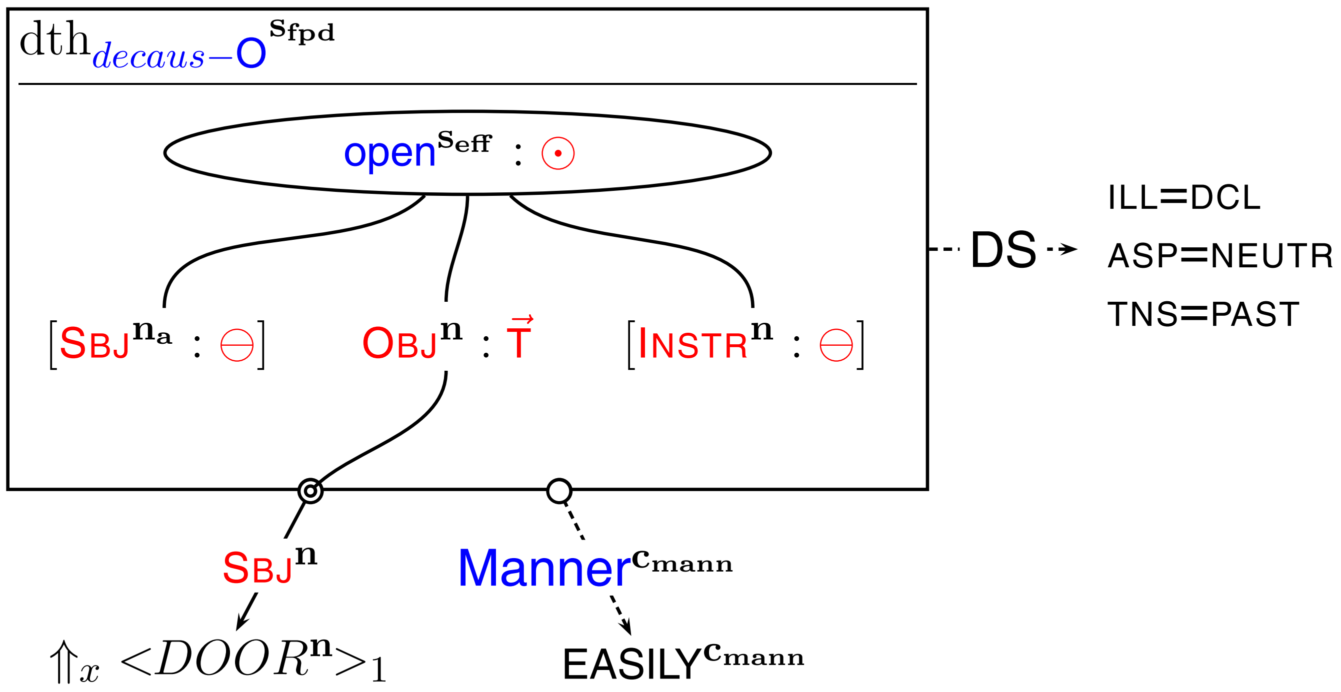




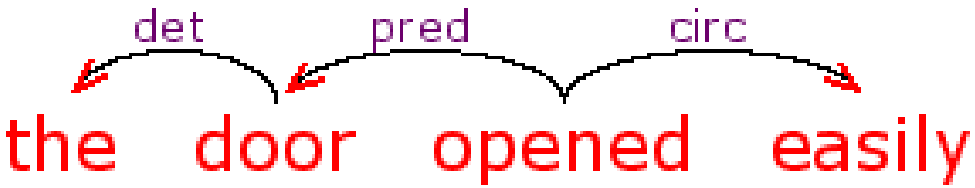
⟪open⟫ : $2^{ème}$ diathèse de décausation. Suite

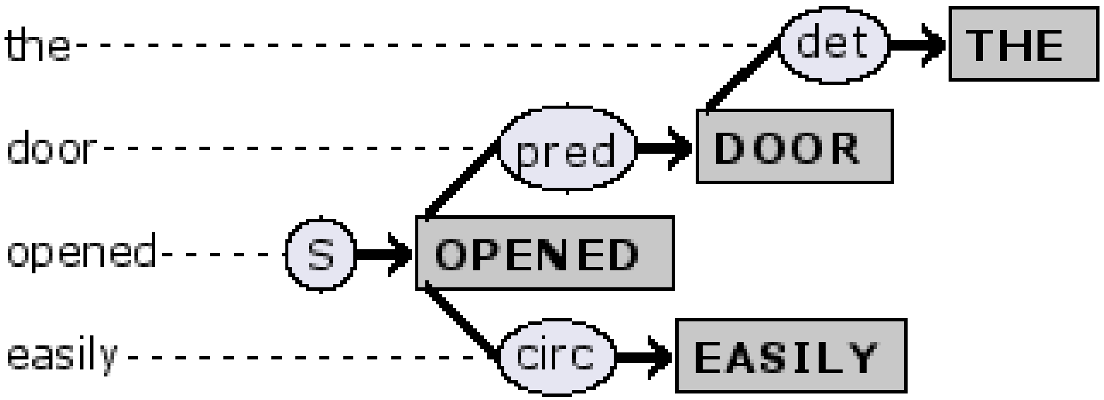
EX: *The door opened easily*

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



AD : $2^{ème}$ diathèse



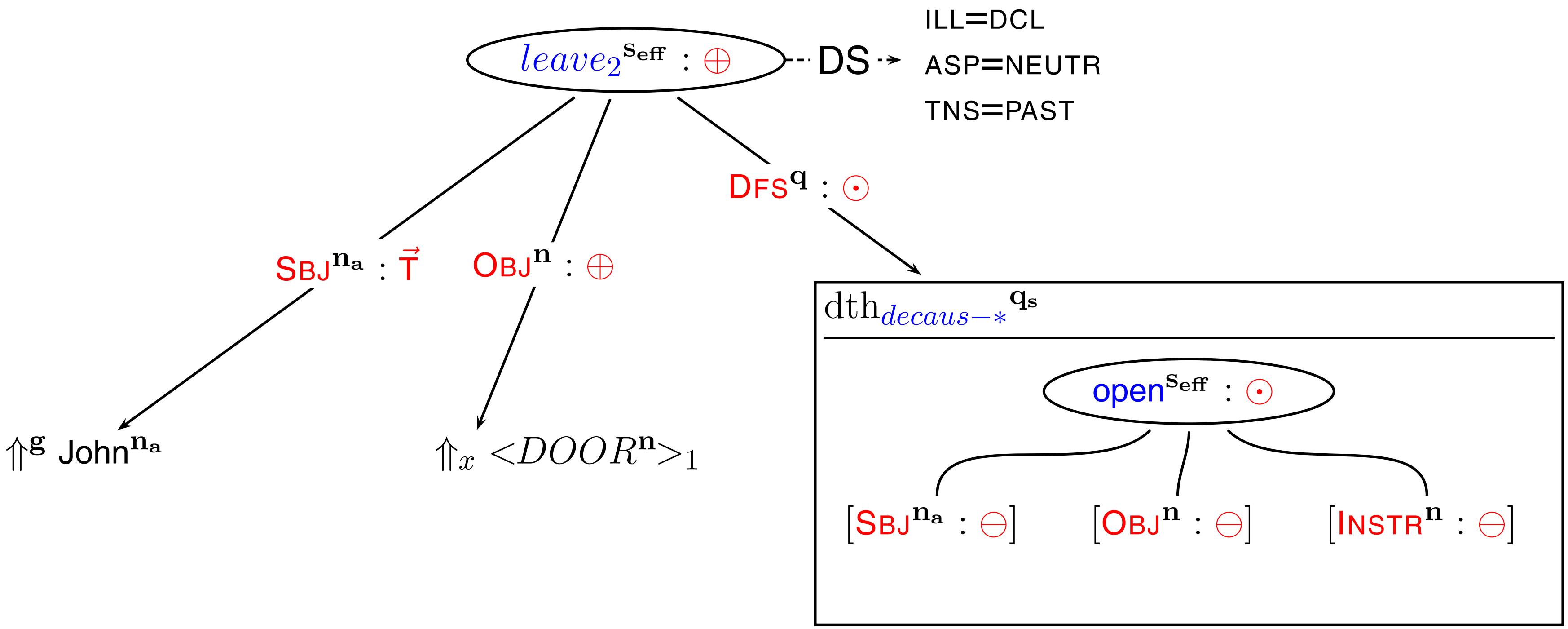


$\langle\langle$open$\rangle\rangle$: $3^{\grave{e}me}$ diathèse de décausation

EX: *John left the door opened*

$\langle\langle$open$\rangle\rangle$: $\langle\langle \mathrm{dth}_{decaus-*}\, (\bigstar_{\odot}, \emptyset \hookleftarrow \mathrm{SBJ}_{\ominus}, \emptyset \hookleftarrow \mathrm{OBJ}_{\ominus}, \emptyset \hookleftarrow \mathrm{INSTR}_{\ominus})^{\mathbf{q_s}} \rangle\rangle$

$\langle\langle leave_2 \rangle\rangle$ canonical profile: $\langle\langle leave_2 (\mathrm{SBJ}^{\mathbf{n_a}}, \mathrm{OBJ}^{\mathbf{n}}, \mathrm{DFS}^{\mathbf{q}})^{\mathbf{s_{eff}}} \rangle\rangle$



AD : $3^{ème}$ diathèse