

Y. MARTEMIANOV et Y. MOUCHANOV
(Moscou, Laboratoire de Traduction
Mécanique de l'Institut des
Langues Etrangères)

Une notation sémantique et ses
implications linguistiques

(Résumé)

1. Une notation est proposée en vue de fixer la structure sémantique d'un texte naturel de façon à en permettre le traitement automatique ultérieur, logique (déduction) ou linguistique (traduction). A la différence d'un langage documentaire (ou d'information), la notation proposée ne vise, en premier lieu qu'à éclaircir et ordonner la structure sémantique d'un énoncé naturel. Il s'agit, pratiquement, d'une forme de présentation, c'est-à-dire d'un "moule" structuré, où est "coulé", après reconstructions, répartitions et regroupements nécessaires, le contenu d'un texte naturel.

2. Pour répondre aux buts poursuivis, la Notation est voulue :

a) dument articulée , c'est-à-dire disposant d'un nombre suffisant de catégories sémantiques générales, régulièrement requisés par les textes naturels ;

b) suffisamment exhaustive - i.e. dégageant d'un énoncé naturel le maximum de données relatives aux catégories établies ;

c) rigoureusement réglementée afin de révéler, à l'analyse, les interprétations fausses ou incomplètes.

d) susceptible de différenciation , i.e. permettant, pour un contenu donné, le passage d'une structuration globale et superficielle à une structuration plus profonde et détaillée ;

e) absolument uniforme , i.e. répartissant les données obtenues, à l'état global ou différencié, parmi une même série de catégories-"facettes".

3. Pour toute description élémentaire (DE, unité fondamentale de la notation) une même série de catégories est prévue, à savoir :

- descripteurs : prédicat à n termes (P^n), sujets (S_i) fixant la valeur des termes, et soi-disant "attributs présumés" (ou "présumés" tout court - (P_i^*)) qui, à la différence du prédicat, désignent les propriétés qu'un sujet doit impliquer pour devenir valeur d'un terme donné ; un présumé spécial (P_0^*) est rattaché au prédicat lui-même.

Ex : Un oiseau sort du nid =

$$= S_1 [\text{un oiseau}] S_2 [\text{le nid}] P_1^* [\text{se déplacer}] P_2^* [\text{espace clos}] P_0^2 [\text{sortir}]$$

A bird flies out of the nest =

$$= S_1 [\text{a bird}] S_2 [\text{the nest}] P_1^* [\text{fly}] P_2^* [\text{space closed}] P_0 [\text{get out}]$$

$$P_0^* [\text{flying}]$$

- restricteurs (R) , désignant les limites temporelles (lim₁ - commencement , lim₂ - fin, lim₃ - suite), date et lieu de l'évènement décrit (α₀ , β₀ ...) , à moins qu'ils ne fassent partie de l'évènement lui-même en tant que valeurs prévues par les termes du prédicat.

Ex. : les moineaux se faisaient rares dans les rues =

$$= T \left[\begin{array}{l} \text{un temps passé} \\ \text{les rues} \end{array} \right] \left. \begin{array}{l} R \\ R \end{array} \right\} \begin{array}{l} \text{date} \\ \text{lieu} \end{array} \left(\begin{array}{l} (T, \alpha_0) \\ (L, \alpha_0) \end{array} \right) \alpha_0 = S_1 [\text{les moineaux}] P_1^* \left[\begin{array}{l} \text{élément} \\ \text{rare} \end{array} \right] P_0$$

Les pierres grandissent en s'approchant =

$$= T [\text{on s'approche des pierres}] R \begin{array}{l} \text{date} \\ \text{lieu} \end{array} \left(\begin{array}{l} (T, \alpha_0) \\ (L, \alpha_0) \end{array} \right) \alpha_0 = S_1 \left[\begin{array}{l} \text{les} \\ \text{pierres} \end{array} \right] P_1^* [\text{objet}] P_0 [\text{grandir}]$$

- Indicateurs (I) - Indiquant le mode d'existence d'une description par rapport au sujet qui relate, imagine, désire ou suppose(etc...) l'évènement décrit.

Ex : Qu'il entre :

$$= I [\text{vouloir } \alpha_0] \alpha_0 = S_1 [\text{il}] S_2 [\text{ici}] P_1^* [\text{se déplacer}] P_2^* [\text{espace clos}]$$

$$P_0^2 [\text{entrer}]$$

Vous auriez... assassiné la République =

$$= I [\text{admettre } \alpha_0 \text{ pour un moment passé et le sachant non réalisé}]$$

$$\alpha_0 = S_1 [\text{vous}] S_2 [\text{la République}] P_1^* [\text{personne}] P_2^* [\text{être animé}]$$

$$P_0^2 [\text{assassiner}]$$

- Quantificateurs (∀, ∃, ∃¹) appliqués aux termes et restricteurs non déterminés ;

- foncteurs (Γ, α, V₁ →) servant à distribuer, parmi les descriptions

élémentaires liées, les valeurs de "vrai" ou "faux".

Ex. : Autrefois les Français ne suivaient guère d'autres lois que celles du point d'honneur. = I [être presque sûr de α_0]

$$\alpha_0 = T[\exists t] \quad R_{\text{date}} \quad (T, \beta_0) \beta_0 = (S_2[\exists \text{ LOIS}] S_1 [\text{les Français}] P_0^2 [\text{suivre}]) \rightarrow (\text{lois} = \text{Lois du point d'honneur}).$$

L'implication et son dérivé -la causalité-(\rightarrow) - permettent de réduire à leur valeur concrète les abréviations conventionnelles qui, dans le langage naturel, animent les notions abstraites : la conscience règne et ne gouverne pas = $S_1[\forall a] P_1^* [\text{personne}] (P_0 [\text{conscient}] \rightarrow S_2[\forall b] P_2^* [\text{société}] (P_0^2 [\text{régner}] \alpha [P_0^2 [\text{gouverner}]]))$.

$$\text{"La curiosité naît de la jalousie"} = S_1[\forall a] P_2^* (\text{personne}) (P_0 [\text{jaloux}] \rightarrow P_0 [\text{devenir curieux}])$$

Remarque : Les notions de "prédicat", "termes", "quantificateurs" et "foncteurs" sont empruntées à la logique, sous réserve de la différence capitale qu'il y a entre leur emploi dans les sciences abstraites (étudiant les classes d'objets construites rationnellement) et dans les descriptions empiriques (objets réels disparates, classés intuitivement).

- Renvoyeurs (V_1 pour S_1 , V_2^* pour S_2^* etc..) - renvoyant à un élément déterminé d'une description antérieure ($V_2^* 1-$) ou postérieure ($V_2^* 1+$)

Ex. : Vous avez changé de langage =

$$= S_1 [V_0^* 1+] S_2 [V_0^* 2+] P_0^2 [\text{différer}]$$

$$F_0^* : T[\exists t] \quad R_{\text{date}} \quad (T, \alpha_0) \alpha_0 = S_2 [V_0] S_1 [\text{vous}] P_1^* [\text{personne}] P_0^2 [\text{parler}] P_0^* [\text{manière}]$$

$$F_0^* : T[t] \quad R_{\text{date}} \quad (T, \gamma_0) \gamma_0 = S_2 [V_0] S_1 [\text{vous}] P_1^* [\text{personne}] P_0^2 [\text{parler}] P_0^* [\text{manière}]$$

c'est-à-dire "diffèrent la manière dont vous parliez autrefois et celle dont vous parlez maintenant (t).

- Refondeurs (F_1 pour S_2 , F_2^* pour P_2^* etc..) - opérateurs présentant une description entière sous forme d'un seul de ses éléments, quitte à incorporer la description ainsi refondue dans une autre description (moyennant un renvoyeur).

Qui sait flatter, sait aussi calomnier =

$$= F_1 : S_1 [\exists a] S_2 [\exists b] P_1^* [personne] P_2^* [personne] P_0^2 [savoir flatter]$$

$$S_1 [\forall c, \forall_1 1-] S_2 [\exists d] P_1^* [personne] P_2^* [personne] P_0^2 [savoir calomnier].$$

(v. les transformations ultérieures à la page 6)

- opérateurs d'insistance : thème (\perp) - rhème (\exists), auxquels un sens logique précis est attaché, à savoir :

$$\perp a_0 (P(a_0, b_0)) \stackrel{\text{df}}{\equiv} \forall c (c = a_0 \rightarrow P(c, b_0))$$

$$\exists b_0 (P(a_0, b_0)) \stackrel{\text{df}}{\equiv} \exists c (P(a_0, c) \wedge c = b_0).$$

Ex. : l'oiseau, lui, vit dans un nid =

$$= S_1 (S_1 [\forall \text{oiseau}] S_2 [\exists \text{nid}] P_1^* [\text{\u00eatre anim\u00e9}] P_2^* [\text{espace clos}] P_0^2 [\text{vivre dans}]) =$$

$$\equiv P_1^* [\forall \text{\u00eatre anim\u00e9}] S_1 [\forall \text{oiseau}] ([\text{\u00eatre anim\u00e9}] = [\text{oiseau}] \rightarrow$$

$$\rightarrow S_2 [\exists \text{nid}] S_1 [\text{\u00eatre anim\u00e9}] P_0^2 [\text{vivre dans}]).$$

C'est dans un nid que l'oiseau vit =

$$= \exists S_2 (S_1 [\forall \text{oiseau}] S_2 [\exists \text{nid}] \dots P_2^* [\text{espace clos}] P_0^2 [\text{vivre dans}]) \equiv$$

$$\equiv S_1 [\forall \text{oiseau}] P_2^* [\text{espace clos}] (S_1 [\text{oiseau}] S_2 [\text{espace clos}] P_0^2 [\text{vivre dans}] \& [\text{espace clos}] = [\text{nid}]).$$

4. Foncteurs renvoyeurs et refondeurs constituent le minimum indispensable dont dispose la Notation pour marquer les relations syntaxiques entre les descriptions \u00e9l\u00e9mentaires, la syntaxe des cat\u00e9gories g\u00e9n\u00e9rales, inh\u00e9rentes \u00e0 la description, \u00e9tant compl\u00e8tement d\u00e9termin\u00e9e par leur signification m\u00eame.

Ex. : Nos grandes vertus finissent par nous lasser =

$$= F_0$$

$$\rightarrow (1+, 2+) \quad S_1 [\text{nous}] P_1^* [\text{personne}] P_0 [\text{vertueux}] - P_0 [\text{exc\u00e9der une norme}]$$

$$S_1 [\forall_0 1-]$$

$$\begin{aligned} \mathfrak{A}_T : T [\quad V_0 T \quad 1+] R \lim_2 (T, \alpha_0) \alpha_0 &= S_1 [\text{nous}] P_1^* [\text{personne}] \\ &P_0 [\text{vertueux}] \\ FT : T [\overset{0}{t}] R \lim_1 (T, \beta_0) \beta_0 &= S_1 [\text{nous}] P_1^* [\text{\AA}tre anim\AA] P_0 [\text{las}] \end{aligned}$$

i.e. Nous sommes vertueux et cela d'une façon tellement excessive qu'un moment vient où nous sommes plus vertueux et c'est le moment où nous commençons à en être las.

5. La séquence de descriptions élémentaires :

- ordonnées en colonne, avec leurs données réparties uniformément selon les catégories-facettes ;
- leurs relations mises en évidence ;
- la distribution de foncteurs, refoncteurs et opérateurs d'insistance permettant de distinguer les descriptions principales et secondaires de l'énoncé ;
- et l'ordre de descriptions dans la colonne symbolisant la marche de l'exposé.

semble, d'une part, assez similaire aux propositions logiques pour ouvrir un passage à la déduction et, d'autre part, assez conforme à la structure sémo-syntaxique de l'énoncé naturel (SSS) pour servir de forme-cible à l'analyse sémantique.

6. L'analyse d'un énoncé visant à la notation est assurée, en premier lieu, par la notation elle-même, qui :

- possède une structure suffisamment rigoureuse pour rejeter les résultats incorrects ;
- dispose d'une certaine synonymie qui facilite le passage d'un énoncé naturel à une structure éminemment logique par exemple :

$$\begin{aligned} \text{a) } S_1 [\forall a] P_1^* [M] P_0 [N] &\Leftrightarrow S_1 [\forall a, a \in M] P_0 [N] ; \\ \text{b) } S_1 [a, a \in \{ F_1 : S_1 [\exists b] S_2 [\exists c] P_1^* [M] P_2^* [N] P_0 [Q] \}] &\Leftrightarrow \\ &\Leftrightarrow S_1 [a] S_2 [\exists c] P_1^* [M] P_2^* [N] P_0 [Q] ; \\ \text{c) } S_1 [\forall a, a \in \{ F_1 : S_1 [\exists b] P_1^* [M] P_0 [N] \}] P_0 [Q] &\Leftrightarrow \\ &\Leftrightarrow S_1 [\forall a] P_1^* [M] (P_0 [N] \rightarrow P_0 [Q]). \end{aligned}$$

Ex. : La structure sémo-syntaxique :

$$F_1 : S_1 [\exists a] S_2 [\exists b] P_1^* [\text{personne}] P_2^* [\text{personne}] P_0^2 [\text{savoir flatter}]$$

$$S_1 [\forall d, d \in \{V_1 \text{ 1-}\}] S_2 [\exists c] P_1^* [\text{personne}] P_2^* [\text{personne}] P_0^2 [\text{savoir calomnier}]$$

obtenue pour l'énoncé "Qui sait flatter, sait aussi calomnier" (voir page 4 , devient l'objet des transformations suivantes :

$$S_1 [\forall d, d \in \{F_1 : S_1 [\exists \text{ personne}] S_2 [\exists \text{ personne}] P_0^2 [\text{savoir flatter}]\}] \\ \equiv S_1 [\forall \text{ personne}] (S_2 [\exists \text{ personne}] P_0^2 [\text{savoir flatter}] \rightarrow S_2 [\exists \text{ personne}] P_0^2 [\text{savoir calomnier}])$$

(i.e. "Si une personne sait flatter quelqu'un, il existe des personnes qu'elle saura calomnier").

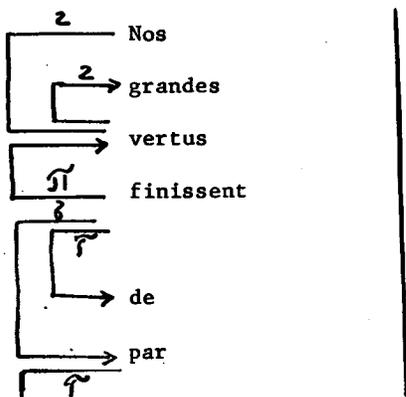
7. D'autre part, l'analyse s'appuie sur l'information sémantique rattachée à chaque mot et qui doit :

- être suffisamment articulée pour entrer, toutes les opérations nécessaires faites, dans une SSS correcte ;
- posséder, réunies dans une structure globale, toutes les structures actualisables, y compris les structures incompatibles.

Le souci principal de l'analyse sémantique d'un énoncé consiste, en gros, à éliminer, de l'information d'un mot pris dans son contexte syntaxique élucidé, les structures partielles

- a) qui ne sont pas appuyées par les valences syntaxiques actualisées,
- b) qui sont incompatibles avec d'autres structures, considérées comme plus convenables de par leur sens ;

Ex. : ($\mathcal{K}, \rho, \delta, \tau, \nu$) symbolisant les valences syntaxiques, resp. sujet, objet) direct, objet indirect, objet non-marqué, valence non-marquée ; sont surmontées d'un trait les structures partielles, éliminées de l'information sémantique du mot dans le contexte syntaxique donné : $\delta \downarrow$ - mot dépendant de δ , $\pi \uparrow$ valence empruntée à un mot dominant) :



$$S_1 [\text{nous}] S_2 [\nu] P_1^* [\text{personne}] P_0 [\nu/\text{avoir}] \\ S_1 [\nu] P_0 [\text{excéder une norme}] \\ S_1 [\nu] P^* [\text{personne}] P_0 [\text{vertueux}] \\ S_1 [\pi] P_0 [\exists d/\pi] \rightarrow T [\delta \downarrow] R \lim_2 (T, \alpha_0) \\ \alpha_0 = S_1 [\pi] P_0 [\rho/\tau/\pi] \\ \dots \\ P_0 [\nu] P_0^* [\tau]$$

9. L'appartenance à de diverses parties du discours des mots sémantiquement identiques, pas plus que le découpage syntaxique différent que subit une même information dans l'énoncé, ne sont pris en considération que pour autant que ces faits de langue permettent de desceller soit la subordination des descriptions élémentaires, soit la distribution d'opérateurs d'insistance (v. les exemples aux pages 3-5).

10. Le travail sur la Notation , entrepris pour orienter l'analyse automatique et permettre la déduction logique ultérieure, débouche vers la synthèse de textes naturels et de l'analyse documentaire, où des solutions, encourageantes se profilent.