MATHÉMATIQUES ET SCIENCES HUMAINES

A. CULIOLI
J. P. DESCLES

Traitement formel des langues naturelles. Deuxième partie : dérivations d'exemples

Mathématiques et sciences humaines, tome 78 (1982), p. 5-31
<http://www.numdam.org/item?id=MSH_1982__78__5_0>

Math. Sci. hum. (20^e année, n°78, *1982*, p.5-31)

TRAITEMENT FORMEL DES LANGUES NATURELLES

deuxième partie : DERIVATIONS D'EXEMPLES.

A. CULIOLI*

J.P. DESCLES**

La première partie de l'article (voir Math. Sci. Humaines, n° 77, 1982) a présenté quelques concepts qui seront maintenant utilisés. Cette seconde partie va donner deux exemples de dérivation d'énoncés. Le problème est celui des énoncés dits à "double sujet", c'est-à-dire des énoncés où un terme, (voire plusieurs) est thématisé. La notation utilisée (qui s'apparente au "lambda-calcul" de Church) offre le moyen de représenter dans l'écriture le concept même de thématisation ou du moins de montrer, en le notant, qu'il conviendrait mieux d'envisager différents degrés de thématisation plutôt qu'une opposition tranchée "thématisé / non thématisé". En effet, prenons les familles d'énoncés très simples :

> *Jean est à Paris / Jean, il est à Paris / Il y a Jean qui est à Paris / C'est Jean qui est à Paris /*
>
> *Jean a embrassé Marie / Jean, il a embrassé Marie / Jean, Marie, il l'a embrassée / Marie, Jean l'a embrassée / C'est Jean qui a embrassé Marie / C'est (bien) Marie que Jean a embrassée ...etc*

Il ne s'agit pas simplement de dire qu'un terme est thématisé (ou pas) pour décrire avec précision ces deux familles. Il est nécessaire de montrer comment se constituent les énoncés, par quelles opérations enchaînées le linguiste peut reconstruire les énoncés dans lesquels un terme est posé pour ensuite se voir attribuer une propriété rhématique. La seconde famille montre bien que le fameux ordre canonique du français S-V-O (Sujet-Verbe-Objet) est complètement disloqué puisque l'on peut commencer tantôt par un terme, tantôt par un autre mais, selon le "terme de départ", il y aura des reprises qui établissent des liens dans la relation prédicative entre les termes agencés. On peut évoquer les phénomènes de redoublement de l'objet

---

* DRL, ERA 642 CNRS, Université de Paris 7.

** UER de mathématiques et informatique, ERA 642 CNRS, Université de Paris 7

(albanais, bulgare, grec moderne, roumain...etc) mais ces phénomènes ne sont pas le propre des langues balkaniques. Pour raisonner sur ces questions, le Calcul des prédicats est un système de représentations fort efficace. Nous visons à indiquer quel est le système de notation employée et comment, dans ce système métalinguistique, on peut effectuer des déductions de théorèmes à partir de schémas d'axiome et de schémas de règles. Les problèmes évoqués (degrés de thématisation, redoublement de l'objet, énoncés à "double sujet") seraient tous traités avec les concepts introduits . Le lecteur reconnaîtra dans les questions abordées la problématique que l'Ecole de Prague désignait sous l'expression "aktuální členění větné" que l'on peut rendre en français par "décomposition énonciative"*. Pour Mathésius l'énoncé doit être analysé du point de vue de "l'intention communicative" de l'énonciateur; il se présente donc sous forme d'un thème ou "point de départ" de l'énoncé et d'un rhème, appelé (par Mathésius) "noyau". Le thème apparaît comme une sorte de support de la prédication.
L'article propose un traitement formel et déductif de la "décomposition énonciative" en précisant les écritures métalinguistiques, les règles de manipulation de ces écritures et en indiquant des déductions simulables éventuellement par ordinateur. De plus, une distinction entre, d'un côté, le repère constitutif et d'un autre côté, le terme de départ est explicitement faite. Ceci évite de réduire la décomposition énonciative à un thème conçu comme une sorte de rappel destiné à introduire l'information nouvelle et à un noyau qui serait l'élément nouveau d'information concernant le thème.
Il serait intéressant de comparer notre approche formelle des solutions qu'apporterait par exemple la Grammaire Relationnelle (de Postal et Johnson) ou les formalismes issus de la "Grammaire universelle" (à la Montague).

---

* Le lecteur français pourra se reporter au livre de J. Fontaine : Le cercle Linguistique de Prague, Mame, Paris, 1974, (p. 105, 132).

1.FRAGMENT D'UN SYSTEME FORMEL

A plusieurs reprises, nous avons évoqué le problème de la thématisation, que ce soit en français avec des énoncés comme : il y a Jean, son bras il est cassé ou il y a Jean qui aime Marie ; en japonais avec des énoncés comme Kazuko wa me ga ookii desu (— "Quant à Kazuko, ses yeux sont grands" "les yeux de Kazuko sont grands"; il y a Kazuko, elle a de grands yeux (ou "ses yeux sont grands") ou comme : Tanakasan wa okusan ga byooki desu ("Monsieur Tanaka a sa femme (qui est) malade"). En arabe, nous avons des familles comme :

1. qāma ?abū zaydin — Le père de Zayd s'est levé
   "s'est levé père Zayd-de"
2. ?abū zaydin qāma — Le père de Zayd, il s'est levé
   "père Zayd-de s'est levé"
3. zaydun qāma ?abūhu — Zayd, son père s'est levé
   "Zayd s'est-levé père-lui"
4. zaydun ?abūhu qāma — Zayd, son père, il s'est levé
   "Zayd père-lui s'est levé"
5. ?inna zaydan qāma ?abūhu — Il y a Zayd, son père s'est levé
   "Il y a Zayd s'est levé père-lui"
6. ?inna zaydan ?abūhu qāma — Il y a Zayd, son père, il s'est levé
   "Il y a Zad père-lui s'est levé"
7. ?inna ?abā zaydin qāma — Il y a le père de Zayd, il s'est levé
   "Il y a père Zayd-de il-s'est levé"

Pour rendre compte de cette dernière famille et des différences entre chacun des énoncés, pour expliquer la (ou plutôt les) valeur(s) du marqueur ?inna, pour montrer comment est constituée la "phrase nominale" en arabe classique et étudier précisément la catégorie du mubtada? il est en fait indispensable d'avoir des opérations abstraites et des règles qui enchaînent ces opérations de façon à produire un énoncé en : (a) lui attribuant une représentation métalinguistique pertinente et (b) en l'insérant dans une famille paraphrastique d'énoncés eux-mêmes munis de leurs représentations.

D.E. Kouloughli dans un article écrit en 1976 et s'inspirant directement des règles qui seront données plus loin, a montré que :

1° La"phrase nominale" de la tradition grammaticale arabe ne peut pas se ramener à la "phrase nominale" ("sans verbe") de la tradition grammaticale indo-européenne ;

2° La "phrase nominale" arabe peut être essentiellement définie comme une phrase à mubtada ? ne saurait être assimilée à la catégorie linguistique gréco-latine de "sujet".

En fait, le mubtada ? (mubtada ? un bihi, littéralement "ce par quoi on commence") doit être rendu comme terme de départ abstrait de l'énoncé et non pas, comme l'ont (mal) compris des lecteurs superficiels des grammairiens arabes comme un premier terme (concret) de l'ordre total exprimé par une phrase.

Il est donc nécessaire de bien distinguer, d'une part, un ordre total de positionnement des termes linguistiques dans l'énoncé ou la phrase, ce positionnement résultant d'une simple concaténation et, d'autre part, d'un ordre* qui résulte d'une orientation de la relation prédicative, cette orientation résulte d'une prise en compte de la constitution de la relation prédicative à partir d'un premier terme (terme de départ). Le terme de départ ne peut être mis en évidence que si l'on analyse l'histoire constitutive de la relation prédicative et donc de l'énoncé qui en résulte.

La catégorie du mubtada ? s'analyse alors comme le résultat d'une opération métalinguistique d'orientation (règle 1 dans la liste des règles données plus loin).

---

* ordre séquentiel

1.1 NOTATIONS

Considérons une lexis λ présentée de façon simplifiée sous la forme d'une occurrence d'une relation prédicative binaire :

$$< a\ r\ b >$$

où a, b, d'une part et r, d'autre part, sont des termes linguistiques déjà construits et catégorisés respectivement en termes nominaux terme prédicatif. Nous ne donnons pas,ici, les règles qui à partir d'une notion (qui rappelons-le, ni verbale, ni nominale) engendreraient des termes nominaux ou des termes prédicatifs. Nous ne donnons pas ici non plus les règles qui montreraient comment est constituée la lexis λ présentée sous forme d'une occurrence de relation. Le détail de ces règles n'est pas pertinent pour la suite de l'exposé.

Introduisons une notation. A la lexis $\lambda \equiv < a\ r\ b >$ sont associées deux lexis non saturées :

$$\lambda^*_1 \equiv < (\ )\ r\ b >$$

$$\lambda^*_2 \equiv < a\ r\ (\ ) >$$

où ( ) représente l'occurrence d'une place d'argument dans la relation prédicative.

Une lexis non saturée est analogue à une forme propositionnelle (nécessairement non saturée).

Si λ désigne une lexis saturée, λ^* désigne une lexis non saturée, c'est à dire l'une des lexis λ^*_1 ou λ^*_2. La lexis doublement non saturée λ^{**} associée à λ sera une occurrence de relation de la forme : $< (\ )\ r\ (\ ) >$

C'est une forme propositionnelle avec deux places d'arguments non instanciées.

Remarque 1 :

En utilisant la "lambda-notation" de Church (CHU,44), il est clair que si une lexis $\Lambda \equiv < a\ r\ b >$ est une "lambda expression", alors :

$$\lambda\ \xi.\ < \xi\ r\ b >$$

$$\lambda\ \xi.\ < a\ r\ \xi >$$

sont des "lambda-expressions" qui représentent deux lexis non saturées.

D'une façon générale, désignons par Λ^* une "lambda expression" où la variable ξ est présente (occurrence de ξ) : nous avons la nouvelle "lambda-expression :

$$\lambda\ \xi.\ \Lambda^*$$

Il est bien entendu que le symbole λ dans chaque "lambda-expression" préfixée comme :

$$\lambda\ \xi.\ \langle\ \xi\ r\ b\ \rangle$$

n'a rien à voir avec les symboles λ ,λ^* qui désigne des lexis.

La <u>saturation</u> d'une lexis non saturée s'effectue par une double opération :

1) repérage d'une lexis non saturée λ^* par rapport à un terme nominal x;
2) identification de la place vide dans la lexis non saturée avec l'objet nominal x qui sert de repère.

Le résultat de cette double opération est noté :

$$\langle\ \lambda^*\ \underline{\in}\ x\ \rangle$$

en utilisant un lien d'identification $\ulcorner\!=\!\urcorner$ entre la place vide dans λ^* et son identificateur.

<u>Remarque 2 :</u>

En utilisant la "lambda-notation" de Church, le résultat de la saturation est exprimé par la "lambda-expression" $(\lambda\ \xi.\ \Lambda^*)\ x$.

En introduisant la β-réduction de H.B. Curry (CUR,58), nous voyons que la "lambda-expression" précédente est un <u>redex</u>, c'est-à-dire l'expression d'une opération "à effectuer" qu'il faut distinguer de son <u>contractum</u> obtenu en substituant l'occurrence de x aux occurrences de ξ dans Λ^* et en effectuant, si possible, l'opération qui en résulterait.

Pour les deux lexis non saturées λ^*_1 et λ^*_2 nous obtenons :

$$\langle\ \langle(\ )\ r\ b\ \rangle\ \underline{\in}\ a\ \rangle$$

$$\langle\ \langle\ a\ r\ (\ )\ \rangle\ \underline{\in}\ b\ \rangle$$

1.2 REGLES ET AXIOMES.

1.2.1 Le schéma d'axiome est de la forme :

$$< \lambda \ \underline{\in} \ \text{Sit}\ (\mathcal{S}_0, \mathcal{T}_0) >$$

"Toute lexis (et toute relation prédicative initiale) est repérée par rapport à la situation initiale d'origine : Sit $(\mathcal{S}_0,\mathcal{T}_0)$". Ce schéma d'axiome revient à repérer la lexis λ et toutes les relations prédicatives qui en seront issues par rapport à l'origine de l'énonciation, c'est-à-dire que la lexis λ est insérée dans l'espace énonciatif subjectif de l'énonciateur.

1.2.2. Nous introduisons quatre règles (ou schémas de règle) qui concerneront uniquement les modes de constitution des relations prédicatives. Le symbole ' → ' désigne une relation de préordre qui joue le rôle d'un symbole de réecriture (ou de déduction). Chaque règle comporte une partie gauche, une formule F_g, et une partie droite une formule F_d, séparées par ' → '. La partie gauche est réécrite par la partie droite, d'où l'expression de la règle :

$$F_g \rightarrow F_d$$

Nous considérons des variables formelles (non libres et non liées) :

u,v,... de type $\underline{n}$ (nom)

V,W,... de type $\underline{p}$ (proposition)

Nous utilisons le symbole de repérage $\underline{\in}$ dont les occurrences dans une formule sont désignées par :

$$\underline{\in}_1, \ldots, \underline{\in}_k, \ldots, \underline{\in}_m, \ldots$$

Les symboles de "bonne formation" de sous-formules sont étiquetés. Nous considérons un ensemble de tels symboles appariés entre eux et étiquetés :

$$\{ <_1 , {}_1> , \ldots , <_m , {}_m> , \ldots , \quad \}$$

1.2.2.1. Règle 1 : règle d'orientation

$$\lambda \rightarrow < \lambda * \underline{\in} \; x >$$

"Toute relation prédicative λ est nécessairement orientée et comporte un terme de départ x".

1.2.2.2. Règle 2 : Règle du repère constitutif

$$< \lambda * \underline{\in} \; x > \rightarrow < < \lambda * \underline{\in} \; x > \underline{\in} < y \; \underline{\in} \; \mathrm{Sit} \; (\mathcal{S}_0 , \mathcal{T}_0) > >$$

"Toute relation prédicative orientée $< \lambda * \underline{\in} \; x >$ est repérée par rapport à un repère constitutif y, lui-même repéré par rapport à la situation d'énonciation d'origine".

1.2.2.3. Règle 3 : Schéma d'intrication

$$<_1 \ldots <_p \ldots <_m V \underline{\in}_i \; x \; {}_{m'}> \ldots {}_{p'}> \underline{\in}_j <_q \ldots <_n x \; \underline{\in}_k \; W \; {}_{n'}> \ldots {}_{q'}> \ldots {}_{1'}>$$

$$\rightarrow <_1 \ldots <_p \ldots <_m V \; \underline{\in}_i \; <_q \ldots <_n x \; {}_{m'}> \ldots {}_{p'}> \underline{\in}_k \; W \; {}_{n'}> \ldots {}_{q'}> \ldots {}_1>$$

1.2.2.4. Schéma de règle 4 : schéma d'absorption

$$<_1 \ldots <_j \ldots <_i \ldots \underline{\in} \; \mathrm{Sit} \; {}_i> \underline{\in} \; \mathrm{Sit} \; {}_j> \ldots {}_1> \rightarrow <_1 \ldots <_i \ldots \underline{\in} \; \mathrm{Sit} \; {}_i> \ldots {}_1>$$

1.2.2.5. Schéma de règle 5 : schéma de dualité

Soient deux expressions E et F,

$$< E \; \underline{\in} \; F > \rightarrow < F' \; \underline{\ni} \; E' >$$

avec les règles suivantes :

i) soit H une sous-expression "bien parenthésée", c'est-à-dire dont tous les chevrons sont appariés deux à deux ; si H présente une occurrence dans E ou dans F, dans une partie gauche de la règle, alors H reste inchangé dans la partie droite de la règle ;

ii) si un chevron ouvrant ou fermant présente une occurrence non appariée dans E ou dans F dans la partie gauche de la règle, alors ce chevron apparaît en position miroir dans la partie droite de la règle et s'il était ouvrant (respectivement fermant) il devient fermant (respectivement ouvrant).

Les expressions F' et E' sont des expressions <u>pseudo-miroirs</u> de F et E obtenues en prenant le miroir des sous-expressions qui ne sont pas des sous-expressions "bien parenthésées". Nous pouvons reformuler la règle ainsi. Désignons par $\underline{m}$, $\underline{n}$, $\underline{p}$, $\underline{q}$, une occurrence d'un chevron (ouvrant ou fermant) et par $\underline{m}'$, $\underline{n}'$, $\underline{p}'$, $\underline{q}'$, une occurrence du chevron correspondant apparié (si $\underline{p}$ est un chevron ouvrant, alors $\underline{p}'$ est fermant...). Supposons que les occurrences des chevrons $\underline{m}$, $\underline{n}$, $\underline{p}$, $\underline{q}$, ...n'aient pas de chevrons appariés dans les expressions E et F. Désignons par $<A>$ et par $<B>$ des occurrences de sous-expressions "bien parenthésées" dans E ou dans F. Le schéma de règle de dualité s'exprime formellement :

$$\underline{m} \ldots < A > \ldots \underline{n} \ldots \ \underline{\in} \ \ldots \underline{p} \ldots < B > \ldots \underline{q} \ldots$$

$$\rightarrow \ldots \underline{q}' \ldots < B > \ldots \underline{p}' \ldots \ \underline{\ni} \ \ldots \underline{n}' \ldots < A > \ldots \underline{m}' \ldots$$

<u>Exemples</u>

(a) $< < A > > > \underline{\in} B > \rightarrow < B \ \underline{\ni} \ < < < A > >$

(b) $< < a \underline{\in} b > \underline{\in} < c > \rightarrow < c > \underline{\ni} < a \underline{\in} b > >$

(c) $< < < A > \underline{\in} < B > > \rightarrow < < B > \underline{\ni} < A > > >$

1.2.3. <u>COMMENTAIRES SUR LES REGLES</u>

1.2.3.1. Etant donnée une lexis λ, celle-ci exprime un certain type d'agencement entre des termes qui ont déjà un fonctionnement assigné. Cette lexis va pouvoir être exprimée à son tour de différentes façons selon les modes de constitution des relations prédicatives dérivées. La première opération consiste à choisir un <u>terme de départ</u> (voir plus haut ce que nous avons dit à propos de la catégorie arabe du <u>mubtada ?</u>). A partir de ce terme de départ, toute la relation prédicative va s'organiser progressivement.

1.2.3.2. La règle 2 permet d'assigner un repère constitutif qui est directement repéré par rapport à l'origine de l'énonciation. En français, nous avons des marqueurs comme <u>il y a</u>, <u>en ce qui concerne</u>... ou simplement des marqueurs prosodiques qui témoignent, en tant que trace repérable, du bien fondé de la règle. Beaucoup de problèmes, mais pas tous, de thématisation sont dans le champ de cette règle. "Terme de départ" et "repère constitutif" ne sont pas toujours identiques comme cela est indiqué dans les énoncés :

<u>Jean, son bras, il lui fait mal</u>

<u>Cette maison, moi, je ne suis pas acheteur</u>

où <u>Jean</u>, <u>cette maison</u> ont un rôle de repères constitutifs ou de thèmes à propos desquels la relation prédicative va exprimer un contenu rhématique c'est-à-dire une prédication.

1.2.3.3. Le schéma de règle 3 a pour but <u>d'intriquer</u> un même terme x, qui serait inséré dans deux relations distinctes, dans une seule formule où il ne présenterait alors plus qu'une seule occurrence. Ce schéma de règle fait passer de formules "bien formées" (c'est-à-dire "bien parenthésées") à des formules où les parenthèses ouvrantes et fermantes se chevauchent. Il est alors nécessaire de bien gérer les chevrons ouvrants ' < ' et les chevrons fermants ' > '.

Reformulons la règle d'une façon plus formelle.

Désignons par $\underline{i}$, $\underline{j}$,...$\underline{p}$,$\underline{q}$,.. des occurrences de chevrons ouvrants des sous-expressions et par $\underline{i}'$, $\underline{j}'$,...$\underline{p}'$,$\underline{q}'$,... les occurrences appariées des chevrons qui ferment les sous-expressions.

La règle en question s'écrit alors sous la forme :

$$\underline{l}\cdots\underline{p}\cdots\underline{m}\ V\ \underline{\epsilon_i}\ x\quad \underline{m}'\cdots\underline{p}'\cdots\ \underline{\epsilon_j}\quad \underline{q}\cdots\underline{n}\quad x\ \underline{\epsilon_k}\ W\ \underline{n}'\cdots\underline{q}'\cdots\underline{l}'$$

$$\rightarrow\ \underline{l}\cdots\underline{p}\cdots\underline{m}\ V\ \underline{\epsilon_i}\ \underline{q}\cdots\underline{n}\quad x\quad \underline{m}'\cdots\underline{p}'\quad \underline{\epsilon_k}\ W\ \underline{n}'\cdots\underline{q}'\cdots\underline{l}'$$

"Oublions" par effacement les symboles qui ne sont pas des chevrons, les expressions à gauche et à droite de la règle ont les formes respectives :

$\underline{l}\ldots\underline{p}\ldots\underline{m}\ldots\underline{m}'\ldots\underline{p}'\ldots\underline{q}\ \ldots\underline{n}\ldots\underline{n}'\ldots\underline{q}'\ldots\underline{l}'$

$\underline{l}\ldots\underline{p}\ldots\underline{m}\ldots\underline{q}\ldots\underline{n}\ldots\underline{m}'\ldots\underline{p}'\ldots\underline{n}'\ldots\underline{q}'\ldots\underline{l}'$

Exemples

$<_1 \quad <_2 \quad <_3 \quad <_4 \ {}_4'> \quad {}_3'> \quad <_5 \quad <_6 \ {}_6'> \quad {}_5'> \quad {}_2'> \quad {}_1'>$

$<_1 \quad <_2 \quad <_3 \quad <_4 \quad <_5 \quad <_6 \ {}_4'> \quad {}_3'> \quad {}_6'> \quad {}_5'> \quad {}_2'> \ {}_1'>$

Cette formulation de la règle met en évidence que si un terme x est inséré dans deux sous formules "bien parenthésées",

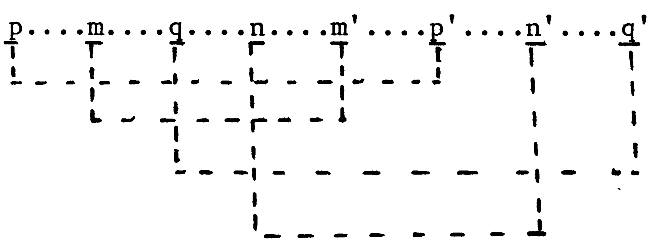
$\underline{p}\ldots\underline{m}\ldots\underline{m}'\ldots\underline{p}'$

$\underline{q}\ldots\underline{n}\ldots\underline{n}'\ldots\underline{q}'$

alors ce terme x peut-être intriqué dans une seule sous-formule qui n'est plus "bien parenthésée" et de la forme :

$\underline{p}\ldots\underline{m}\ldots\underline{q}\ldots\underline{n}\ldots\underline{m}'\ldots\underline{p}'\ldots\underline{n}'\ldots\underline{q}'$

La structure d'appariemment de cette sous-formule est :



1.2.3.4. Le schéma de règle d'absorption est introduit puisqu'il n'y a pas changement des <u>coordonnées énonciatives</u> de la relation prédicative ; il y a absorption à droite des coordonnées identiques.

1 2.3.5. Nous avons déjà introduit la transformation de dualité (1.2.2.5) La formulation retenue ici est essentiellement une généralisation métalinguistique qui autorise la transformation d'une expression engendrée en une expression pseudo-miroir.

1.2.3.6. Remarque 3

Tout C-langage L (ou langage algèbrique ou encore langage de type 2 dans la hiérarchie de Chomsky,, cf. (GRO/LEN.1967) est tel que l'on peut lui associer : 1° un langage de Kleene standard Q

2° un langage de Dyck D*

3° un homomorphisme φ de langage tel que $L = \varphi(D^* \cap Q)$

Un langage D* de Dyck est construit à partir d'un bi-alphabet de lettres appariées.

Désignons par $A \cup A'$ ce bi-alphabet, d'où le monoïde libre engendré par celui-ci et noté $(A \cup A')^*$ dont le neutre est désigné par '1'

Désignons par G(A) "le groupe libre" engendré par ce bi-alphabet dont le neutre est noté : '0'.

Considérons l'homomorphisme canonique h de $(A \cup A')^*$ dans G(A) tel que

$$h(1) = 0 \; ; \; h(a) = a \; ; \; h(a') = -a$$

en désignant par '-a' l'inverse, dans G(A), de l'élément a.

Le noyau h^{-1} de cet homomorphisme est le langage de Dyck D^*. Il comprend tous les mots construits avec le bi-alphabet $A \cup A'$ qui sont "réductibles" à 1 par effacement de lettres contiguës et appariées.

Tout C-langage L s'obtient alors par homomorphisme d'un C-langage standard convenablement choisi. Le langage de Dyck D* s'interprète directement comme un langage d'expressions "bien parenthésées" où les parenthèses sont toutes appariées deux à deux. On voit immédiatement que les règles d'intrication "transforment" des expressions d'un langage L (éventuellement un C-langage) en des expressions d'un langage L' qui lui n'est pas en général un C-langage. Il serait intéressant de situer les langages L' ainsi obtenus dans la hiérarchie de Chomsky. Plus exactement, les langages L' obtenus par une "transformation d'intrication" à partir de C-langage L forment-ils une sous-famille stricte des langages récursifs ? des langages de type 1 ?

Remarque 4

Reprenons le schéma de dualité. Supposons que les $\mathcal{H}$ que nous avons appelées "expressions bien parenthésées" soient toutes engendrées par une C-grammaire de Chomsky. Ces expressions appartiennent alors toutes à un certain C-Langage L.

Une expression d'un langage L' obtenu à partir de L, étant construite éventuellement à l'aide d'une (ou plusieurs) intrication(s), peut être transformée à son tour par dualité. Ce type de transformation opère entre expressions de L'. La formulation de la règle de dualité devient simple : Toute "sous-expression bien parenthésée" ou expression de L reste inchangée dans E' ou F', le reste de la formule se transforme en son miroir.

2. DERIVATIONS FORMELLES D'ENONCES

2.1 Considérons la lexis initiale :

$$\lambda = <_{o}\ a\ r\ b\ _{o}>$$

Appliquons dans l'ordre les règles (1), (2), (3) puis (4), nous avons la déduction suivante :

$$
\begin{array}{l|l}
1 & <\ \ <_0 a\ r\ b\ _0> \underline{\in}\ \mathrm{Sit}\ (\mathcal{S}_0,\mathcal{T}_0)\ > \\
2 & <\ \ <_1\ <_0 \overbrace{(a)\ r\ b\ _0> \underline{\in}\ a}^{=}\ _1> \in \mathrm{Sit}\ (\mathcal{S}_0,\mathcal{T}_0)\ > \\
3 & <\ \ <_1\ <_0 \overbrace{(\ )\ r\ b\ _0> \underline{\in}\ a}^{=}\ _1> \in\ <_2 a\ \underline{\in}\ \mathrm{Sit}\ (\mathcal{S}_0,\mathcal{T}_0)\ _2> \underline{\in}\ \mathrm{Sit}\ (\mathcal{S}_0,\mathcal{T}_0)> \\
4 & <\ \ <_1\ <_0 \overbrace{(\ )\ r\ b\ _0> \underline{\in}\ <_2\ a}^{=}\ _1> \underline{\in}\ \mathrm{Sit}\ (\mathcal{S}_0,\mathcal{T}_0)\ _2> \underline{\in}\ \mathrm{Sit}\ (\mathcal{S}_0,\mathcal{T}_0)\ > \\
5 & <\ \ <_1\ <_0 \overbrace{(\ )\ r\ b\ _0> \underline{\in}\ <_2\ a}^{=}\ _1> \underline{\in}\ \mathrm{Sit}\ (\mathcal{S}_0,\mathcal{T}_0)\ _2>\ \ >
\end{array}
$$

Nous en déduisons la formule en omettant les chevrons externes :

$$\mathrm{I} \qquad \boxed{<_1\ <_0 \overbrace{(\ )\ r\ b\ _0> \in\ <_2\ a}^{=}\ _1>\ \in\ \mathrm{Sit}\ (\mathcal{S}_0,\mathcal{T}_0)\ _2>}$$

Supposons que nous partions de la lexis :

$$\lambda \equiv <\ /\mathrm{Jean}/,\ /\mathrm{conduire}/,\ /\mathrm{voiture}/\ >$$

La formule (I) s'interprete avec :

a ≡ /Jean/ ; b ≡ /conduire/ ; c ≡ /voiture/

pour l'énoncé :

Il conduit la voiture, Jean

Le marqueur il est la trace d'une identification entre une place vide insérée dans une relation prédicative (non saturée) et son identificateur Jean . La virgule qui est insérée entre il conduit la voiture et Jean est la trace d'un marqueur prosodique (pause) nécessaire.

En utilisant le schéma de dualité dans la sous-formule $\langle_2 \ldots {}_2\rangle$ (pas 6 et 7) en reportant dans la formule générale (pas 8) nous obtenons en faisant fonctionner la dualité sur ${}_1\langle_0 \ldots {}_0\rangle \;\underline{\in}\; \langle_2 \ldots {}_2\rangle$ la formule du pas 9 :

5 $\quad \langle\ \langle_1 \langle_0\ (\ )\ r\ b\ {}_0\rangle\ \underline{\in}\ \langle_2\ a\ {}_1\rangle\ \underline{\in}\ \text{Sit}\ (\mathcal{S}_0, \mathcal{T}_0)\ {}_2\rangle\ \rangle$

6 $\quad \langle_2\ a\ {}_1\rangle\ \underline{\in}\ \text{Sit}\ (\mathcal{S}_0, \mathcal{T}_0)\ {}_2\rangle$

7 $\quad \langle_2\ \text{Sit}\ (\mathcal{S}_0, \mathcal{T}_0)\ \underline{\ni}\ \langle_1\ a\ {}_2\rangle$

8 $\quad \langle\ \ \langle_1\ \ \langle_0\ (\ )\ r\ b\ {}_0\rangle\ \underline{\in}\ \langle_2\ \text{Sit}\ (\mathcal{S}_0, \mathcal{T}_0)\ \underline{\ni}\ \langle_1\ a\ {}_2\rangle\ \ \rangle$

9 $\quad \langle\ \ \langle_2\ \text{Sit}\ (\mathcal{S}_0, \mathcal{T}_0)\ \underline{\ni}\ \langle_1\ a\ {}_2\rangle\ \underline{\in}\ \langle_0\ (\ )\ r\ b\ {}_0\rangle\ \ {}_1\rangle\ \ \rangle$

En omettant les chevrons externes nous obtenons la représentation métalinguistique :

$$\text{I'} \qquad \boxed{\langle_2\ \text{Sit}\ (\mathcal{S}_0, \mathcal{T}_0)\ \underline{\ni}\ \langle_1\ a\ {}_2\rangle\ \underline{\ni}\ \langle_0\ (\ )\ r\ b\ {}_0\rangle\ \ {}_1\rangle}$$

qui représente l'énoncé :

<u>Il y a Jean qui conduit la voiture</u>

ou encore (avec une pause)

<u>Il y a Jean, il conduit la voiture</u>

Le marqueur <u>il y a Jean</u> est représenté dans (I') par :

$\langle_2\ \text{Sit}\ (\mathcal{S}_0, \mathcal{T}_0)\ \underline{\ni}\ a\ {}_2\rangle$

En effet, une analyse détaillée de *il y a* montrerait que, dans ce contexte, cette unité textuelle est la trace linguistique d'une localisation qui fait intervenir la situation d'origine Sit($\mathcal{S}_0$, $\mathcal{T}_0$) créée par l'énonciateur $\mathcal{S}_0$ en $\mathcal{T}_0$. La représentation choisie n'est donc pas arbitraire mais au contraire argumentable. L'écriture représentative a pour but de noter et de résumer les propriétés sémantiques du marqueur *il y a* dans ce contexte.

Le marqueur *qui* est la trace d'une identification entre une place vide dans la relation prédicative et un identificateur. Des considérations plus fines sur la constitution des énoncés en tenant compte notamment de la détermination des termes a et b montreraient que les conditions de représentation de *il* et de *qui* - que nous n'avons pas donné ici, laissant au lecteur le soin de les formuler - permettraient d'expliquer pourquoi en français

**Il y a Jean il conduit la voiture*

est inacceptable sans modulation prosodique spécifique entre *Jean* et *il* (...); de même ces conditions doivent éliminer :

**qui conduit la voiture, Jean*

qui est mal constitué donc inacceptable.

Les relations entre une formule métalinguistique, déduite d'une relation prédicative initiale au moyen de règles explicites et sa réalisation sous forme d'une suite textuelle attestable, doivent être données dans une composante (dite de réalisation). Réciproquement, tout énoncé reçoit une représentation métalinguistique au moyen de règles stables assignant une représentation à chaque marqueur.

Nous n'avons pas tenu compte dans l'exemple précédent de la constitution du nom propre *Jean*, du verbe conjugué *conduit*, du complément *la voiture* à partir des notions /Jean/, /conduire/, /voiture/, la constitution de ces catégories faisant appel à des régles différentes.

2 ..2 Considérons un autre exemple. Prenons l'énoncé japonais :

japonais :	Kazuko	wa	me	ga	ookii	(desu)
mot à mot :	Kazuko	particule	oeil yeux	particule	grand(s)	(copule)

équivalents français	Quant à Kazuko, ses yeux, ils sont grands En ce qui concerne Kazuko, ses yeux, ils sont grands Il y a Kazuko, ses yeux, ils sont grands Les yeux de Kazuko sont grands Kazuko a de grands yeux

Il nous faut donner une représentation métalinguistique aux marqueurs wa et ga dans ce contexte. Certes, il ne s'agit pas ici d'une représentation qui spécifierait une fois pour toute le statut général de wa et de ga. Nous cherchons seulement à représenter les occurrences de ces marqueurs et non à décrire toutes les valeurs associées à ces marqueurs (la description et la représentation théorique prendrait beaucoup plus de place).

Ce type d'énoncé a été qualifié de "phrase à double sujet ". On trouve des phénomènes analogues dans des énoncés français tels que *Il y a Jean, quand il est revenu, sa mère était déjà partie* ou encore *Jean, son frère, les mobylettes, il les répare drôlement bien.* Sur le problème d'énoncés de même type, dans d'autres langues, on pourra se reporter utilement, pour le breton, à Pierre Trépos, Grammaire bretonne, Rennes, p. 244; à Teng Shou-Hsin, "Double nominatives in Chinese", Language 50, (1974), p.455-473, pour le chinois; pour le cambodgien, à Judith M. Jacob, Introduction to Cambodian, Oxford University Press 1968, p. 148, où l'on trouve la remarque suivante: "in these last sentences, the independent noun construct represents neither the subject nor the object, but a general sentence topic, the chief point of interest in the sentence"; pour le japonais, à Susumu Kuno, The structure of the Japanese Language, MIT Press 1973 ; etc.

Comment engendrer l'énoncé japonais de façon à ce que la représentation métalinguistique obtenue permette de reconstruire (en partie au moins) "l'histoire" de la représentation et donc de reconstruire les relations prédicatives sous-jacentes ? Posons, sans nous poser de nouveau le problème de la constitution du nom propre et plus généralement des catégories nominales et verbales

a ≡ me (oeil, yeux)

b ≡ ookii (grand(s))

c ≡ Kazuko

Considérons les deux lexis

$\lambda_1 \equiv < a \underline{\in} b >$

$\lambda_2 \equiv < a \underline{\in} c >$

glosées respectivement en français par :

"oeil (yeux) sont grand(s)"

"des yeux sont à Kazuko";"Kazuko a des yeux"

La première lexis λ_1 est une attribution d'une propriété à un objet nominal, la seconde λ_2 une lexis de possession ramenée à une localisation (on sait que localisation et possession sont des phénomènes étroitement corrélés dans la plupart des langues, sinon toutes)

Considérons maintenant la lexis complexe :

$\lambda \equiv < \lambda_1 \underline{\in} \lambda_2 >$

qui établit une relation inter-lexis visant à privilégier la lexis λ_2 qui sert de repère à la lexis λ_1. Là encore, des considérations linguistiques montreraient que les relations interpropositionnelles (parce que, puisque, et, si...) sont représentées par les formules métalinguistiques dérivées de lexis complexes initiales mettant en jeu l'opérateur de repérage $\underline{\in}$.

Nous allòns omettre les chevrons externes pour simplifier les écritures des formules. Nous appliquons la règle d'orientation à λ_1 (pas 5), puis la règle 2 en introduisant un repère constitutif c (pas 6). Le schéma d'intrication est ensuite appliqué deux fois (pas 7 et 8), puis le schéma de règle d'absorption (pas 9). Nous en déduisons la formule II.

1 $\lambda \underline{\in} \mathrm{Sit}\ (\mathcal{S}_0, \mathcal{T}_0)$

2 $\lambda = <_0 \lambda_1 \underline{\in} \lambda_2 \ {}_0>$

3 $\lambda_1 = <_1 a \ \underline{\in} \ b \ {}_1>$

4 $\lambda_2 = <_2 a \ \underline{\in} \ c \ {}_2>$

5 $<_0 \ <_3 \ <_1 \overbrace{(\ ) \ \underline{\in} \ b \ {}_1> \ \underline{\in} \ a}^{=} \ {}_3> \underline{\in} <_2 a \ \underline{\in} \ c \ {}_2> {}_0> \underline{\in} \ \mathrm{Sit}\ (\mathcal{S}_0, \mathcal{T}_0)$

6 $<_0 <_3 <_1 \overbrace{(\ ) \underline{\in} b \ {}_1> \underline{\in} a}^{=} \ {}_3> \underline{\in} <_2 a \underline{\in} c \ {}_2> \ {}_0> \underline{\in} <_4 c \underline{\in} \mathrm{Sit}(\mathcal{S}_0, \mathcal{T}_0) {}_4> \underline{\in} \mathrm{Sit}(\mathcal{S}_0, \mathcal{T}_0)$

7 $<_0 <_3 <_1 \overbrace{(\ ) \underline{\in} b \ {}_1> \underline{\in} a}^{=} \ {}_3> \underline{\in} <_2 a \underline{\in} <_4 c \ {}_2> \ {}_0> \underline{\in} \mathrm{Sit}(\mathcal{S}_0, \mathcal{T}_0) {}_4> \underline{\in} \mathrm{Sit}\ (\mathcal{S}_0, \mathcal{T}_0)$

8 $<_0 <_3 <_1 \overbrace{(\ ) \underline{\in} b \ {}_1> \underline{\in} <_2 a}^{=} \ {}_3> \underline{\in} <_4 c \ {}_2> \ {}_0> \underline{\in} \ \mathrm{Sit}\ (\mathcal{S}_0, \mathcal{T}_0) {}_4> \underline{\in} \mathrm{Sit}\ (\mathcal{S}_0, \mathcal{T}_0)$

9 $<_0 <_3 <_1 \overbrace{(\ ) \underline{\in} b \ {}_1> \underline{\in} <_2 a}^{=} \ {}_3> \underline{\in} <_4 c \ {}_2> \ {}_0> \underline{\in} \ \mathrm{sit}\ (\mathcal{S}_0, \mathcal{T}_0) \ {}_4>$

II $\boxed{<_0 <_3 <_1 \overbrace{(\ ) \underline{\in} b \ {}_1> \underline{\in} <_2 a}^{=} \ {}_3> \underline{\in} <_4 c \ {}_2> \ {}_0> \ \underline{\in} \ \mathrm{Sit}\ (\mathcal{S}_0, \mathcal{T}_0) \ {}_4>}$

Cette formule métalinguistique peut être glosée en français par :

"Il est b, le a de c"

" ils sont b, les a de c"

ou en revenant aux termes linguistiques représentés par a,b,c :

<u>ils sont grands les yeux de Kazuko</u> .

Utilisons maintenant le schéma de règle de dualité aux sous- expressions suivantes :

$$<_4\ \dots\ _4>$$

$$<_3\ \dots\ _3>$$

puis finalement à l'expression :

$$<\ <_3\ \dots\ _3>\ \underline{\in}\ <_4 \dots\ _4>$$

Nous avons la déduction :

1	1	$<_4\ c\ _2>\ _0>\ \underline{\in}\ \mathrm{Sit}\ (\mathcal{S}_0,\mathcal{T}_0)\ _4>$
2	2	$<_4\ \mathrm{Sit}\ (\mathcal{S}_0,\mathcal{T}_0)\ \underline{\ni}\ <_0\ <_2\ c\ _4>$
3	1	$<_3\ <_1\ (\ )\ \underline{\in}\ b\ _1>\ \underline{\in}\ <_2\ a\ _3>$
4	2	$<_3\ a\ _2>\ \underline{\ni}\ <_1\ (\ )\ \underline{\in}\ b\ _1>\ _3>$
5		$<_0\ <_3\ a\ _2>\ \underline{\ni}\ <_1\ (\ )\ \underline{\in}\ b\ _1>\ _3>\ \underline{\in}\ <_4\ \mathrm{Sit}\ (\mathcal{S}_0,\mathcal{T}_0)\ \underline{\ni}\ <_0\ <_2\ c\ _4>$
6		$<_4\ \mathrm{Sit}\ (\mathcal{S}_0,\mathcal{T}_0)\ \underline{\ni}\ <_0\ <_2\ c\ _4>\ \underline{\ni}\ <_3\ a\ _2>\ \underline{\ni}\ <_1\ (\ )\ \underline{\in}\ b\ _1>\ _3>\ _0>$

Nous obtenons finalement la formule métalinguistique :

II' $\boxed{<_4\ \mathrm{Sit}\ (\mathcal{S}_0,\mathcal{T}_0)\ \underline{\ni}\ <_0\ <_2\ c\ _4>\ \underline{\ni}\ <_3\ a\ _2>\ \underline{\ni}\ <_1\ (\ )\ \underline{\in}\ b\ _1>\ _3>\ _0>}$

que l'on peut gloser en français par :

"il y a c, son a , il est b"

ou en revenant aux termes linguistiques :

<u>il y a Kazuko, ses yeux, ils sont grands</u>

Nous remarquons que les formules II et II' nous permettent de reconstituer les lexis λ_1 et λ_2 initiales. En effet, nous avons :

$$\breve{\lambda}_1 = <_2\ c\ \underline{\ni}\ a\ _2> \Rightarrow <_2\ a\ \underline{\in}\ c\ _2> = \lambda_1$$

en vertu de la dualité.

Par ailleurs de :

$$<_3 a\ \underline{\ni}\ <_1 (\ )\ \underline{\in}\ b\ _1>\ _3>$$

nous déduisons :

$$<_1 a\ \underline{\in}\ b\ _1>$$

en considérant que la première expression est une opération "à effectuer" alors que la seconde expression est obtenue après l'effectuation de l'opération, à savoir le remplacement de la place vide dans la relation prédicative $<_1 \ldots _1>$ par le terme 'a' qui lui sert d'identificateur. Nous noterons:

$$<_3 a\ \underline{\ni}\ <_1 (\ )\ \underline{\in}\ b\ _1>\ _3> \ \rhd\ <_1 a\ \underline{\in}\ b\ _1>$$

en utilisant le symbole '$\rhd$' de réduction entre, d'un côté "l'opération à effectuer" et, d'un autre côté, son résultat.

Remarquons également qu'aucune des règles utilisées n'a changé les ordres primitifs donnés par λ_1, λ_2 et λ. A chaque étape du calcul, on peut reconstituer l'ordre primitif. En effet, à l'étape 5 nous avons, avec les mêmes notations que précédemment :

$$<_3 <_1 (\ )\ \underline{\in}\ b\ _1>\ \underline{\in}\ a\ _3> \ \rhd\ <_1 a\ \underline{\in}\ b\ _1>$$

et la même chose à toutes les autres étapes.

Finalement, nous pouvons représenter l'énoncé japonais :

<u>Kazukowa me ga ookii (desu)</u> .

par la représentation formelle :

$$<_4 \text{Sit}\ (\mathcal{S}_0, \mathcal{T}_0)\ \underline{\ni}\ <_0<_2 c\ _4>\ \underline{\ni}\ <_3 a\ _2>\ \underline{\ni}\ <_1 (\ )\ \underline{\in}\ b\ _1>\ _3>\ _0>$$

'<u>Kazuko</u> <u>wa</u> <u>me</u> <u>ga</u> <u>ookii</u> (<u>desu</u>)

2.3 Les règles de réalisation apparaissent à ce stade comme de simples approximations qui ne règlent -comme nous l'avons dit plus haut- ni le problème général des valeurs de wa et de ga , ni le problème de leur relation avec les gloses françaises et anglaises telles que, en français : "il y a" ou en anglais : "as for".Seule, une argumentation fondée sur des observations minutieuses et un raisonnement rigoureux permettraient de traiter ce problème de façon satisfaisante.

Les formules précédentes appartiennent à une famille de formules dérivées d'un même schéma initial ; les étapes précédentes du calcul ne sont qu'un des calculs possibles car il est clair qu'il existe d'autres étapes d'un calcul qui conduiraient, de proche en proche, à engendrer toute une famille paraphrastique d'énoncés.

Nous avons défini une grammaire, ou plutôt un fragment de grammaire G, comprenant un ensemble fini et totalement ordonné de règles avec la contrainte d'enchaînement formulée ainsi :

"appliquer dans l'ordre la première règle, puis la seconde éventuellement et réitérer la troisième règle autant de fois qu'il est possible puis procéder aux absorptions et aux permutations par dualité".

2.4. L'engendrement d'une formule métalinguitique s'effectue au moyen de règles de déduction ou de réécritures de la forme :

$$F_g \rightarrow F_d$$

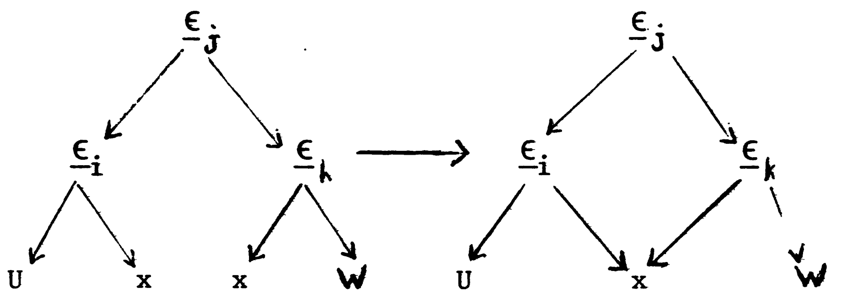
où F_g et F_d sont deux formules.

Chaque formule est représentable par un graphe : le graphe de l'agencement constitutif de la formule.

Une règle est alors une transformation entre graphes. Il se trouve que les graphes des agencements constitutifs ne sont pas arborescents.

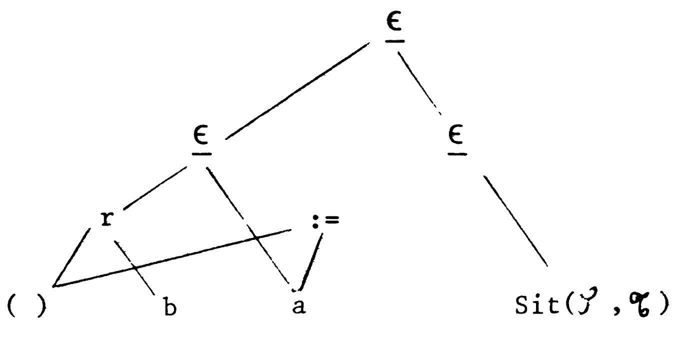
Ce ne sont pas des graphes quelconques. Ces graphes constituent une sous-classe des graphes sans circuits dont les sommets de chaque niveau sont totalement ordonnés. L'application d'une règle consiste à rechercher, dans un graphe donné, un sous-graphe qui représente la partie gauche de la règle pour remplacer ensuite ce sous-graphe par le sous-graphe représentatif de la partie droite.

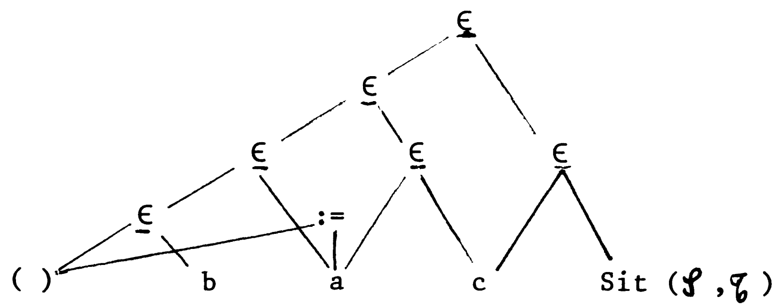
Dans le cas qui nous intéresse ici, les règles (1), (2) et (3) sont des règles d'expansion. Seule la règle (4) d'intrication présente quelques difficultés puisque la présentation linéaire de cette règle est incomplète. Une formulation plus rigoureuse de cette règle s'exprime en effet par la transformation entre graphes (sans circuits):



La partie droite de la règle permet de noter l'intrication de x dans deux relations: $< U_i \ \underline{\in}_i \ x_i >$ et $< x_k \ \underline{\in}_k \ \ _k >$.

Les graphes (ou treilles) obtenus par l'application des transformations de graphes et qui représenteraient les formules (I) et (II) sont visualisés par respectivement :





Cette grammaire nous a permis d'engendrer par un calcul déductif une formule métalinguistique F. L'examen de cette formule permet de "reconstituer" une "histoire de constitution" en reconstruisant les lexis primitives. Tout co-énonciateur (ou auditeur) doit procéder, à la réception de l'énoncé, à cette reconstruction pour retrouver les lexis primitives qui portent en elles-mêmes une certaine information. On remarquera au passage que la signification d'un énoncé n'est pas réductible à cette information puisque toute énonciation impose une modulation paraphrastique porteuse elle aussi de signification.

Une autre grammaire de réalisation assure la correspondance entre les formules engendrées par la première grammaire et des occurrences de marqueurs grammaticaux. Lors de la compréhension de l'énoncé, le co-énonciateur doit procéder à une reconnaissance automatique en attribuant un statut métalinguistique aux marqueurs grammaticaux et en recomposant une formule métalinguistique.

Une présentation plus technique qui préciserait les conditions d'application d'une règle permettrait de mieux décrire l'engendrement formel d'un énoncé et d'une famille d'énoncés.

En effet, chaque transformation est en fait une opération définie sur l'ensemble des treilles. Toute grammaire ou ensemble d'opérations potentielles peut être formulée en terme de grammaire de graphes. Il s'agit d'une classe spécifique de grammaire de graphes qu'il nous faudra étudier. Nous y reviendrons dans un prochain article. Nous avons seulement cherché à présenter ici le concept d'intrication et à l'illustrer par un exemple linguistique précis.

3. CONCLUSION

Il est temps de nous effacer en demandant à d'autres chercheurs de conclure. D'un côté, J.Cl. Chevalier, dans son magistral ouvrage sur La notion de complément chez les grammairiens (Droz, 1968), nous rappelle (page 22) l'importance de la construction théorique:

"On jugera peut-être ce programme trop ambitieux; nous ne le croyons pas. Nous avons voulu tenter une description largement synthétique qui permît de saisir à quel prix une notion pouvait être élaborée et une telle démarche concerne toute l'histoire des sciences; (...)

On critiquera sans doute aussi telle ou telle carence de notre exposé; mais nous nous rallions entièrement à la critique très juste qu'établissait récemment A.J. Greimas de l'état d'esprit naïf qui désire des inventaires exhaustifs (...); ce n'est pas tant l'impossibilité pratique qu'il faut souligner que la faute de méthode; l'exhaustivité et son corollaire, la statistique, supposent que la théorie sortira tout armée des faits et que le seul rôle du chercheur sera un rôle de classificateur. Nous tenterons de montrer que l'équilibrage des paramètres suppose une analyse plus profonde qui se justifie par sa seule cohérence; les faits ne sont que des émergences qui permettent de vérifier des hypothèses appuyées sur des rapprochements d'ensembles; aussi les avons-nous disposés de façon fort inégale, tantôt insistant sur le développement immanent de la grammaire, sur l'effort des grammairiens pour résoudre des contradictions inhérentes au système lui-même, tantôt nous avons montré qu'un système ne s'établissait que dans un autre système, qu'un assemblage de faits, si bien établi soit-il, ne prenait sa cohésion que quand les termes d'une problématique nouvelle étaient posés.

Ce phénomène nous a paru extrêmement sensible en grammaire; on a pu voir pendant des décennies des grammairiens accumuler les faits et les inventaires sans qu'il en ressorte rien d'autre qu'une monstrueuse complication. Le grand grammairien est celui qui édifie un système fondé sur les faits à l'intérieur d'une épistémologie rigoureuse; les solutions peuvent être diverses (il serait d'un positivisme naïf de croire qu'une seule "se dégage" des faits)."

De l'autre, M. Gross et A. Lentin écrivaient dès 1967, dans leur ouvrage Notions sur les grammaires formelles (Gauthier-Villars) (page 192) :

"Dans ces conditions, qui ne sont que des conditions normales pour une science jeune, il n'est pas facile de voir quelle sera la nature des études abstraites de modèles de langues naturelles, mais indéniablement les essais systématiques extrêmement instructifs, même quand les résultats sont négatifs, si bien que le linguiste doit être conduit à mener de front les études abstraites et les études empiriques".

Nous nous efforçons de respecter ce double conseil, notamment en conduisant des analyses systématiques et approfondies sur des langues nombreuses et variées. Car, faut-il le répéter, nous n'oublions pas que la linguistique est une science d'observation, où l'on ne saurait opposer le théoricien et l'enquêteur sur le terrain. On comprendra donc que le présent exposé ne cherche pas à être manifeste, mais un appel raisonné à une entreprise collective, où il n'y aura jamais assez d'efforts convergents.

BIBILOGRAPHIE

CHURCH, A., The calculi of lambda conversion, Ann. of Math. Studies, 6 Princeton, N.J., 1941.

CULIOLI,A. & DESCLES,J.P., Traitement formel des langues naturelles, première partie : mise en place des concepts, Math. et Sciences Humaines, 1982.

CURRY,H.B., & Alii, Combinatory logic, NorthHolland Publishing Company, 1958.

GROSS,M. & LENTIN,A., Notions sur les grammaires formelles, Gauthier-Villars, Paris, 1967.

KOULOUGHLI,D.E., "Sur la phrase nominale en arabe classique. Contribution à l'étude formelle de la catégorie de "mubtada ?"", Rapport PITFALL 43, Paris, 1976, à paraître dans T-a Informations 1, Klincksieck, 1982.