

J. B. GRIZE

Quelques problèmes logico-linguistiques

Mathématiques et sciences humaines, tome 35 (1971), p. 43-50

http://www.numdam.org/item?id=MSH_1971__35__43_0

© Centre d'analyse et de mathématiques sociales de l'EHESS, 1971, tous droits réservés.

L'accès aux archives de la revue « Mathématiques et sciences humaines » (<http://msh.revues.org/>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
<http://www.numdam.org/>

QUELQUES PROBLÈMES LOGICO-LINGUISTIQUES

par

J. B. GRIZE ¹

En présence d'une discipline comme la linguistique, un logicien peut apporter une aide précieuse au spécialiste de la discipline en tentant de construire un système formel dont l'interprétation sera jugée acceptable par le spécialiste (voir l'article « l'attitude formalisante en linguistique ») et en cherchant à définir l'instrument logique le plus intéressant pour entreprendre une lecture rigoureuse et fructueuse des faits à expliquer, lecture qui peut révéler certains problèmes importants qui étaient auparavant passés inaperçus. Toutes les sciences, surtout celles qui s'appuient sur un empirique observable, se posent le problème des « unités ». La linguistique a besoin d'unités. Mais quelles unités prendre et comment les définir ? Les grammaires traditionnelles admettent l'existence de catégories linguistiques (noms, verbes, adjectifs, articles...). L'auteur, à la suite de K. Adjukiewicz et Y. Bar-Hillel, et en employant comme eux l'instrument de la « logique combinatoire », veut expliciter et construire le réseau de relations entre les différentes catégories linguistiques (phrase, nom, verbe, protoverbes, opérateurs verbaux, déterminants, adjectif, adverbe) avec l'intention de participer à la construction d'une grammaire de reconnaissance.

J. P. Desclés

I. INTRODUCTION

En présence d'un corps de faits, le logicien peut se proposer deux tâches. Ou bien il tente de formaliser ces faits, c'est-à-dire qu'il construit un système formel propre à leur servir de modèle, ou bien il se sert de la logique comme d'un instrument connu qui doit lui permettre une lecture plus rigoureuse des faits.

Les linguistes, et même déjà les grammairiens traditionnels, ont réuni une riche collection de faits relatifs au langage et c'est quelques-uns d'entre eux qui vont servir d'objet à cette étude. Il existe en effet des chaînes parlées et écrites qu'il est possible de traiter comme des unités et dont on peut se proposer de reconnaître l'organisation. Plusieurs directions sont possibles. On peut partir des chaînes observées et, au moyen d'une grammaire de reconnaissance, on tente d'en rendre exhaustivement compte. C'est cette dernière perspective que je vais adopter et elle conduit d'emblée à poser deux questions :

- 1) Comment savoir ce dont on peut rendre compte ?
- 2) Comment déterminer la nature de l'instrument à utiliser ?

1. Centre de Recherches Sémiologiques, Université de Neuchâtel.

1) Dans un article récent, A. Culioli rappelle que ni les structures de surface, ni les données naïves ne sont formalisables [1] ; il s'ensuit, d'une part qu'il serait prématuré de viser la construction d'un système formel et, d'autre part, qu'il faut tirer parti des concepts théoriques qui existent déjà. Or, toutes les grammaires s'accordent sur ce point qu'il existe des catégories linguistiques, sans lesquelles la langue ne fonctionnerait même pas. Ce sont elles qui serviront de point de départ non naïf à cette étude.

2) Quant à l'instrument, ce sera celui de la fonctionnalité [2]. On sait en effet, par des travaux comme ceux de K. Adjukiewicz [3] ou de Y. Bar-Hillel [4], par exemple, qu'il permet une analyse fructueuse. On a montré par ailleurs l'équivalence (au sens faible) entre les « grammaires catégorielles » de Y. Bar-Hillel et les grammaires « indépendantes du contexte » de N. Chomsky [5] et [10].

K. Adjukiewicz et Y. Bar-Hillel partent de deux catégories primitives à l'aide desquelles ils reconstruisent les autres. L'intérêt d'un tel travail est indéniable, en ce sens que, d'une part, il explicite tout un réseau de relations entre les catégories de la langue et que sa rigueur permet, d'autre part, de mettre le doigt sur certaines particularités du fonctionnement des langues naturelles.

Soucieux dans ce qui suit d'apporter une certaine contribution au problème de la reconnaissance du français :

- 1) Je traiterai du français ;
- 2) Je distinguerai deux types de catégories primitives : des *catégories d'objets* et des *catégories d'opérateurs*.

II. DESCRIPTION DE L'INSTRUMENT D'ANALYSE

Donnons-nous des *objets* que nous désignerons par les lettres a, b, c , etc., ainsi qu'une opération qui applique un objet sur un autre. On aura donc des expressions de la forme : (ab) , $((ab) c)$, ..., $(a (bc))$, $((a (bc)) d)$, etc.

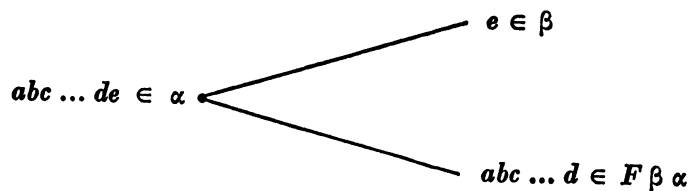
Convenons, pour éviter l'encombrement des parenthèses, de supprimer éventuellement les couples de parenthèses qui convergent à gauche. Les expressions ci-dessus pourront donc s'écrire : ab , abc , $abcd$, ..., $a (bc)$, $a (bc) d$.

Donnons-nous ce que nous appellerons des *catégories*, que nous désignerons par les lettres α, β, γ , etc., ainsi que la relation $\in$ telle que les objets et les chaînes d'objets peuvent être en relation $\in$ avec les catégories. On aura par exemple : $a \in \alpha$, $a (bc) \in \beta$, etc.

Soit enfin un opérateur F , appelé *flèche*, et qui engendre de nouvelles catégories selon la règle suivante :

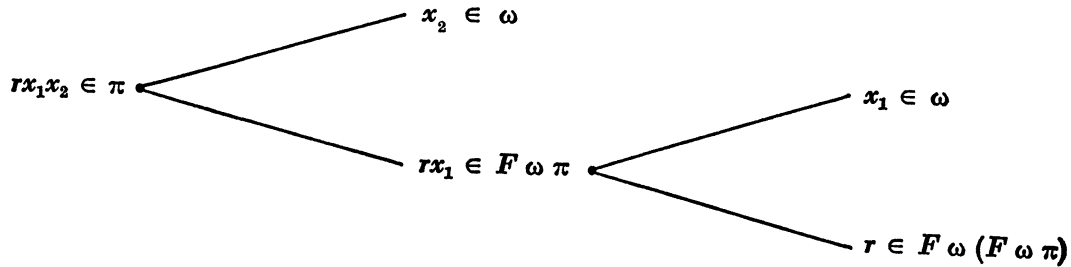
Règle

Si $abc \dots de \in \alpha$ et si $e \in \beta$, alors $abc \dots d \in F \beta \alpha$. Il est commode d'écrire la règle sous la forme suivante qui offre l'avantage d'engendrer un arbre :

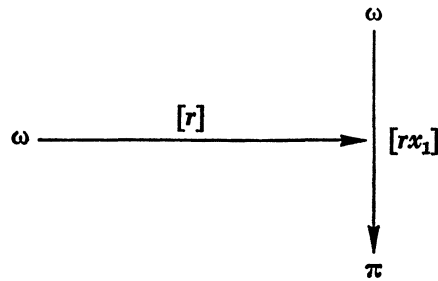


Exemple

Si l'on transcrit dans le formalisme de la logique des prédicats la proposition « l'objet x_1 est en relation r avec l'objet x_2 », on obtient la suite rx_1x_2 . Soit π la catégorie des propositions et ω celle des objets. Une double application de la règle ci-dessus donne :



Si l'on désigne par $[X]$ la catégorie à laquelle appartient l'objet X et si l'on représente l'opérateur F par une flèche, on aura le diagramme suivant :



L'objet rx_1 appartient à la catégorie $F \omega \pi$, catégorie des *prédicats unaires* et l'objet r à la catégorie $F \omega (F \omega \pi)$, catégorie des *prédicats binaires*.

III. POSTULATS ET CATÉGORIES PRIMITIVES D'OBJETS

L'analyse qui précède n'a été possible que parce que l'écriture logique jouit d'un certain nombre de propriétés, dont les quatre suivantes :

- 1) On connaît la catégorie des propositions et il existe un algorithme pour décider si une expression donnée est ou n'est pas une proposition.
- 2) Une proposition est composée d'objets distinguables.
- 3) On connaît la catégorie à laquelle appartiennent certains des objets.
- 4) Il existe un ordre canonique pour écrire une proposition.

Il ne sera donc possible de procéder à une analyse des phrases que si des conditions analogues sont satisfaites. C'est ce à quoi servent les six postulats suivants.

Postulat 1. Toute chaîne de la langue est composée d'objets distinguables.

En pratique ces objets seront des mots, au sens naïf.

Postulat 2. Il est toujours possible de décider si une suite donnée d'objets est ou n'est pas une phrase. La notion de phrase recouvre pratiquement celle de « phrase minimale » [6]. Elle correspond à l'idée

d'expression bien formée et il est commode de ne pas l'appeler « proposition » pour réserver ce terme à l'être logique correspondant. Il est clair d'autre part que la procédure de décision qui est postulée ne peut, à ce moment du travail, que reposer sur la compétence des locuteurs. Ceci ne porte nullement sur la cohérence de l'analyse, mais uniquement sur sa pertinence.

Notion 1. La catégorie primitive P des *phrases*.

Pour continuer, il faut tenir compte d'une différence fondamentale entre l'écriture logique et l'écriture du français. En logique, un même signe ne peut jamais appartenir à deux catégories, ce qui n'est pas le cas en français. Ainsi, selon la fonction que le signe « touche » aura au sein d'une phrase, il sera verbe ou substantif. Il est donc indispensable de postuler qu'il est possible d'assigner des fonctions définies aux objets et aux suites d'objets.

Postulat 3. Les objets d'une phrase peuvent être groupés selon deux types de fonctions : une *fonction prédicative* R et quatre *fonctions complétives* C_0 , C_1 , C_2 et C_3 .

Postulat 4. Les objets et groupes d'objets C_1 appartiennent à une même catégorie, dite catégorie des noms.

Notion 2. La catégorie primitive N des *noms*.

Si *Pierre sait le russe* et *Pierre sait que Jean dort* sont reconnues comme des phrases, alors *le russe* et *que Jean dort* sont des noms. Il est évident qu'ils ne le sont pas de la même façon et, en particulier, que le sens de *sait* diffère d'un exemple à l'autre. Ceci ne fait toutefois que manifester l'effet de sens que le cadre syntaxique confère à un objet linguistique.

Les postulats qui précèdent, et en particulier le deuxième, vont permettre de placer par commutation des objets attestés dans diverses catégories. Ainsi, à partir de la phrase *Pierre sait le russe*, pourra-t-on encore placer dans la catégorie N , non seulement *Pierre* et *le russe*, mais *mon ami Pierre*, *l'anglais d'Oxford*, etc. On pourra aussi placer dans une catégorie α , pour l'instant inconnue, *sait*, *ignore*, *sait à peu près*, etc. Le problème n'est pas bien entendu de donner les catégories en extension. Il consiste simplement à en reconnaître l'existence et à les introduire, soit comme de nouvelles catégories primitives, soit en les ramenant à des catégories déjà connues.

Il reste encore à poser deux conventions d'ordre qui doivent assurer le bon fonctionnement de l'instrument d'analyse.

Postulat 5. L'analyse adoptera l'ordre canonique $RC_3C_2C_1C_0$.

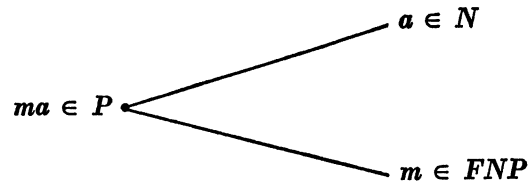
La permutation de l'ordre de la chaîne linguistique, que Bar-Hillel, lui, respecte, ne porte pas ici à conséquence. D'une part, en effet, les fonctions sont supposées connues et d'autre part, dans la perspective adoptée, le problème de la génération ne se pose pas.

Postulat 6. Les objets d'une catégorie à analyser seront écrits avant ceux d'une catégorie connue.

Comme l'opérateur F est binaire, l'ordre des objets des catégories connues sera déterminé de proche en proche. Notons encore que les postulats 5 et 6 sont compatibles entre eux.

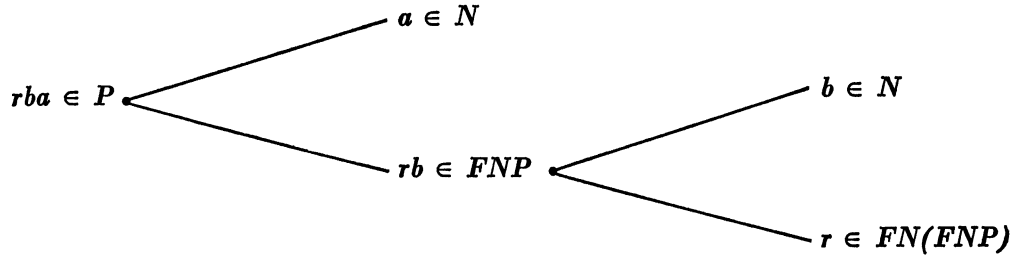
IV. VERBES ET PROTOVERBES

Les verbes sont les objets ou groupes d'objets susceptibles d'assumer la fonction prédicative R . Soit alors la phrase *Albert marchait*. Si nous convenons dès maintenant d'utiliser les initiales des mots pour les abrégés, les postulats 2, 3 et 5 permettent d'affirmer que $ma \in P$. Il vient alors :



Notion 3. $V_1 = \text{df } FNP = \text{df}$ catégorie des *verbes à une place*.

Soit encore la phrase *Albert regarde Brigitte*. Nous aurons :



Notion 4. $V_2 = \text{df } FN(FNP) = \text{df}$ catégorie des *verbes à deux places*.

Il est facile de généraliser en posant :

Notions 5. $\begin{cases} V_1 = \text{df } FNP \\ V_n = \text{df } FNV_{n-1}. \end{cases}$

Ceci définit la classe des catégories de verbes.

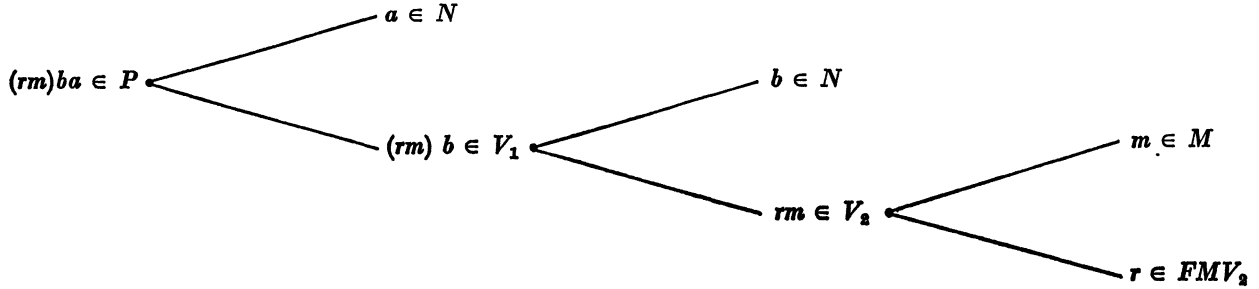
Remarquons aussi qu'un même objet peut appartenir à plus d'une catégorie. Ainsi, les trois objets *regardait* dans les trois phrases *Albert regardait*, *Albert regardait Brigitte* et *Albert regardait Brigitte dans les yeux* appartiennent respectivement à V_1 , V_2 et à V_3 . Contrairement donc à ce qui se passe en logique, l'appartenance à une catégorie n'est pas « context-free ».

Le traitement des verbes attributifs et de quelques autres phénomènes du français ne permet pas de s'arrêter ici, même au sein de ce qui ne veut être qu'une esquisse. Partons de ce que ni *Albert regarderont*, ni *Aristote philosopher* ne sont des expressions bien formées (des phrases). Une façon de faire face à ce problème consiste à se donner une catégorie d'opérateurs qui, appliqués à ce que je nommerai des protoverbes, engendreront des verbes.

Notion 6. La catégorie primitive des *opérateurs verbaux* M .

Les éléments de M sont traités comme des opérateurs inconnus, mais il n'est pas impossible d'en définir certains, en sortant naturellement du cadre de cet article. C'est ainsi, par exemple, que F. Bugniet [7] définit formellement des opérateurs d'orientation du prédicat et des opérateurs d'affirmation et de négation. Les premiers permettent de rendre compte de la distinction entre *regarde* et *est regardé* et les seconds de la différence entre *regarde* et *ne regarde pas*. Ces opérateurs ne suffisent évidemment pas et il en faudra en tout cas d'autres pour les modes, les temps et les aspects. De plus, il faut encore définir une loi de composition interne. Ce n'est pas mon propos et, dans ce qui suit, je me contenterai d'appeler m le résultat de la composition des opérateurs qui, à partir d'un protoverbe comme, par exemple REGARD ou PHILOSOPH permet de rendre compte d'objets observés comme *regardait* ou *est philosophe*.

Soit alors de nouveau la phrase *Albert regard-ait Brigitte*. Conformément au postulat 6 et au sens de m , nous l'écrivons $(rm)ba$. Il vient alors :



Notion 7. Les catégories des *protoverbes* V_i^* , telles que $V_i^* = \text{df } FMV_i$.

Ceci permet, en principe, de traiter des verbes attributifs. On rendra compte de *Pierre est courageux* à l'aide du protoverbe $c \in V_1^*$ par une composition d'opérateurs m et de *Pierre semble courageux* par une composition m' dans laquelle aura figuré un opérateur du type « en apparence ».

Il est aussi possible de distinguer le sens littéral et le sens figuré de phrases du genre *Jean prend la porte*. Si nous posons $x = \text{df } la\ porte$, on aura :

sens littéral : $(pm)xj \in P$ et $m \in V_2^*$,
sens figuré : $(px)mj \in P$ et $px \in V_1^*$.

Cela revient finalement à considérer que *prend-la-porte* appartient à la même catégorie que *prend*.

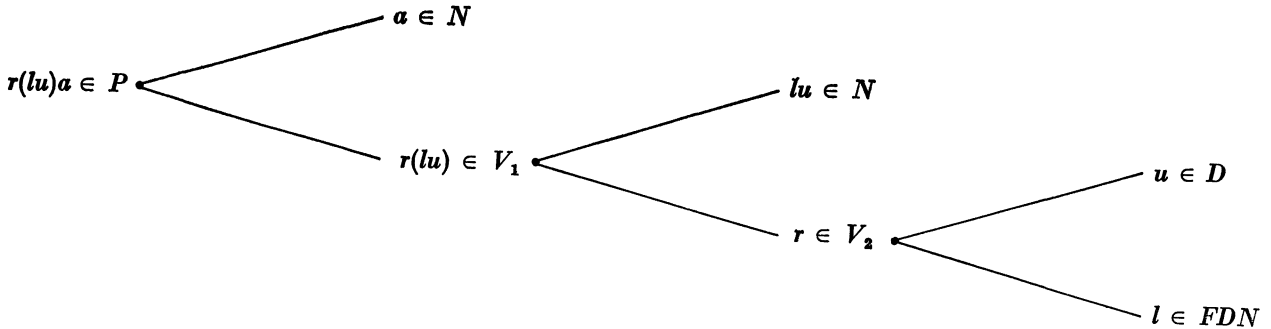
V. NOMS ET SUBSTANTIFS

De même qu'il y a intérêt à analyser les verbes comme résultant de l'application de certains opérateurs à des protoverbes, de même il est intéressant de concevoir que certains noms résultent de la détermination de ce que j'appellerai des substantifs par des opérateurs. Ni les uns, ni les autres ne sont encore entièrement connus, mais leur introduction explicite a au moins l'avantage de poser le problème.

Partons de l'expression *Albert regarde livre*. Le postulat 2 permet de dire que ce n'est pas une phrase et que, en conséquence, *livre* $\notin N$. En revanche, *un livre, ce livre, le livre dont je vous ai parlé* sont des noms. Quant au problème des déterminants, s'il n'est pas encore entièrement résolu, il est cependant en bonne voie et il semble que des opérateurs d'extraction, de fléchage, de parcours combinés entre eux ne sont pas loin d'en rendre compte [8]. Quoi qu'il en soit, j'introduirai la catégorie d'opérateurs D , dans laquelle entreront, au niveau de la langue, les articles, les adjectifs démonstratifs et les adjectifs numéraux.

Notion 8. La catégorie primitive des *déterminants* D .

Soit alors la phrase *Albert regarde un livre*. On aura, en n'introduisant pas, pour simplifier, le préverbe : $r(lu)a$. Et il vient :



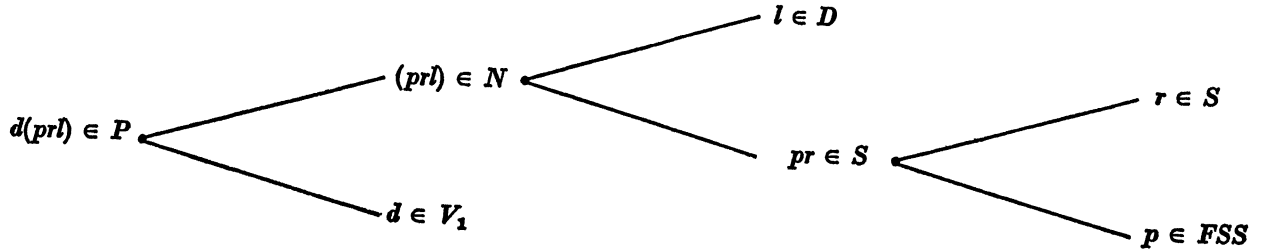
Notion 9. La catégorie des *substantifs* S , telle que $S = \text{df } FDN$.

Il faut d'ailleurs s'attendre à trouver en D d'autres objets encore et, en particulier, des adjectifs que je n'ai pas mentionnés. Ainsi, puisque *petit ruisseau deviendra grand* est une phrase, *petit* dans ce contexte sera élément de D . Il ne le sera cependant pas dans le contexte *le petit ruisseau a débordé*.

QUELQUES CATÉGORIES DÉRIVÉES

Il semble que les catégories primitives P , N , M et D puissent suffire à pousser l'analyse assez loin. Elles permettent en tout cas de définir des catégories d'adjectifs et d'adverbes.

Reprenons la phrase *le petit ruisseau a débordé* sans, pour simplifier, nous intéresser au protoverbe. Conformément au postulat 6, nous écrirons $d(prl)$. Il vient :

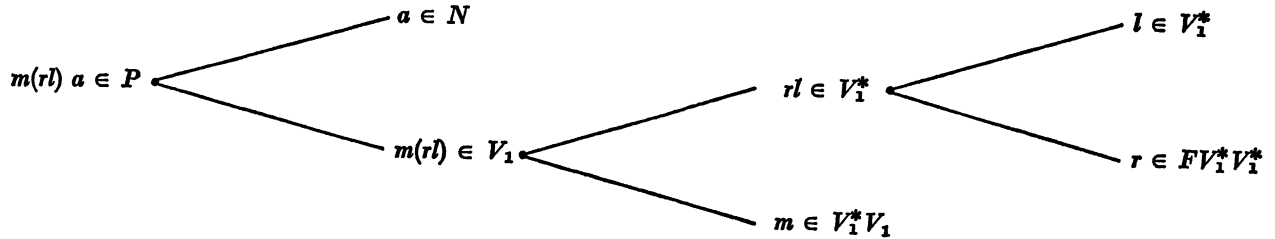


Notion 10. La catégorie des *adjectifs* $AS = \text{df } FSS$.

Il faut noter que dans *petit Jean*, l'objet *petit* appartiendra à une autre catégorie adjectivale, la catégorie $AN = \text{df } FNN$. Il s'agit là, une fois encore, d'effets de sens très fins qui résultent de la syntaxe.

Si l'on convenait de distinguer des unités plus petites que les mots, certains suffixes appartiendraient aux catégories AS et AN . Ainsi, le suffixe *-ette*, qui permet de passer du substantif *maison* au substantif *maisonnette* et du nom *Pierre* au nom *Pierrette*.

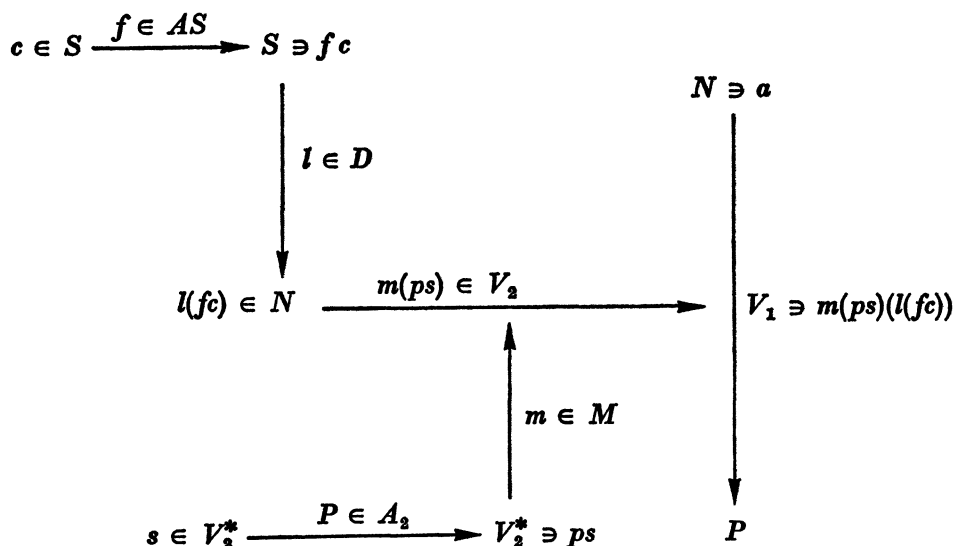
Partons enfin de la phrase *Albert lisait rapidement*. On aura $m(rl)a$ et il vient :



Notion 11. Les catégories des *adverbes* A_i , telles que $A_i = \text{df } FV_i^*V_i^*$.

Notons qu'il existe d'autres types d'adverbes. Ainsi l'objet *très* dans le groupe *très rapidement* est élément de la catégorie FA_iA_i et dans le groupe *très joli* de la catégorie $F(AS)(AS)$.

L'exemple suivant fait usage de toutes les catégories introduites et peut servir de résumé. *Alice salua poliment le chapelier fou*, soit $m(ps) (l (fc)) a$.



On remarquera, pour terminer, que les catégories sont indifféremment représentées par des flèches et par des sommets et que les termes qui s'introduisent dans les lexis [9] sont les objets formés d'une seule lettre et qui appartiennent à l'union des catégories N, S, AN, AS, V_i^* et A_i^* .

BIBLIOGRAPHIE

- [1] CULIOLI, A., FUCHS, C., PÊCHEUX, M., *Considérations théoriques à propos du traitement formel du langage*. Centre de Linguistique quantitative de la Faculté des Sciences de l'Université de Paris, n° 7, 1970.
- [2] Voir l'exposé formel de CURRY, H. B., FEYS, R., *Combinatory logic*, Amsterdam, North-Holland, 1968.
- [3] AJDUKIEWICZ, K., "Die syntaktische Konnexität", *Studia Philosophica*, 1, 1935, pp. 1-27.
- [4] BAR-HILLEL, Y., "Une notation quasi arithmétique destinée aux descriptions syntaxiques", *Langages*, 9, 1968, pp. 9-22. (L'article original a paru en anglais dans *Language*, 29, 1953, pp. 47-58.)
- [5] CHOMSKY, N., MILLER, G. A., *L'analyse formelle des langues naturelles*, Paris, Gauthier-Villars et Mouton, 1968.
- [6] Cf. par exemple, DUBOIS, J., *Grammaire structurale du français : Le verbe*, Paris, Larousse, 1967.
- [7] BUGNIET, F., *Remarques sur les notions d'assertion linguistique et de proposition logique*. Centre de Recherches Sémiologiques de l'Université de Neuchâtel, n° 4, 1970.
- [8] et [9] Voir l'étude [1] ci-dessus.
- [10] BAR-HILLEL, Y., GAIFMAN, G., SHAMIR, E., "On categorial and phrase-structure grammars", *Bulletin of the research council of Israel*, June 1960, Vol. 9, n° 1.