

TRAVAUX DU CENTRE DE RECHERCHES SÉMIOLOGIQUES

Colloque sur l'analyse du discours

« DIVERGENCES ET CONVERGENCES »

Yanouchka Opper (éd.)

No 19 — Février 1974

UNIVERSITÉ DE NEUCHÂTEL
Centre de recherches
sémiologiques

Archives

UNIVERSITÉ DE NEUCHÂTEL

Centre de recherches sémiologiques
Avenue du Premier-Mars 26
2000 Neuchâtel (Suisse)

COLLOQUE SUR L'ANALYSE DU DISCOURS

"Divergences et Convergences"

Yanouchka OPPEL (éd.)

No 19 - février 1974

TABLE DES MATIERES

page

0. INTRODUCTION	1
Liste des participants	3
<u>TEXTES et PROPOS:</u>	
I. L'AAD, par Catherine FUCHS et Michel PECHEUX	4
1. C. FUCHS: Quelques constatations	5
2. M. PECHEUX: Sur l'AAD	9
3. M. PECHEUX: L'expérience Mansholt (propos recueillis par Y. Oppel)	12
4. C. FUCHS: Problèmes de l'analyse syntaxique (propos recueillis par Y. Oppel)	14
II. RECHERCHES LIEES A L'ANALYSE AUTOMATIQUE DU DISCOURS (A.A.D.), par l'équipe TAL sous la direction de Jacques ROUAULT	17
III. REMARQUES SUR LES RAPPORTS ENTRE STRUCTURE PROFONDE ET PROPOSITION, par Narciso PIZARRO	26
IV. ESSAI D'ANALYSE DE DISCOURS POUR METTRE EN EVI= DENCE DES REPRESENTATIONS ECONOMIQUES, par Pierre VERGES	33
V. APERCU DE L'ETAT ACTUEL D'ELABORATION D'UNE CON= CEPTION DE LA THEORIE DU TEXTE, par Janos S. PETÖFI	42
VI. POSSIBILITES D'UNE ANALYSE DISTRIBUTIONNELLE D'UN TEXTE THEATRAL, par Patrice THOMPSON	51

0. INTRODUCTION

Les 29 et 30 juin 1973, le Centre de Recherches Sémiologiques a réuni dans ses locaux plusieurs chercheurs⁽¹⁾ venant de pays et d'horizons théoriques divers. Le but de cette rencontre était de confronter des démarches différentes en analyse du discours, et, si possible, de dégager leurs points de convergence et de divergence.

Le colloque, ouvert par Jean-Blaise GRIZE, présente d'emblée plusieurs sources possibles de divergence. En effet, une différenciation peut se produire

- I. dans le choix de l'angle d'attaque:
 1. la linguistique (pour la plupart)
 2. la logique
 3. la rhétorique (une rhétorique généralisée)
- II. dans celui des voies d'approche:
 1. analyse formelle
 2. analyse matérialiste
- III. dans celui des visées:
 1. description de l'organisation de certains éléments textuels ou discursifs s'inscrivant dans une typologie du discours;
 2. mécanismes de fonctionnement, par exemple l'argumentation;
 3. explication des problèmes de signification visant une théorie du discours.

Il me semble que la dichotomie la plus tranchée est celle qui - grosso modo - oppose ceux qui ne consi-

(1) Cf. Liste des participants, p.3

dèrent que le texte en lui-même et ceux qui incluent dans leur champ d'observation les facteurs déterminant les textes et les effets qui en résultent. D'autres oppositions ont été relevées par J.-B. GRIZE - à qui incombait la difficile tâche de faire le point à la fin de cette rencontre -, parmi lesquelles on peut citer:

- un texte vs une famille de textes, comme objet d'étude;
- le texte vu comme un objet vs un procès;
- le texte étudié - pour lui-même (théorie textuelle)
 - pour autre chose - pour une théorie du discours
 - pour une autre théorie (politique, idéologique, p.ex).

C'est autour du rôle et des possibilités de la linguistique que le débat a été souvent le plus ardent.

Dans ce cahier⁽¹⁾, nous présentons les textes qui ont constitué la base des discussions, à part celui de Pierre FIALA et Charles RIDOUX qui a fait l'objet du cahier no 17.⁽²⁾ Remarquons toutefois que les points de vue rapportés et, plus particulièrement les exemples utilisés, sont loin de représenter tous les points de vue des auteurs ou des équipes auxquelles ils appartiennent - je pense spécialement à Patrice THOMPSON et au CdRS. Le temps disponible nous a contraints à quelques aspects seulement.

Nous exprimons ici nos très vifs remerciements à tous les participants du colloque et aux auteurs des textes que voici.

Yanouchka OPPEL

(1) Dans un prochain cahier collectif, nous prévoyons de donner notre point de vue sur les démarches discutées au cours de ce colloque.

(2) P. FIALA, Ch. RIDOUX, Essai de Pratique Sémiotique. Travaux du CdRS, no 17, 1973, Univ. de Neuchâtel.

Liste des participants au colloque "Divergences et Convergences"

Mmes	M. DUPRAZ	(Grenoble)
	C. FUCHS	(Paris)
	F. NOËL	(Montréal)
MM.	P. BIZARD	(Grenoble)
	J.-J. COURTINE	(Grenoble)
	P. FIALA	(Paris/Neuchâtel)
	J.-B. GRIZE	(Neuchâtel)
	M. PÊCHEUX	(Paris)
	J. PETÖFI	(Bielefeld)
	N. PIZARRO	(Montréal)
Ch.	RIDOUX	(Paris)
	J. ROUAULT	(Grenoble)
	E. ROULET	(Neuchâtel)
	P. THOMPSON	(Neuchâtel)
	P. VERGÈS	(Aix-en-Provence)

et les membres du Centre de Recherches Sémiologiques.

I. L' AAD

par Catherine FUCHS et Michel PECHEUX

1. C. FUCHS : Quelques constatations
2. M. PECHEUX : Sur l'AAD
3. M. PECHEUX : L'expérience Mansholt
4. C. FUCHS : L'analyse syntaxique

Catherine FUCHS "Quelques constatations"

1) La sémantique constitue à l'heure actuelle la préoccupation dominante des linguistes, des théoriciens du discours, des utilisateurs d'analyse textuelle, ...et aussi le lieu d'affrontement théorique le plus "chaud".

2) Entre la position selon laquelle "tout ce qui est sémantique est linguistique" et celle selon laquelle "rien de ce qui est sémantique n'est linguistique", s'échelonnent une série de positions intermédiaires, dont l'un des soucis principaux consiste à définir les rapports entre syntaxe et sémantique.

3) Toutes ces positions ont en commun la caractéristique suivante: feignant réalisé un consensus sur ce que l'on entend par "sémantique", on se demande de quel champ théorique elle constitue l'étude; à qui revient la sémantique? qui a droit d'en parler? qui est le mieux armé pour le faire?

Faisons à présent l'hypothèse que la problématique ainsi formulée est mal posée, que la sémantique n'est pas un objet bien connu qu'il s'agirait de "caser" et de décrire quelque part, mais qu'elle constitue bien plutôt, en creux, une série de questions que pose, à diverses disciplines, l'existence de l'activité de langage en tant que conduite signifiante spécifiquement humaine. Nous ne pouvons plus dès lors esquiver des interrogations comme: le langage ainsi conçu constitue-t-il l'objet de la seule linguistique? comment s'articulent le champ de la linguistique et ceux des autres "sciences sociales"? en définitive, qu'est-ce qui est constitutif des phénomènes sémantiques?

A partir de là, nous préciserons notre position, qui consiste à refuser:

- 1) toute insularité de la linguistique, où l'objet d'étude serait conçu comme un tout homogène, découpable en niveaux successifs, du phonologique au sémantique (inclus), et les mêmes méthodes d'analyse (ex.: décomposition des unités en traits) transposables de niveau en niveau. Nous nous opposons par conséquent à une conception de la sémantique comme une organisation systématique (un réseau) de concepts universels et a priori qui "traduirait" le monde. Nous pensons qu'il s'agit là d'une position idéaliste qui méconnaît le rôle essentiel des facteurs socio-historiques; sur ce point nous sommes d'accord avec la thèse de Pêcheux selon laquelle "le lien qui relie les 'significations' d'un texte aux conditions socio-historiques de ce texte n'est nullement secondaire, mais constitutif des significations elles-mêmes."

- 2) toute dichotomie du type: code (Objet de la linguistique)/ utilisation du code (objet d'une autre "science humaine"). Ceci vise aussi bien:
 - l'opposition langue/parole de Saussure, où la linguistique se voit attribuer, au plan sémantique, l'étude de la valeur, tandis que la signification renvoie à l'acte de parole du sujet individuel
 - l'opposition compétence/performance de Chomsky qui introduit indûment l'étude de la signification au niveau du code conçu comme universel (cf. point 1), la contrepartie du sujet idéal étant encore le sujet individuel au compte de qui sont mis les "écarts par rapport à la norme".
 - d'une certaine façon l'opposition langue (stock lexical + systèmes de règles syntaxiques)/ processus de production du discours (constitutifs des significations de Pêcheux). En effet, si cette position a l'énorme mérite d'éliminer les notions du sujet idéal, sujet individuel, créativité, au profit de systèmes de représentation liés à des positions de classe, nous pensons néanmoins que d'une part sa conception de la langue (objet de la linguistique) est réductrice et mécaniste, et que d'autre part l'élimination

du sujet entraîne une dichotomie factice là où la prise en compte des phénomènes d'énonciation permettrait une articulation plus dialectique au coeur même des processus sémantiques.

Expliquons ce dernier point:

- 1/ Nous entendons par "énonciation" un certain nombre d'opérations, ayant en surface des traces plus ou moins explicites, par lesquelles tout énoncé est toujours repéré par rapport à un système de points d'ancrage que nous appellerons "situation d'énonciation", constitué par le "moi-ici-maintenant" du sujet énonciateur. Ces opérations ont pour effet de faire porter sur l'énoncé une série de déterminations (modalités, assertion, aspects, voix, thématization, déterminants,...). Pour nous, parler de délinéarisation du texte, c'est dire que la distinction fondamentale ne passe pas entre "structures profondes" et "structures de surface", mais entre schéma non encore déterminé (pré-assertif) et énoncé muni de déterminations et qu'une méthode d'analyse qui, à un niveau "profond" se donne déjà les correspondants de tous les éléments qui figurent en surface, est incapable de rendre compte des processus par lesquels les énoncés se constituent et se déterminent progressivement à la fois par rapport à la situation d'énonciation et entre eux. Ajoutons par parenthèse que l'étude du langage infantin montre bien que dès les premières productions, l'énoncé comporte (explicitement ou non) un "actualisateur" qui met en situation.
- 2/ Quel statut théorique attribuons-nous à la situation d'énonciation? Nous parlerons d'"illusion nécessaire". Expliquons-nous:
Nous pensons - qu'il s'agit d'un concept nécessaire pour l'analyse des phénomènes linguistiques (ce n'est pas une création mythique du linguiste),
- mais qu'il demande à être articulé tout à la fois sur une théorie des processus de production des

discours et sur une théorie du sujet.

Pour l'analyse linguistique, le "moi-ici-maintenant" constitue un point de référence stable; si cette stabilité est mise en question, c'est l'activité même de langage qui se trouve atteinte. Par contre, du point de vue d'une théorie du sujet, par exemple, ou d'une théorie des idéologies, il est clair que cette stabilité apparaît comme une illusion, par laquelle le sujet se place comme centre de référence, et qui joue un rôle de masque idéologique.

En résumé, mes travaux portent donc sur les trois régions suivantes:

- 1) étude linguistique des phénomènes d'énonciation (aspects, temps, thématization, voix,...) et élaboration théorique du cadre conceptuel de l'énonciation, dans la perspective de A. Culioli;
- 2) étude de l'articulation entre le plan de l'énonciation et le plan discursif (par exemple sur le problème précis des relatives);
- 3) étude critique des développements récents de la sémantique générative et de la sémantique interprétative, visant à tracer une ligne de partage, parmi l'hétérogénéité des phénomènes "sémantiques" abordés par ces deux courants, entre d'une part ce qui me semble relever de l'énonciation, et d'autre part ce que l'on ne saurait prétendre, selon moi, traiter sur des bases linguistiques: les effets de sens produits à partir des systèmes de représentation dont les énonciateurs sont les supports.

Michel PECHEUX : Sur l'A.A.D.

Il apparaît nécessaire aujourd'hui de faire une sorte de bilan des recherches théoriques et pratiques menées sous le sigle "AAD", et de caractériser à la fois le point où ces recherches sont parvenues, les difficultés qui subsistent (pas seulement des difficultés de détail mais aussi des "questions de fond" qui menacent l'ensemble de l'entreprise), et les orientations nouvelles.⁽¹⁾

L'enjeu qui est sous-jacent à cette réflexion est triple. Il y va en effet de questions d'orientation décisives qui sont les suivantes:

1) Le tout premier enjeu concerne la détermination d'une orientation correcte dans la théorie des idéologies, pour éviter le faux-fuyant idéaliste par lequel la théorie des idéologies se verrait réduite à une théorie des systèmes de pensée (tels qu'ils apparaissent dans des paroles et des écrits), ce qui aurait pour effet à peu près inévitable de superposer langue et "idéologie" par l'intermédiaire d'une hypothèse du type SAPIR-WHORF. Le terme de discours semble, de ce point de vue, avoir l'avantage de permettre (je ne dis pas d'impliquer) la position d'un rapport (ou du moins de commencer à le faire) dans la mesure où le discursif représente "la linguistique" à l'intérieur des processus idéologiques et inversement (cf. sur ce point l'article HENRY, HAROCHE, PECHEUX de Langages no 24, 1971).

(1) cf. un texte récent ronéotypé: C. FUCHS, M. PECHEUX, Analyse du discours et Linguistique. Mises au point et perspectives à propos de l'AAD, octobre 1973 (Note de Y. Opperl).

2) Le deuxième enjeu, c'est celui d'une orientation théorique correcte en linguistique, avec les conséquences pratiques de cette orientation: il semble que, sur ce point, les questions décisives concernent à la fois les conceptions de la sémantique d'une part, de l'énonciation d'autre part, et que les conséquences pratiques portent essentiellement sur le rapport entre syntaxe, sémantique et énonciation, au niveau de l'analyse concrète d'un texte ou d'une famille de textes. Là encore, les conceptions idéalistes de l'énonciation, qui conduisent à superposer discours et parole dans la liberté du sujet parlant face au "système" de la langue, doivent être "mises hors d'état de nuire", ce qui, on le verra, n'est pas toujours facile.

3) Enfin le troisième et dernier enjeu consiste à observer et contrôler les effets produits au niveau de la lutte (à la fois théorique et politique) qui traverse aujourd'hui des disciplines comme l'Histoire ("la révolution française a-t-elle eu lieu?"), la sociologie ou l'économie politique, dans leur rapport avec l'analyse des textes. Ce dernier enjeu va par exemple se cristalliser autour du livre récent de Régine ROBIN "Histoire et Linguistique", et des thèses matérialistes qui y sont soutenues.

Ces trois enjeux ont en fait déterminé le plan de la présente communication, en déterminant les deux points suivants, sur lesquels il semblait important de s'expliquer.

a) Le premier point concerne l'existence concrète actuelle de l'AAD: il semble important de montrer que, avec ses limites actuelles, ce dispositif d'analyse n'est pas resté un pur projet inapplicable en pratique et dépourvu d'applications concrètes. Il s'agit ici de donner aux chercheurs qui s'intéressent à ces questions des "pièces à conviction" à partir desquelles la distinction se fasse dans leur esprit entre les mises en cause qui contribuent à permettre une nouvelle percée, et celles qui constituent

en fait un repli et une régression. La première partie de cet exposé sera pour cette raison consacrée à la présentation d'une analyse concrète, qu'on peut résumer par le titre "Expérience Mansholt".

b) Le deuxième point porte sur la pratique de l'AAD, et en particulier sur les présupposés de cette pratique: contre l'idée d'une pratique définitivement figée dans un rituel qui reflèterait un dogme (j'emploie ces mots à dessein!), il apparaît utile certes d'énumérer les changements qui ont été réalisés, et ceux qui sont prévus, mais aussi de désigner les transformations de nature globale qui peuvent et doivent être envisagées. Ce point fera l'objet de la deuxième partie de l'exposé, dans laquelle Catherine FUCHS exposera le contenu de récentes mises au point linguistiques concernant la phase dite "d'analyse syntaxique".

Michel PECHEUX: L'expérience Mansholt (Propos recueillis
par Y. Oppel)

Cette expérience a débuté lors d'une réunion à Grenoble. Son but était de confronter différents angles d'attaque, différentes méthodes d'analyse du discours. Plusieurs groupes tenteraient l'analyse d'un même texte: la lettre de Mansholt.

L'hypothèse de base du groupe AAD consiste dans l'idée qu'il est possible d'étudier l'effet d'une position politique au niveau de la production et de la lecture d'un texte à partir des phénomènes discursifs observables dans ce texte. Il n'existe pas de processus discursif unique. Selon QUI énonce, les mots diffèrent ou changent de sens, ainsi que les configurations de mots, c'est-à-dire la formation, au niveau du syntagme, des groupes d'énoncés. On peut observer, entre autres, des effets de diffraction, de création d'images divergentes à partir d'une source unique. Il s'agit alors d'étudier les mécanismes et les contenus - ceux-ci étant liés - en relation avec les prises de position politiques des sujets.

Trois hypothèses sont formulées, se référant chacune à l'ambiguïté:

1. Le texte choisi doit être politiquement ambigu pour pouvoir donner lieu à une double lecture, en présentant divers points d'identification possibles. Pour la lettre de Mansholt, cette ambiguïté a été vérifiée (il s'agit du point de vue de la Sozial-Democratie).
2. Pour pouvoir filtrer les prises de position gauche/droite, il faut choisir des sources politiquement plausibles (la PCF et l'UDR, par exemple, ne l'auraient pas été). On a présenté le texte pour la droite comme étant l'extrait d'un rapport d'un groupe d'études de républicains giscardiens, représentants de la majorité; et, pour la

gauche, comme extrait d'un rapport CFDT.

3. Les sujets soumis à l'expérience doivent avoir une position politique ambiguë qui les rende capables d'obéir aux deux fonctionnements (en l'occurrence, des cadres en cours de recyclage hésitant, par conséquent, sur leur position dans la société).

La consigne est la suivante:

Le texte est lu par les sujets. On leur demande d'en faire un résumé, en insistant sur l'aspect "documentaire", dans le but de les contraindre à présenter des résumés objectifs s'en tenant à la littéralité du texte, en excluant, en principe, au niveau conscient les prises de position.

Puis on applique l'AAD aux deux corpus de textes recueillis.

Catherine FUCHS : Problèmes de l'analyse syntaxique
(propos recueillis par Y. Oппel)

Dans la phase constituée par l'analyse syntaxique de l'AAD, la pratique actuelle rencontre plusieurs points d'achoppement, sur lesquels la linguistique semble en défaut:

1. Le terme de discours est ambigu, passible de plusieurs interprétations:
 - a. dans un sens limité, il désigne un texte entier, pris comme un tout, dans lequel, on étudie les mécanismes inter-phrastiques;
 - b. il est utilisé comme synonyme de "parole", en faisant appel à l'intervention du sujet, renvoyant par là aux phénomènes de l'énonciation. Le risque couru, c'est d'en faire une dimension supplémentaire ajoutée à la langue;
 - c. dans le sens de M. Pêcheux, il désigne des formations discursives constituant des ensembles de textes, des discours à partir d'une matière première: la langue.

2. L'analyse syntaxique de l'AAD pose un certain nombre de problèmes. La première difficulté se rencontre dans l'établissement du point de départ. Le choix de la phrase est délicat (p.ex.: problème de l'automatisation des relations anaphoriques et des contrastes de temps), mais il semble difficile de la remplacer par autre chose. Au lieu de la phrase, l'AAD part des propositions. C'est ce que fait traditionnellement la grammaire. Mais le choix repose sur un sémantisme implicite. Le travail consiste alors à construire une représentation en énoncés élémentaires d'après un schéma fixe, canonique, qui permet la comparaison entre énoncés, en rétablissant les anaphores (encore manuellement pour l'instant).

F	d ₁	N ₁	V	Adv.	P	d ₂	N ₂
---	----------------	----------------	---	------	---	----------------	----------------

Le schéma ci-dessus est formé de 8 cases. Les catégories représentées par ces cases offrent une certaine hétérogénéité: tantôt l'analyse est faite en termes de parties du discours, tantôt en termes de fonctions; ce qui signifie que l'analyse formelle repose sur une sémantique non-explicitée. Il en va de même pour les connections entre schémas. De fait, aucune syntaxe n'est neutre et il serait intéressant de dégager la sémantique implicite des catégories grammaticales traditionnelles. Dans le cas de l'AAD, le schéma des énoncés élémentaires se ramène à une situation de type SVO enchaînée mécaniquement par les connecteurs interpropositionnels.

Le problème de l'adverbe: l'adverbe offre la possibilité de fonctions syntaxiques multiples. Il peut modifier plusieurs points de la proposition. A cet endroit, il faudrait pouvoir avoir recours à une théorie des adverbes.

Le problème de la préposition: là aussi, il faudrait disposer d'une théorie plus élaborée des verbes et des types de compléments. On pourrait, par exemple, partir des verbes comme "atomes" et, procédant de proche en proche dans la détermination des éléments qu'ils conditionnent, faire éclater l'unité de la phrase.

Mise à part l'absence d'univocité qui rend souvent l'analyse difficilement contrôlable (les sujets ne coïncident pas de la même façon), l'hétérogénéité, on vient de le voir, pose le problème de fond d'un sémantisme implicite qui conditionne un postulat de base de l'AAD, l'idée que les effets de sens sont avant tout d'origine lexicale. Ceci revient à poser une délimitation nette séparant:

LEXIQUE (stock lexical) / SYNTAXE (règles de combinaison)

Il serait donc possible de traverser la syntaxe - une syntaxe prétendument neutralisée - pour faire surgir les effets de

sens. Mais la frontière est-elle vraiment si bien tracée? (ex. le passage de fonction-sujet à NOM évacue le problème de l'énonciation. De même le contenu de la case F ne sert qu'à coder des marques et n'explique rien sur la fonction).

Si on tient compte du fait que la syntaxe contribue aux effets de sens, il faut alors rendre compte des phénomènes d'énonciation qu'il est possible de formuler en termes d'opération, en caractérisant avec précision ce que l'on entend par énonciation, et par opération ; ce qui permettrait d'explicitier la sémantique de cases telles que d_1 par exemple.

Il faut se donner d'autres unités que S V O pour ancrer les opérateurs d'énonciation. On peut alors traiter comme des systèmes d'opération en relation les uns avec les autres les phénomènes de modalité, d'aspect et de temps, de thématization et de voix, les déterminants (au sens large) et les personnes.

Il est primordial, non pas de dresser des listes, mais de construire un système. Par ailleurs, il faut se demander sur quel type de schéma abstrait appliquer un tel système. ⁽¹⁾

(1) Cf. le travail cité en note 1 p.9, où l'on nous propose (p. 52-55) une reconfiguration de l'énoncé élémentaire aboutissant à une suite de trois éléments N_1VN_2 (N_2 facultatif), et à lui associer un groupe FD_1D_2 . On aura donc $(Fd_1 d_2)+(N_1VN_2)$. En poursuivant, la notion de proposition se trouve éclater en: "schéma morpho-syntaxique" et "schéma d'énoncé" (A R D) dérivé du premier (N_1VN_2). (Note de Y.O.).

II. RECHERCHES LIEES A L'ANALYSE AUTOMATIQUE DU
DISCOURS (A.A.D.)

par l'équipe TAL sous la direction de
Jacques ROUAULT

Equipe de traitement automatique des langues: P. BIZARD,
J.J. COURTINE, M. DUPRAZ et J. ROUAULT (directeur).

Recherches liées à l'Analyse Automatique du Discours (A.A.D.)

Dans le cadre de l'équipe, l'Analyse du discours constitue à la fois une voie de recherche autonome et le champ d'application de procédures d'analyse linguistique mises au point parallèlement (mais pas tout à fait indépendamment...).

Nous examinons successivement dans ce rapport les travaux déjà réalisés, les problèmes de l'analyse linguistique et l'étude de la méthode A.A.D. elle-même.

I - TRAVAUX REALISES

1- La première tâche a consisté à écrire entièrement un programme d'A.A.D. dans le but (1) de connaître dans ses moindres détails la procédure algorithmique, (2) de se livrer à des expériences sur des corpus que nous aurions réunis.

En fait l'écriture du programme nous a conduits à une étude algorithmique de la méthode et nous a amenés à modifier très sensiblement l'algorithme lui-même. Disons de façon très schématique que l'algorithme que nous avons réalisé favorise l'axe syntagmatique alors que le programme réalisé à Paris favorise, lui, l'axe paradigmatique.

Les résultats de ce travail sont consignés dans le document : M. DUPRAZ. (*)

2- Parallèlement à ce travail nous avons étudié les différentes méthodes d'analyse de contenu, leur position par rapport à la linguistique et leur validité théo-

(*) "Caractéristiques du programme "AADP" d'analyse automatique du discours écrit en Algol W, d'après la méthode de M. PECHEUX=

rique. Les conclusions de cette étude critique figurent dans le document suivant : J.J. COURTINE (*)

et nous ont confirmé dans l'idée que l'A.A.D. était la seule méthode théoriquement fondée qui puisse s'intégrer dans le travail de l'équipe.

3- Expérimentation sur plusieurs corpus

- a) Deux expériences prévues depuis le début n'ont pu être menées suffisamment loin pour que l'on puisse dire que les deux problèmes ont avancé de façon substantielle. Il s'agit
- 1/ d'une tentative d'étude rigoureuse des textes de présentation des résultats d'une analyse factorielle. Le but poursuivi est le suivant: rechercher les critères qu'utilisent les expérimentateurs pour dépouiller ces résultats. Pour cela soumettre de façon indépendante à plusieurs personnes entraînées les données et les résultats d'une analyse factorielle, faire produire à chacune de ces personnes un texte d'interprétation des résultats et soumettre ces textes à l'A.A.D. Malheureusement, nous n'avons pu jusqu'ici réunir un corpus de textes de ce type.
 - 2/ Une autre expérience que nous souhaiterions poursuivre est l'étude des questionnaires comportant à la fois des questions ouvertes et des questions fermées. Peut-on imaginer des méthodes de dépouillement utilisant à la fois l'analyse factorielle (pour les questions fermées) et l'A.A.D. (pour les questions ouvertes) et établissant certaines corrélations entre ces deux types de résultats ?

Un début d'étude de ce problème a été mené à partir des réponses au questionnaire élaboré par A. BOLLE, historien à l'Université des Sciences Sociales de Grenoble. Le questionnaire était consacré au "Jour du Seigneur" dans les paroisses protestantes du Sud-Est. Malheureusement les réponses ouvertes que nous avons analysées à l'aide de l'A.A.D. correspondaient à une question trop

(*) "Linguistique et analyse du discours".

vague, de sorte que les résultats n'étaient pas probants.

b) Une expérience purement A.A.D. a été réalisée à partir de textes publicitaires d'eau minérale parus dans des journaux féminins. Les conclusions de l'analyse automatique paraissent cette fois probants et interprétables. Les résultats figurent dans le document de J.J. COURTINE, cité plus haut.

c) Deux autres expériences sont en cours; elles portent:

1/ Sur un corpus recueilli en situation expérimentale et déjà traité à Paris (corpus "Manshold"). La comparaison des résultats du programme "Paris" et du programme "Grenoble" doit nous apporter un certain nombre de renseignements sur les différences entre les deux méthodes.

2/ Sur les réponses à une enquête faite auprès des élèves de classes terminales de Lycées Grenoblois; la question était "d'exprimer l'image que vous vous faites des Sciences Sociales".

Une première partie des réponses a déjà été soumise à l'A.A.D. et a fourni des résultats suffisamment intéressants pour que le corpus soit augmenté de façon importante.

Un document est prévu sur ce sujet pour juin 1974.

4- Critique de l'A.A.D.

A la lumière des expériences réalisées nous avons pu dresser une liste des insuffisances, imprécisions et incohérences de la méthode A.A.D.

a) La démarche linguistique souffre de certains défauts de codage auxquels on peut remédier dans une optique à court terme. Mais elle souffre aussi d'insuffisances liées à la forme du codage et donc à l'analyse linguistique -et ces défauts ne peuvent être corrigés que par une refonte de la méthode. Parmi ceux-ci citons la forme rigide de l'"énoncé": les huit places sont en fait un codage issu de

la théorie de la lexis (faisant intervenir un schéma à trois places). Cette optique n'est justifiée du point de vue opératoire que si l'on se donne, entre autres, un moyen linguistiquement fondé de décomposer les verbes comme "donner" en une combinaison de schémas à trois places. Ce qui ne nous semble pas être totalement le cas actuellement.

De plus la "forme" de l'énoncé a un statut incertain et ne se réfère qu'à des marques de surface difficilement interprétables.

- b) Le point de vue algorithmique est à modifier en ce qui concerne notamment le choix des objets d'entrée (cf. a), le choix des proximités et la création des domaines et hyper-domaines.
- c) Enfin la validité des résultats n'est pas assurée au sens suivant:
- Comment varient les résultats en fonction de la taille du corpus?
 - Quelle est l'influence du fait que l'on choisit (ou non) de comparer les énoncés à l'intérieur d'un même discours?
 - etc...

II - LE PROBLEME DE L'ANALYSE LINGUISTIQUE

Remarquons d'abord que le problème de la place de la linguistique dans la méthode A.A.D. est réel et qu'il n'est peut-être pas encore totalement résolu. Remarquons aussi que nous sommes, en gros, d'accord avec les principes posés dans (FUCHS-PECHEUX 1973). En particulier nous reconnaissons l'importance des opérations d'énonciation dans la théorie linguistique. Ce qui implique que l'on définisse l'unité minimale d'énonciation ainsi que les liens possibles entre de telles unités.

L'unité minimale d'énonciation (ou énoncé) s'articule autour du verbe et, à ce sujet, deux problèmes se posent :

- (1) Celui de la structure syntaxique d'un énoncé, ce qui revient à mettre en course l'énoncé actuel à 8 places et à leur trouver une structure plus souple, par exemple sous forme d'un arbre dont la racine (étiquetée par le verbe) a un nombre variable de descendants.

Ceci suppose une étude précise sur le comportement syntaxique de chaque verbe, sur la clarification des compléments, etc...

Ce travail a été réalisé et formulé sous la forme d'une hypothèse dans: Equipe de Traitement Automatique des Langues: "Noyau verbal et classes de verbes".

D'autres études sont en cours sur le groupe nominal et le groupe adverbial.

Par ailleurs un programme d'analyse syntaxique est en cours de réalisation pour :

- a) tester les hypothèses linguistiques
- b) réaliser l'analyse syntaxique d'une phrase.

- (2) L'autre problème est l'étude théorique des marques d'énonciation ainsi que le repérage de ces marques de surface.

Sans entrer dans les détails de travaux qui absorbent la plus grosse part de l'activité de l'équipe, disons que nous nous sommes orientés peu à peu vers une "linguistique expérimentale" que l'on peut résumer ainsi: nous avançons un certain nombre d'hypothèses linguistiques. Elles sont choisies en fonction d'un but déterminé (par exemple analyser -en le définissant- le "noyau verbal") et des impératifs propres au traitement automatique des langues. Ces hypothèses sont ensuite testées sur un corpus déterminé de façon à arriver à un processus simple et efficace de reconnaissance automatique de certains phénomènes linguistiques.