



Etudes sur les représentations économiques de personnes en formation

Arlette Apkarian-Lacout, Pierre Vergès, Pierrette Vergès

► To cite this version:

Arlette Apkarian-Lacout, Pierre Vergès, Pierrette Vergès. Etudes sur les représentations économiques de personnes en formation. [Rapport de recherche] C.N.R.S. 1972, pp.80. halshs-03639862

HAL Id: halshs-03639862

<https://shs.hal.science/halshs-03639862>

Submitted on 13 Apr 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

C. N. R. S.

**Action Thématique Programmée
sur les méthodes pédagogiques
dans l'initiation économique des adultes**

A. T. P. 6101

E. R. A. 284

ETUDES SUR LES REPRESENTATIONS ECONOMIQUES DE PERSONNES EN FORMATION

Rapport méthodologique

Le 31 Décembre 1972

**Arlette LACOUT
Pierrette VERGES
Pierre VERGES**

C.N.R.S.

Action Thématique Programmée
sur les méthodes pédagogiques
dans l'initiation économique des adultes
A.T.P. 6101
E.R.A. 284

ETUDES SUR LES REPRESENTATIONS ECONOMIQUES
DE PERSONNES EN FORMATION

RAPPORT METHODOLOGIQUE

Le 31 Décembre 1972

Arlette LACOUT
Pierrette VERGES
Pierre VERGES

Parler méthode est à la fois nécessaire et insatisfaisant. Un rapport méthodologique, qui décrit l'enchaînement entre une théorie générale et des outils très particuliers, utilise un mélange de langages à la fois divers et spécifiques. Il n'y a rien de commun entre les algorithmes de programmation et l'argumentation conceptuelle.

Ce rapport a un double objet : d'une part donner les éléments qui rendent notre démarche cohérente, c'est-à-dire la chaîne des hypothèses qui petit à petit orientent les outils employés; d'autre part donner les détails pratiques et les traitements informatiques qui conditionnent la forme des résultats que les chercheurs ont à interpréter. Le plan de ce rapport découle de cette démarche, et se particularise de plus en plus au fur et à mesure de la définition de méthodes particulières.

Nous présentons d'abord le cadre théorique qui oriente notre approche méthodologique. La première application de ce cadre se situe au niveau de la détermination du corpus et du recueil des données. Sa seconde application se situe au niveau de l'analyse du discours des personnes interviewées. Les hypothèses intermédiaires posées dans la présentation de l'analyse du discours nous conduisent à une méthode de transcription permettant la constitution d'une base de données. Enfin nous présentons les programmes permettant l'interrogation de ces données.

PLAN DU RAPPORT

1	CADRE THEORIQUE ET METHODE	p. 6
1.1.	Les représentations sociologiques	p. 7
1.2.	Représentation et Pratique économique	p. 11
2	LES PROBLEMES DE METHODE POSES PAR LA CONSTITUTION DU CORPUS	p. 12
2.1.	Le plan de sondage	p. 12
2.2.	Le recueil des données	p. 18
3	L'ANALYSE DU DISCOURS	p. 26
3.1.	Les prémisses linguistiques	p. 27
3.2.	Implication d'une démarche sociologique	p. 30
3.3.	Choix nécessaires et étapes de l'analyse	p. 32
4	LA METHODE DU SYSTEME DE TRANSCRIPTION ..	p. 37
4.1.	Système documentaire, codage T.R.T.	p. 37
4.2.	La double situation d'un terme dans la chaîne syntaxique et dans la suite des enchaînements	p. 40
4.3.	Les enchaînements de sens	p. 48
4.4.	Les traitements informatiques : codage et programme de création de la base de données	p. 51
5	METHODE ET PROGRAMME D'INTERROGATION DE LA BASE DE DONNEES LIES AUX PREMIERES HYPOTHESES	p. 62
5.1.	Traitement du vocabulaire	p. 62
5.2.	Traitement des arbres	p. 69
5.3.	Conclusion sur le traitement informatique	p. 74

FICHE SYNTHETIQUE SUR LE CORPUS

GROUPEs	Nbre d' entret.	Nbre d' arbres	Nbre mots ds texte codé (environ)	Nbre rela- tions ma- jeures et ex. mini- maux (environ)
Agriculteurs	18	610	24.000	9.500
Cadres diplômés	9	312	15.000	6.400
Cadres autodidactes	10	340	18.000	7.700
Ouvriers C.G.T.	10	325	15.900	6.000
Ouvriers C.F.D.T.	9	307	15.700	6.800
Ouvriers entreprise	9	293	11.000	4.700
Travailleurs sociaux	9	250	10.500	4.300
TOTAL	74	2.437	110.100	45.400

Le recueil : Le recueil s'est effectué à Lyon, Paris, Aix-en-Provence, Marseille, et dans la Drôme. L'ensemble des enregistrements sur bande magnétique fait environ 75 heures. Ils ont nécessité 430 heures de décriptage, puis 750 heures de frappe pour environ 2.800 pages dactylographiées.

Le codage : a duré environ 2.000 heures et a donné lieu globalement à la perforation de 31.150 cartes.
La vérification du codage a duré 200 heures et a fait perforer 3.000 cartes.

FICHE SYNTHETIQUE SUR LES PROGRAMMES

(écrits en FORTRAN)

	<u>Nbre de cartes</u>	<u>Temps sur 370/165</u>	<u>"Taille mémoire"</u>
<u>Opération de transcription</u>			
PG1 - Création du vocabulaire de base	150	12 sec	72 K
PG2 - Mise sur bande des données par TRI standard (PA012) .	10	1 sec/1000 cartes	100 K
PG3 - Sélection des entretiens à traiter et introduction de modifications	100	6 sec/1000 cartes	30 K
PG4 - Création de la base de donnée et écriture des arbres	1350	20 sec/1000 cartes	182 K
<u>Opération de traitement de la base de donnée</u>			
PG5 - Distribution d'un terme dans son contexte minimal	750	35 sec/1000 cartes	140 K
PG6 - Constitution du fichier inverse : les arbres pour un terme donné	120	7 sec/1000 cartes	200 K
PG7 - Distribution statistique des termes par entretien et groupe d'entretiens....	80	8 sec/1000 cartes	30 K
PG8 - Distribution statistique de la succession des rela- tions	140	7 sec/1000 cartes	50 K
PG9 - Analyse en Treillis des différents champs	300	30 sec	110 K
PG10- Analyse des champs séman- tiques	160	20 sec	95 K

1. CADRE THEORIQUE ET METHODE

Pour situer le cadre théorique où s'inscrit notre méthode de recherche, nous reprenons rapidement quelques éléments d'analyse qui sont largement développés dans la première partie du Rapport d'Etude. Nous renvoyons donc à ce rapport pour tous les développements et les éclaircissements concernant les représentations.

Définir une méthodologie permettant l'analyse des "représentations économiques", renvoie aux divers sens donnés à ce terme de "représentations". Dans le rapport d'étude nous analysons la définition donnée par les formateurs et la demande sociale qui a, implicitement, motivé cette étude. Nous essayons ensuite de mettre en place les éléments qui nous ont conduits à affirmer que les représentations ne sont pas la première étape d'une analyse qui aboutirait, par condensations (1) successives à la mise en évidence de types idéologiques c'est-à-dire de vision de la société. (Dans la suite de ce rapport nous désignerons cette démarche comme la théorie des "représentations-idéologies".)

Loin de mettre l'idéologie au terme d'une analyse qui débiterait par des représentations, il nous a paru nécessaire de postuler, au contraire, l'existence, au sein de la structure sociale, d'un processus de création des représentations. Ces dernières sont alors considérées comme les produits, les

(1) Condensation = résumé qui permet de condenser un certain nombre de termes en un seul, même quand il s'agit de mots qui ne sont pas tout à fait synonymes, mais qui ont la même signification dans le discours.

traces de ce processus et l'organisation des représentations ne peut que renvoyer à ce "processus-idéologie". Il faut immédiatement remarquer que notre champ d'analyse et le matériel utilisé ne nous permettent pas d'atteindre le niveau de ce processus de création des représentations au sein de la structure sociale. Ce n'est d'ailleurs pas notre objectif d'atteindre ce niveau. Seules les analyses théoriques de la structure sociale se donnent un tel objectif.

Cette prise de position n'est, cependant, pas neutre vis-à-vis de la méthode que nous avons employée. Il convient donc de poser les éléments essentiels du cadre théorique dans lequel se situent les diverses opérations méthodologiques auxquelles nous avons procédé.

1.1. Les représentations en sociologie

Chez les sociologues, les diverses définitions du terme représentation sont assez vagues.

Une première méthode consiste à concevoir les représentations comme les données dont la condensation conduit à une systématisation nommée idéologie. On procède ainsi : Au premier palier, le discours des acteurs sociaux demeure l'objet immédiat de la recherche de l'analyste qui ne fait que redécouvrir son donné. Au deuxième palier l'analyste met en évidence, par condensation, par élimination des résidus, les modèles idéologiques provenant des groupes sociaux. Dans ce deuxième palier également l'analyste reste proche du discours des acteurs sociaux. Il ne fait que dénommer un objet analogue à ce discours.

Cette première méthode n'a donc fait que reproduire le discours en le classant, en l'organi-

sant et tout procédé de formalisation, un graphe (1) par exemple, apparaît comme la vérité du discours. Il paraît alors naturel de considérer l'idéologie comme l'ensemble des représentations car on passe par étapes successives du discours aux représentations et des représentations à l'idéologie. La description des données devient ainsi de plus en plus concise sans qu'il soit nécessaire de faire appel à un ensemble d'hypothèses théoriques pour justifier cette méthode de condensation. En fait, une telle méthode repose sur un postulat causaliste implicite, même s'il n'apparaît pas dans le processus de condensation, postulat qui attribue à la situation sociale des acteurs sociaux-sujets la production de l'idéologie. C'est la pratique sociale de ces acteurs-sociaux-sujets qui engendre leurs représentations et leur idéologie. La pratique sociale explique l'idéologie sans grande ambiguïté.

Une série de travaux récents, qui se réclament d'Althusser, ont mis en évidence une démarche radicalement différente. L'idéologie est, ici, définie comme représentant "le rapport imaginaire des individus à leurs conditions réelles d'existences" (2). Les individus se représentent non pas le système des rapports réels qui gouvernent leur existence mais se représentent le rapport imaginaire qu'ils entretiennent avec ces rapports réels. L'idéologique est alors un lieu de "décalage" car le discours qui y est produit est imaginaire, décalé des conditions réelles d'existence. Ce décalage ne tient pas à la conscience des sujets (en terme d'erreur ou de tromperie par exemple) mais c'est un effet de la structure sociale dont les sujets ne sont que supports. Cet effet de la struc-

(1) Graphe = ensemble d'éléments sur lesquels s'applique une relation.

(2) ALTHUSSER : Idéologie et appareils idéologiques d'état in La Pensée N° 151 Juin 1970.

ture sociale fait appel à des éléments de la théorie de la "formation sociale" et en particulier au concept de l'idéologie dominante. Ce concept récuse l'idée qu'une classe sociale, ou une strate sociale, puisse produire directement une idéologie-conception du monde. Si c'était le cas, en effet, on serait amené à des explications "accidentelles" de certains faits importants comme par exemple la domination d'un groupe qui s'accorde fort bien de l'idéologie modifiée d'un autre groupe, ou encore la possibilité permanente de contamination de l'idéologie d'un groupe par l'idéologie dominante. Cet ensemble de faits conduit à poser une certaine autonomie de l'idéologique qui s'oppose alors à la conception classique de la détermination par l'infrastructure (situation économique de classes). Mais l'analyse théorique, à laquelle nous faisons référence, indique que s'il n'y a pas détermination directe il y a, cependant, détermination, en dernière instance, par l'économique; car toute autre conception reviendrait à poser un monde des idées à côté d'un monde matériel, ce qui est contraire au monisme pratiqué par ces auteurs. Dans cette nouvelle perspective l'idéologie n'est plus le simple reflet de la situation des acteurs sociaux, elle est également le lieu d'un travail spécifique au niveau de la structure sociale. Les rapports entre les deux pôles : situation et idéologie sont complexes. Ils ne s'inscrivent plus en terme de causalité linéaire. L'idéologie n'est plus le simple produit de la situation des acteurs sociaux, mais le résultat de tout un mécanisme de production complexe au niveau de la structure sociale.

Il ne s'agit pas alors de reconstituer la synthèse ou la condensation du discours des acteurs sociaux, mais de reconstituer les mécanismes particuliers de fonctionnement idéologique de ce discours

grâce aux traces qu'ils laissent dans les articulations des divers éléments du discours, ou grâce à la manière dont les locuteurs passent d'un domaine à un autre. Notons au passage que nous distinguons nettement le processus idéologique, qui est de l'ordre de la structure sociale (cf. le début de ce rapport) et les mécanismes de fonctionnement de l'idéologie au niveau du langage (cf infra).

Il ne s'agit par exemple de dire qu'un discours est "idéaliste" en signifiant par là que son unité repose sur une idéologie hégélienne. Il nous faut repérer les éléments de fonctionnement idéologique. Par exemple on peut repérer le rôle joué, pour un locuteur, par la valorisation constante, en termes de bien ou de mal, des éléments économiques du discours. Cette valorisation peut avoir une place indispensable dans le discours du locuteur pour lui permettre de parler de l'économie nationale. De la même manière, chez un autre locuteur, cette place peut être tenue par une conception politique de l'appropriation des biens.

Cette mise en perspective de l'idéologie sert de cadre de référence à toutes les opérations méthodologiques, qui font appel, par ailleurs à des hypothèses intermédiaires spécifiques. Elle n'est qu'un cadre lointain, certes, mais très présent. Il n'y a pas, en effet, de lien direct, sauf à créer des analogies naïves, entre les mécanismes de production de l'idéologie et les processus de production du langage. Il n'y a pas davantage de lien entre le décalage structural dont nous avons parlé et les décalages que l'on peut repérer dans les discours entre les appréhensions des différents champs sociaux. Mais l'hypothèse qui se résume un peu rapidement en disant que l'idéologie produit des effets ayant une certaine autonomie dans la structure sociale est essentielle dans le choix d'une méthodo-

logie. Cette prise en compte permet de découvrir un certain nombre de voies (comme, par exemple, la recherche des domaines idéologiques grâce à l'analyse automatique du discours, de M. PECHEUX). Elle permet également d'écarter définitivement certaines méthodes comme la représentativité par échantillonnage (cf infra). Nous reviendrons plus longuement sur ces points tout au long de ce rapport.

1.2. Représentation et pratique économique.

Après avoir cerné un mode possible d'analyse des représentations, il n'est pas possible d'en rester là. Dans la mesure où, à la limite, tout discours peut permettre de parler de représentation, il est important de préciser le propos du discours en fonction de la recherche. Dans notre cas, il est important de préciser le domaine économique où se situeront les discours des sujets pour mettre en évidence un ensemble de représentations économiques adéquates aux objectifs de la recherche. Nous précisons ce domaine en sachant que tout acteur social, même s'il n'a pas un rôle explicite d'acteur économique (planificateur, entrepreneur, commerçant ...) a une connaissance pratique, vécue, de l'économie. C'est dans cette hypothèse que nous situons notre analyse des représentations économiques.

Quand nous parlons de "pratique économique" nous envisageons essentiellement une pratique sociale. Nous envisageons la pratique économique comme une modalité de la pratique sociale. L'une des caractéristiques des sociétés industrielles est de définir des zones socialement signifiées comme économiques. Pour fonder l'objet propre de notre recherche, nous avançons donc l'hypothèse de l'existence d'une catégorie pratique économique pour l'acteur social. Mais si nous appréhendons ainsi l'acteur

social comme acteur économique nous savons bien qu'il ne se perçoit pas comme tel. Il faut noter là que les états de conscience des acteurs n'ont pas d'intérêt pour nous, sinon nous serions conduits à réintroduire subrepticement un acteur social-sujet dans notre démarche. Il s'agit au contraire de postuler une dimension économique impliquée par la structure sociale. D'où l'objectif très général mais essentiel de notre analyse : isoler une dimension économique et mettre en évidence une relation entre cette dimension économique et les dimensions politiques et idéologiques au sein de la structure de la pratique sociale des acteurs. C'est en ce sens que nous essaierons de repérer dans le discours des lieux d'autonomisation, c'est-à-dire une portion du discours uniquement économique, et les modes de traduction entre l'économique et le social englobant.

2. LES PROBLEMES DE METHODES POSES PAR LA CONSTITUTION DU CORPUS.

La constitution du corpus pose simultanément deux problèmes, celui de la détermination des acteurs sociaux interrogés, et celui des modes de recueil.

2.1. Le plan de sondage

Nous devons d'abord préciser que nous éliminons d'emblée l'hypothèse de représentativité. Celle-ci n'a des sens que si la stratification du plan de sondage a quelque rapport statistique avec le phénomène analysé, comme c'est le cas pour le revenu dans l'étude des objets de consommation, ou pour la taille des exploitations agricoles dans le cas d'enquête rurale. Or jusqu'à présent personne n'a pu présenter une telle relation dans le domaine des représentations économiques. Cette relation ne

pourrait d'ailleurs que s'appuyer sur une causalité de la situation sociale dont nous avons montré les limites (cf. 1.1.).

La représentativité n'a aussi de sens que si l'on peut, à partir des résultats statistiques de l'échantillon, inférer d'autres résultats statistiques sur la population globale comme dans le cas des estimations électorales de 20 heures les soirs d'élections. Or, les principes que nous avons énoncés (cf. 1.1.) sur le repérage des articulations du discours nous font résolument quitter le domaine statistique du comptage, pour celui de la mise en évidence de structure ou de mécanismes de fonctionnement. Sur de tels modèles dont le principe est la compréhension de mécanismes, les méthodes statistiques dont le principe est l'extension, n'apporte pas grand chose. Pour cette double et fondamentale raison, (pas de structuration statistique de la population parente, volonté d'analyse structurale des données), il nous faut rejeter toute idée de représentativité. Ce n'est pas pour autant la porte ouverte à n'importe quelle détermination de la population enquêtée.

La demande d'analyse des représentations économiques est faite par des formateurs en milieu adulte en vue de l'amélioration de la formation économique. On peut alors penser que le sujet interrogé se détermine par rapport à l'économie d'abord en référence à la place qu'occupe l'économie dans la société à laquelle il appartient, ensuite en référence à la place que lui-même occupe en tant que producteur économique dans cette société. La population que nous devons interroger, doit donc être stratifiée, non pas statistiquement, comme dans le cas d'échantillon représentatif, mais par rapport à la place que la formation économique occupe dans la

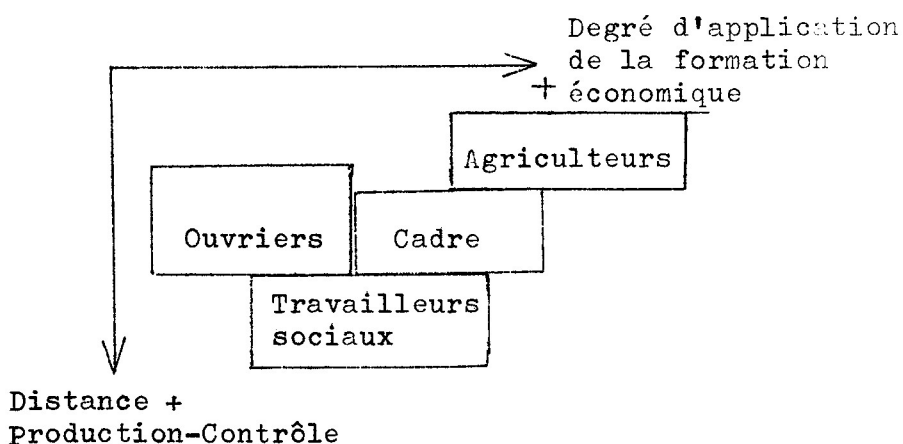
société et à la place du sujet interrogé dans le processus de production. L'interaction de ces deux dimensions est assez complexe, elle peut être analysée sous l'angle de l'enjeu que constitue le développement de la formation adulte, pour les différents groupes (cf. le rapport sur l'enquête Normateurs que nous avons effectué en 1972).

Nous pouvons traduire, opératoirement le premier axe en considérant la plus ou moins grande importance professionnelle de la formation économique. On peut par exemple, dire que les agriculteurs appliquent professionnellement ce qu'ils peuvent apprendre à propos de la formation du prix du lait ou des enjeux de l'Europe Verte. A l'opposé un ouvrier (qui n'est pas responsable syndical) n'emploiera probablement pas les connaissances sur la stagflation qu'on lui aura apprises. Le cas du cadre est intermédiaire. Pour lui la gestion peut être très opératoire mais sa formation économique peut n'être que culturelle (lui permettant de lire correctement les pages économiques des journaux).

Traduire opératoirement le deuxième axe est plus difficile; on devrait s'appuyer sur une théorie de la structure sociale, et de la place qu'y prend l'économie. Disons simplement qu'il nous a paru important de concevoir la distance qui existe pour chaque acteur social entre son acte de production vu sous l'angle de ses conséquences économiques (revenu des agriculteurs, salaires ...) et le contrôle qu'il peut exercer sur cet acte de production en tenant compte de l'environnement économique qui entoure cet acte de production et des décisions (le plus souvent macro-économiques) qui le déterminent. Par exemple, un plan de stabilisation gouvernemental n'interviendra pas beaucoup dans la détermination du volume de travail que fournira un éducateur spé-

cialisé ou un professeur. Alors qu'il aura une implication directe sur les actes du petit commerçant ou de l'agriculteur. Les cadres et les ouvriers occupant une position intermédiaire.

Nous avons choisi en fonction d'impératifs liés au mode de recueil des données de ne pas enquêter toutes les combinaisons possibles du tableau. Pour certaines catégories sociales nous avons la possibilité d'interroger deux types de population; nous avons choisi celle qui fréquente déjà régulièrement les organisations de formation adultes faisant une large place dans leur programme à la formation économique.



Remarquons qu'il ne faut pas interpréter le schéma précédent en terme d'acteur (degré d'application) et de conscience de l'aspect économique de ses actes (distance Production-Contrôle). Car cette interprétation laisse la porte ouverte à des analyses de la formation conçue comme démarche personnelle d'un sujet-acteur. Ce dernier devant devenir, par la formation un véritable acteur économique. La formation serait, alors, une trajectoire sur le plan défini par les deux axes de notre figure. Il n'en est rien. La position sur cette figure est structurelle, elle tient la place occupée par ces groupes au sein de la structure sociale.

Si l'on désire prendre en compte l'aspect dynamique de la formation il faut le faire en posant la formation économique elle-même comme enjeu de la lutte sociale. En effet la formation est liée à une volonté d'évolution de cette structure sociale où d'évolution au sein de cette structure sociale. On peut faire apparaître cela en mettant en évidence pour les populations intermédiaires (cadres, ouvriers) deux autres dimensions essentielles de la formation. La première renvoie au rôle de la formation dans l'ascension sociale, la seconde fait référence au langage économique comme langage dominant dans la structure sociale. Dans le premier cas, on peut parler de promotion individuelle, ou de recyclage, alors que dans le second cas, c'est une analyse de la société qui mène à la formation économique. C'est par exemple le patronat qui engage des actions d'information économique de son personnel pour que la rationalité économique de l'entreprise l'emporte, dans son esprit, sur les arguments idéologiques. C'est encore la nécessité pour les syndicalistes de connaître les éléments essentiels de lecture du bilan de leur entreprise ou de lecture du budget de l'Etat pour étayer leur dénonciation de l'appropriation des profits. On peut analyser ces exemples en termes de langage dominant et d'une accession nécessaire à ce langage pour l'élaboration concrète d'un projet social.

Pour concrétiser la première remarque, nous avons distingué dans le groupe des cadres et celui des ouvriers deux sous-groupes. Le premier est celui pour lequel il nous a semblé, a priori, que la formation économique jouait un rôle de promotion individuelle. Il est constitué des ouvriers non qualifiés, des cadres autodidactes. Le second nous a semblé rechercher davantage dans la formation économique une culture, un vernis, ou à l'opposé

un langage dans l'action sociale. Il est constitué par les ouvriers qualifiés, les cadres diplômés.

Pour concrétiser la deuxième remarque nous avons distingué chez les ouvriers, ceux qui suivaient une formation organisée par la direction de l'entreprise, au sein de l'entreprise, et ceux qui suivaient une formation syndicale, dans le cadre de leur responsabilité syndicale. (On a aussi distingué C.G.T. et C.F.D.T. pour tenir compte de deux types de formation économique différents). Le premier groupe est un peu un groupe contrôle, vis-à-vis des militants syndicaux.

On obtient le plan de sondage suivant, compte tenu des impératifs du recueil des données (cf. 2.2.) et des aléas de ce recueil.)

GROUPES	NOMBRES
Agriculteurs	18
Cadres autodidactes	9
Cadres diplômés	10
Ouvriers C.G.T.	10
Ouvriers C.F.D.T.	9
Ouvriers entreprise	9
Travailleurs sociaux	9
Total	74

Cette stratification ne se justifie pas simplement par rapport aux indicateurs classiques, comme les catégories socio-professionnelles, ou par rapport aux publics des formateurs, mais elle renvoie de manière prioritaire aux hypothèses que nous avons formulées sur le rôle de la formation économique pour chacun de ces groupes.

2.2. Le recueil des données

Notre analyse relative aux représentations économiques (cf. 1.2.) nous conduit à ne pas nous situer par rapport à la science économique. En effet, il peut paraître légitime qu'une science sociale particulière se découpe un champ dont elle étudie le fonctionnement. Mais ce découpage, qui répond à des exigences scientifiques et pratiques constitue toujours une opération formelle au regard de la réalité. Placer la science économique comme point de départ du questionnement, c'est la constituer en norme, les discours n'étant analysés, à travers cette grille de lecture, qu'en terme de vues-bévues vis-à-vis d'elle. Cette voie serait valable si on voulait tester des connaissances. Or en se donnant comme thème de recherche les représentations économiques on entend mettre l'accent sur les personnes en formation économiques et non pas sur la science économique. Il s'agit donc de s'interroger sur les représentations de ces personnes en formation et de considérer ces représentations comme le point de départ de l'analyse et non pas de les juger en fonction de leur conformité à une grille d'analyse économique.

Néanmoins nous écartons une autre perspective, de type psycho-sociologique axée sur le rapport des représentations avec les motivations de l'action. Une telle démarche se centrerait sur des acteurs sociaux-sujets, alors que notre cadre théorique (cf. 1.1.) est différent. Sans nier cet aspect des choses, de même qu'on ne peut pas nier que tout discours peut être analysé en terme de connaissance-méconnaissance, nous nous proposons une démarche différente qui essaiera de repérer dans le discours d'acteurs sociaux, les systèmes de représentations et les modes de raisonnement.

En disant cela, nous savons que nous nous situons simplement à un niveau de définition pratique, afin de bâtir un système d'hypothèses intermédiaires qui introduisent notre méthode de recueil des données. Prenant comme point de départ les représentations, nous sommes amenés assez naturellement à considérer un support privilégié : la verbalisation. (1). La verbalisation par le sujet de ses conduites, ou de tout autre phénomène social est déjà, en effet, une représentation. Ce qui n'est pas directement le cas d'une observation concrète de ces mêmes conduites (cf. 1.2.)

Nous ne voulons cependant pas dire que ce soit la verbalisation qui soit le véritable mécanisme producteur des représentations. Elle n'en est que le support. En outre, la verbalisation a un autre intérêt pour nous, elle permet d'accéder au raisonnement, c'est-à-dire à la mise en relation des "sens", des significations.

Or, le raisonnement ou plus exactement l'articulation des significations nous paraît être essentiel dans le repérage des liaisons du système des représentations. Par exemple, ce qui nous paraît intéressant, c'est de montrer que les acteurs sociaux, s'ils n'ont pas la connaissance scientifique ou lexicale des termes économiques, sont du moins en mesure de décrire des objets, des notions, et de raisonner en terme de calcul, de prévision. C'est donc le repérage de chaînons plus étendus que les mots qui nous intéresse dans la mise en évidence d'un

(1) Nous employons ici le terme de verbalisation à la place de celui de discours pour exprimer une procédure orale de recueil de données alors que le terme de discours a été employé jusqu'à présent dans ce texte en un sens plus large : toute production d'information par le sujet (écrits, tracts ...).

système, d'un fonctionnement. Ce souci se retrouve chez certains formateurs. André VOISIN (1) par exemple décrit une session de formation dont la première phase consiste en l'élaboration par les participants d'une liste de termes dits économiques. En fait, lorsque le pédagogue lit cette liste, il ne peut rien dire sur la signification de chacun de ses éléments. C'est seulement lorsque ces termes s'articulent entre eux par des relations spécifiques (conséquence, inclusion ...) au sein de graphiques que le pédagogue peut analyser ces termes, et constater soit qu'ils mettent en évidence une autonomie de l'économie, soit au contraire qu'ils révèlent une liaison entre l'économie ou le social, ou encore une traduction de l'économie par le social. C'est cette étape là qui constitue véritablement, à notre sens, une mise en évidence des représentations. Notre matériau est différent, mais notre propos est analogue : rechercher des configurations types (éléments et relations entre éléments) qui permettront de révéler un système de représentation.

Ayant choisi l'expression orale d'un entretien comme support des données, il nous faut situer maintenant le contenu d'un tel recueil. Là encore l'hypothèse de pratique économique intervient. Nous avons déjà dit que nous envisageons la pratique économique comme une modalité de la pratique sociale. Si l'acteur social peut être ainsi appréhendé comme acteur économique, il ne se perçoit pas, généralement comme tel. D'une part, nous pouvons penser que le sujet vit des phénomènes économiques mais qu'il ne

(1) A. VOISIN. Contribution à l'élaboration d'une méthodologie de la formation économique. Une expérience d'initiation économique des délégués d'un Comité Central d'entreprise. Rapport de l'A.T.P. Pédagogie économique. F.O.R.S. 1972.

les exprime pas forcément comme tels lorsqu'il est amené à réfléchir sur ces phénomènes; car sa réflexion peut revêtir de multiples formes. D'autre part, certains phénomènes, isolés et considérés comme économiques par l'économiste peuvent n'être jamais appréhendés comme économiques par l'acteur social; en effet bien que la modalité économique de ces phénomènes ne soit pas absente, elle est toujours intégrée à la structure, beaucoup plus vaste et complexe, que constitue la pratique sociale de l'individu (exemples : l'eau, la pollution). Isoler une dimension n'est que le résultat abstrait d'une analyse mais au départ, lorsqu'on analyse l'acteur économique, on analyse l'acteur social. L'ensemble de ses pratiques et l'ensemble de ses représentations sont liés au social. D'où l'objectif très général mais essentiel de notre recueil de données : approcher une dimension économique afin de pouvoir mettre en évidence, au niveau de l'analyse, une relation entre cette dimension et les dimensions politiques et idéologiques au sein de la structure de la pratique sociale des acteurs sachant que l'agent économique est agent économique au sein d'une pratique sociale qu'il vit comme telle et non comme une pratique économique.

Maintenant, se pose à nous le problème de la mise en oeuvre d'instruments à la fois théoriques (hypothèses) et pratiques (recueil des informations, fabrication d'un vocabulaire, etc...) pour opérer cette approche d'un champ économique au sein des pratiques des acteurs sociaux. Nous nous sommes heurtés ici à la difficulté d'élaborer un ensemble précis d'hypothèses a priori concernant la nature du découpage et de l'articulation de l'économique dans la structure sociale. Notre travail aurait été plus facile en effet si, ayant déjà un certain nombre de connaissances concernant l'analyse de la spé-

cificité de l'économique au sein du champ social, nous avons pu dès le départ déterminer un certain nombre de conduites économiques pertinentes pour l'acteur social. Pertinentes au sens où elles découleraient d'une analyse de la place de l'économique dans la structure sociale. Nous aurions pu analyser de manière plus rapide et plus précise le système de représentation lié à ces conduites. Or, l'analyse sociologique dans ce domaine n'est pas très riche. Peu de recherches se sont données pour objectif de définir l'économique dans le champ social. Ici, nous ne voulons pas dire qu'il n'y a pas de théories analysant la place de l'économique dans la structure sociale (cf. en particulier les développements de la théorie marxiste) mais que peu de recherches se sont données pour objectif de définir, de mettre en évidence, la place de l'économique dans le champ des pratiques sociales. Nous analysons en annexe divers travaux qui ont le mérite de poser une problématique, mais qui sont encore trop fragmentaires pour qu'on puisse en déduire des définitions très précises. Il nous a paru difficile de nous situer par rapport à tous ces différents types de recherche pour élaborer un ensemble d'hypothèses intermédiaires très structuré concernant la nature de notre champ d'analyse. Nous nous sommes limités à adopter cette hypothèse générale et opératoire de l'existence d'une catégorie pratique économique dans le champ de la pratique sociale des acteurs sociaux pour déterminer le type de données à recueillir et le mode de recueil de ces données.

Nous avons établi une grille d'entretien sur le thème de l'argent que nous avons choisi parce qu'il nous semble être à la fois le lieu de pratiques économiques, de conduites économiques assez caractéristiques de la société actuelle, et sous une autre dénomination, l'objet d'analyse pour l'économiste.

Ainsi en même temps qu'il peut nous permettre d'analyser la catégorie pratique économique dont les représentations font partie intégrante, par ce qui nous paraît être un certain nombre de ses éléments constitutants, il pourra ensuite nous permettre de mettre en évidence, dans une confrontation avec les formateurs, les décalages susceptibles d'exister entre la manière dont les acteurs sociaux se représentent les phénomènes liés à l'argent et la façon pour les économistes de les analyser. Le champ de la consommation au sens large du terme nous a paru particulièrement impliquant pour l'ensemble des sujets. Nous avons aussi en son temps, testé le champ de la production : l'entreprise. Mais ce champ fait plus référence chez les sujets à des représentations sociales de promotion, de métier, de camaraderie qu'à des représentations économiques. Au niveau des entretiens tests que nous avons effectués il s'est avéré, sans pour autant que l'on puisse lui donner valeur générale, que le salaire est davantage vécu comme élément du budget de consommation que comme un des éléments du bilan de l'entreprise (même chez les syndicalistes). A l'opposé, nous avons pensé à choisir un thème sur un fait économique patent : "la crise économique". Mais le discours des sujets relève dès lors plus de déterminismes sémantiques que sociologiques. En effet, le terme de crise économique induit sa signification d'une science particulière : poser cette question revenait à tester les polysémies de cette notion. Nous risquions d'aboutir aux représentations d'un mot technique plus qu'aux représentations d'un phénomène réel. Nous risquions aussi de glisser vers l'analyse des idéologies-représentations. Car, en proposant le thème de crise, nous posions au sujet un redoutable problème : qu'est-ce qui est perçu comme crise ? La réponse à ce problème renvoie le plus souvent à une analyse politique de la situation

sociale et nous arrivions à poser l'idéologie comme premier niveau de l'analyse. Il était alors difficile ensuite de la poser comme mode de production des représentations. Nous avons un exemple précis d'un tel risque. L'analyse du Concours du VI^e Plan (1) montre que la division des systèmes de représentations, qui a tendance à privilégier la consommation privée, ou au contraire les consommations collectives, pouvait provenir de la forme du questionnaire qui polarisait ces deux termes. Cette remarque mise à part, les représentations apparues ne pouvaient être que décrites, condensées; à aucun moment il n'était possible de les articuler dans un système de production idéologique et social.

Prenant en considération ces éléments, nous avons retenu en premier lieu, un thème général : celui de l'argent, puis un ensemble de thèmes proches de conduites économiques plus spécifiques : l'organisation du budget, les économies, l'achat à crédit, enfin un thème, sous la forme d'une question inductive, concernant la circulation de l'argent dans l'économie nationale et visant à permettre au sujet d'aborder, ou de se situer, à un niveau plus global de l'économie (cf. annexe : le libellé des questions). Ce que nous avons recherché essentiellement dans le choix de ces thèmes et dans leur libellé, c'est qu'ils permettent au sujet de développer un discours, déjà circonscrit, pensons-nous, dans le champ de l'économique. Nous pensons que ces thèmes rendent possible, tout à la fois, l'appréhension de champ dans un contexte social plus large qui lui donne sens. Nous faisons l'hypothèse

(1) Rapport intermédiaire sur la représentation de l'économie, à travers les résultats au Concours sur le VI^e Plan paru dans Ouest France, en Mai 1971.

que les domaines recouverts par ces thèmes peuvent être le lieu de pratiques spécifiquement économiques mais fondamentalement sociales. C'est aussi pour cette raison que nous avons essayé d'établir une certaine liaison entre le thème général et les thèmes plus particuliers; nous pensons en effet que ces thèmes serviront à tester la validité du thème de départ quant à sa pertinence économique en voyant dans quelle mesure et comment ces thèmes, qui se rapportent à des conduites particulières, se situent dans l'ensemble plus vaste du thème général, et spécifient celui-ci par un certain nombre de pratiques assez caractéristiques de la société actuelle.

Nous avons choisi, pour recueillir nos données, le mode d'entretien guidé, c'est-à-dire l'entretien mené de manière non-directive sur une suite de thèmes différents. C'est une procédure qui, d'une part est assez peu contraignante pour les sujets et permet donc d'obtenir un matériel suffisant pour parvenir à une analyse en profondeur des discours et à l'articulation de leurs différents niveaux. Elle est, d'autre part, intermédiaire entre deux extrêmes (l'entretien totalement non-directif et le questionnaire) et permet à la fois de délimiter un domaine, sans le structurer dès le départ, ce qui répond aux exigences de notre recherche.

Une telle grille et un tel mode d'entretien avaient quelques conséquences sur le plan du sondage : il nous fallait interroger des personnes qui allaient suivre un stage de formation où l'économie avait une part certaine, et il fallait veiller à les interroger avant que la formation économique ne commence, afin de ne pas recueillir des reformulations de l'enseignement reçu. Les entretiens durent environ une heure et représentent chacun de trente à cinquante pages dactylographiées.

Les thèmes ont été testés et repris avec le conseil d'un psycho-sociologue (1) qui a ensuite animé une session de formation de trois jours, au cours de laquelle les enquêteurs (2) se sont familiarisés avec la grille, et ont explicité pour eux-mêmes ce que représentait l'argent. Ces séances non directives ont été essentielles pour le bon déroulement des entretiens.

3. L'ANALYSE DU DISCOURS

Nous développons maintenant les éléments qui nous ont conduits à construire en objet d'analyse, les différents interviews constituant notre corpus. En effet, on ne code pas une verbalisation comme on coderait un schéma ou un questionnaire. Le discours a ses lois propres qui ne sont pas celles de tout message; en particulier on peut remarquer que les signes linguistiques se caractérisent par la linéarité (déroulement de la chaîne parlée). Il est alors nécessaire de prendre en compte les apports essentiels de la théorie linguistique dans ce qu'elle nous apprend sur la manière dont se constitue le sens et la phrase. Nous ne voulons pas faire une analyse linguistique (3) de notre corpus, mais nous servir à un deuxième degré de ses apports. Le choix entre divers moyens d'analyse procède d'une référence interne (une théorie

(1) Pierre MIOLLAN assistant à l'U.E.R. de Psychologie de l'Université de Provence.

(2) Melle B. CADIERGUES, Mme N. BENGUIGUI, Melle C. HENRY, Melle M.C. BARRAL.

(3) Comme on peut la trouver en particulier dans le numéro de la revue *Langages* n° 11 : Socio-linguistique et 23 : Le discours politique.

linguistique) et d'une référence externe (nos hypothèses sociologiques quant à notre démarche). Ces deux références sont étroitement liées, à tel point qu'il est difficile de les **séparer** clairement pour l'exposition.

Nous tenterons ci-dessous de mettre en évidence d'une part notre utilisation de la linguistique, d'autre part les nécessaires implications de la démarche sociologique. Nous renvoyons en annexe à l'analyse bibliographique, que nous avons faite à propos des différentes méthodes d'analyse du contenu.

3.1. Les apports linguistiques

Nous nous référons en premier lieu aux apports de Ferdinand de SAUSSURE (1) lorsqu'il définit les caractéristiques du signe linguistique : le rapport entre signifiant et signifié (2) constitutif du signe linguistique est arbitraire, c'est-à-dire qu'il ne peut y avoir bi-univocité entre signifiant et signifié. Plus précisément un terme n'a de sens que par rapport à un contexte; un terme a plusieurs sens parce qu'il peut être inclus dans des contextes différents. Et sa "valeur" est constituée par un ensemble de rapports et de différences avec son contexte et les autres termes de la langue. Ce système de rapports et de différences qui déterminent la signification d'un terme s'organise autour de deux axes, l'un syntagmatique, qui est l'axe de la chaîne parlée, l'autre paradigmatique qui définit la classe à laquelle le terme peut appartenir.

(1) Cf. Cours de Linguistique Générale, et G. MOUNIN "Ferdinand de SAUSSURE".

(2) On appelle signifiant l'image acoustique, et signifié le concept.

C'est la prise en considération de l'arbitraire du signe linguistique et de la valeur qu'il reçoit de son contexte, qui nous conduit à poser un premier principe, emprunté à la linguistique : la signification de toute unité (terme) doit être référée à des rapports d'identité et de différence avec d'autres unités (termes) dans le discours. Nous choisissons donc une méthode qui n'opère pas une classification des unités linguistiques (termes) a priori, mais qui au contraire situe les termes du langage naturel dans le contexte qui leur donne sens.

Logiquement, on pourrait penser que le contexte minimum est la phrase (unité délimitée entre deux points). Mais nous avons été obligés ici de tenir compte de la contrainte temps : vouloir analyser rigoureusement chaque phrase exigeait un très long travail. Aussi nous avons préféré passer par la définition d'un contexte, dénommé énoncé minimal, construit à partir d'une syntaxe sommaire et transformée. Pourquoi avoir opté pour une syntaxe à l'intérieur de l'énoncé minimal ? C'est là qu'intervient notre second emprunt à la linguistique : la syntaxe traduit des relations de sens entre les termes. Ici nous ne faisons pas simplement référence à l'aspect syntagmatique de la détermination du sens, mais à des théories récentes qui fondent l'analyse du sens à partir d'une centralité de la syntaxe (on la trouve exposée par D. SLAKTA (1) qui emploie la grammaire des cas de Ch. FILIMORE (2) et à l'évo-

(1) "Esquisse d'une théorie lexico-sémantique : pour une analyse d'un texte politique".
D. SLAKTA in *Langages* N° 23 1971.

(2) "Toward a moderne theory of cases" Ch. FILIMORE
in *Moderne studies in English*.
REIBEL & SCHANE 1969. New Jersey, Prenc. H.

lution de la pensée de N. CHOMSKY quand il dit que "la structure superficielle joue aussi un rôle dans la détermination de l'interprétation sémantique" (1). Il faut cependant remarquer que ces recherches s'appuient sur une syntaxe fine alors que la nôtre est assez grossière. Aussi nous ne tirerons pas tout le parti possible de telles analyses mais elles nous permettront certains traitements.

Les deux principes jouent pour nous le rôle d'hypothèses intermédiaires, hypothèses nécessaires à partir du moment où l'on a défini les données. En effet, chaque type de traitement de données ne renvoie pas directement aux hypothèses théoriques (cf. 1.1. et 1.2.), mais nécessite la mise en évidence d'hypothèses intermédiaires. La cohérence entre hypothèses théoriques et hypothèses intermédiaires à travers le mode de recueil de données est essentielle mais difficilement validable a priori, sinon par une analyse effectuée à partir d'un point de vue épistémologique. Nous pouvons dire seulement, par exemple, que nous sommes, avec ces deux emprunts, amenés à penser que l'opération de transcription à effectuer ne doit pas faire apparaître le discours comme porteur d'un sens immédiat ou caché. Il n'est pas ce matériau brut dont il suffirait de sortir la gangue (les termes du langage courant) pour atteindre le minerai pur (le sens) par une simple opération de transcription. Il nous semble que au niveau même du discours, on ne peut mettre en évidence que des chaînes de signifiants, des renvois d'un signe à un autre signe linguistique. L'analyse et l'interprétation de ces renvois, chaînes, articulations pouvant seule dégager un sens. La signi-

(1) "La forme et le sens dans le langage naturel" N. CHOMSKY in Hypothèses (Col. Change) Ed. LaFont 1972.

fication du discours étant encore autre, puisqu'elle est à rapporter avec un en-deça et un au-delà du discours; à rapporter ^{aussi} à tous les discours sociaux sur le même thème et à tous les discours, avant et après l'interview, de la personne enquêtée. Nous ne prétendons obtenir, au terme de l'analyse systématique interne, que des chaînes de signes : énoncés et dépendances. Si nos références linguistiques paraissent accorder une place prépondérante au déroulement de la chaîne parlée, il ne faut pas conclure que, pour nous, ce déroulement traduirait le raisonnement de l'esprit, il n'en est qu'un bon indicateur. Notre propos n'est pas, au niveau infralinguistique, celui du fonctionnement de l'esprit humain. Il est bien plutôt l'interrogation sur des différences d'argumentation qui seront toujours référées à des acteurs historiquement et socialement situés. Nos seuls préalables de départ au niveau de l'analyse réfèrent donc à des théories non mentalistes, car toute analyse en terme de raisonnement de l'esprit pose une question qui, en l'état des recherches nous paraît très hasardeuse. Non pas que cette question doive être rejetée, mais nous pensons qu'il est dangereux de plaquer une technique d'analyse sur un phénomène largement inconnu.

3.2. Implication d'une démarche sociologique

Nous avons tenté jusqu'ici de situer les éléments minimaux (termes) dans un contexte leur donnant sens. Mais nous voulions en même temps faire porter toute notre attention sur les enchaînements entre énoncés, pour parvenir à la mise en évidence de configurations ou de raisonnements. Et ici il ne s'agit plus tellement des mots utilisés, que des constructions dans lesquelles ils se combinent. Nous devons donc passer à un niveau plus étendu que la phrase, et force nous est de constater que le terrain

est laissé vacant par la linguistique car ce n'est pas son propos. Les seules recherches axées sur cet ordre de préoccupations sont à notre connaissance élaborées par des logiciens (par exemple les travaux sur l'argumentation du Centre de recherches sémiologiques de Neuchâtel (1)). Cet appel vers un dépassement de la phrase provient tout autant de notre objet d'étude (les raisonnements) que de notre démarche, qui nous fait appréhender le discours comme un matériel linguistique certes, mais aussi et surtout comme un produit social. Les phénomènes de sens analysés par les sciences sociales sont d'un ordre supérieur à la phrase et même, dirons nous, supérieur à l'entretien, car pour les sciences sociales, l'entretien n'est qu'un élément ponctuel dans l'ensemble des productions sociales. Mais si ce dépassement de la linguistique s'avère nécessaire, il faut remarquer que le passage à ce niveau n'est pas encore le lieu d'analyses, ni très élaborées, ni très systématiques, alors qu'il est souvent le lieu de transpositions hardies.

En effet, les sciences sociales particulières (sociologie, anthropologie) tentées par les apports et la démarche théorique de la linguistique, ne pouvaient se situer qu'à ce niveau supérieur à la phrase. Mais ne pouvant se référer là, à aucune analyse systématique, elles n'ont fait bien souvent, que "transporter" le type d'analyse adéquat au niveau des unités minimales de son ou de sens dans le domaine de la phrase, et du discours, alors qu'il existe des ruptures successives entre les niveaux du phonème et du morphème, entre le niveau de la phrase et celui du discours. (Nous renvoyons, pour une argumentation complète sur ce point, à, entre autres,

(1) J.B. GRIZE "Logique de l'argumentation et discours argumentatif", mars 1971, J.L. GALAY "Esquisses pour une théorie figurale du discours" Septembre 1972.

M. PECHEUX "Analyse automatique du discours", la toute dernière partie du 1er chapitre, et à E. BENVENISTE "Problème de linguistique générale", le chapitre sur "les niveaux de l'analyse linguistique").

Dans un tel contexte scientifique, il nous semble qu'il nous faille prendre des hypothèses minimales, et considérer que la signification d'un discours, autre que sa signification manifeste, ne se saisit qu'à travers des outils spécifiques tels que : l'étude sur des types de relations entre énoncés, des glissements de sens des termes en passant d'un énoncé à un autre, des contradictions, des associations de termes, des effets métaphoriques, des figures rhétoriques, etc... Il nous faut également poser que la théorie qui permet de reconnaître et de situer ces éléments dans un système explicatif ne peut pas être que linguistique (cf. G. MOUNIN : "Introduction à la sémiologie").

3.3. Choix nécessaires, et étape de l'analyse.

Les choix méthodologiques renvoient aux hypothèses énoncées. En particulier nous pouvons dire que le discours est fait de deux éléments irréductibles, non confondables et non séparables: le sens et l'expression linguistique. Ils sont donnés l'un avec l'autre. Il ne peut exister de distinction du type contenu/contenant (1). Aussi il est difficile de séparer une première démarche qui serait uniquement linguistique (2) puis une seconde qui se servirait

(1) On peut renvoyer ici à P. RICOEUR
"Analyses du conflit des interprétations"
Seuil 1970

(2) L'E.R.A. au CMPS du CETA de Grenoble, avec qui nous sommes en relations suivies, a pour projet de recherche la mise au point d'un programme qui permettrait de passer du langage courant à des énoncés minimaux de manière automatique. C'est un premier pas dans ce sens.

des résultats de la première analyse pour dégager un nouveau sens.

Cependant, il nous a paru important de laisser la porte ouverte à la possibilité de tester de multiples hypothèses d'analyses sociologiques sur un même matériel sans avoir pour cela besoin de refaire une opération de codage très souvent onéreuse (retour aux 3.000 pages dactylographiées). Aussi, nous avons de manière pratique, décomposé la démarche de constitution des données en deux temps :

- une première opération de transcription qui essaie de rester le plus près possible du texte, de ses termes, sans hypothèses préalables sur une organisation du sens autre que le déroulement du discours. On obtient une transcription formée d'énoncés minimaux, et de relations entre ces énoncés minimaux. L'énoncé minimal, qui ne fait référence qu'à une syntaxe élémentaire, est de la forme : syntagme nominal, syntagme verbal, syntagme nominal objet. Cet énoncé est de longueur variable en fonction des règles de réécriture de chacun de ces syntagmes (cf. infra). Nous avons aussi établi une liste d'opérateurs connectant les énoncés minimaux entre eux. Le choix de ces opérateurs est pratique et lié à plusieurs ordres de déterminations : d'une part si nous conservions tous les opérateurs syntaxiques, nous aboutissions à un trop grand nombre d'opérateurs; d'autre part, des relations de type purement logique n'épuiseraient pas la signification du discours, ou en d'autres termes les liens entre énoncés ne pouvaient pas se traduire par une simple logique formelle. Nous avons donc construit des classes de relations où viennent se ranger soit des opérateurs syntaxiques, soit, si nécessaire, un lien de sens non marqué formellement par une préposition. Ces rela-

tions sont dénommées relations majeures (parce qu'au-delà de l'énoncé minimal).

On peut voir tant au niveau de la simplification de l'énoncé qu'au niveau des relations majeures, que la linguistique n'a pas été notre seul guide au niveau des opérations pratiques. On peut dire en particulier que la syntaxe de l'énoncé minimal recouvre le type le plus simple de la phrase française, et qu'elle peut aussi traduire des relations de type logique, comme : l'agent de l'action, l'action, l'objet de l'action. De manière plus générale, ces opérations pratiques de transcription, bien que difficilement validables, renvoient à ce qui peut être utilisable pour notre analyse et à ce qui n'est que variation de surface au regard de la classe des hypothèses que l'on veut tester dans une deuxième étape. L'ensemble des opérations de transcription aboutit à la constitution d'une base de données.

- Une deuxième opération, le traitement de la base de données, est très directement liée à l'interprétation du matériel. Elle est donc la traduction opératoire d'hypothèses sociologiques. Mais cette traduction n'ignore pas les principes essentiels de l'opération de transcription, comme celui de la détermination contextuelle du sens. Il nous semble qu'il faut, ici, poser un principe essentiel. Nous pensons que l'on ne peut pas considérer les discours comme des données élaborées qu'un simple processus automatique, comme le calcul statistique (cf. en annexe, la lexicographie), ou l'analyse de tableau de co-liaison, pourrait traiter pour faire surgir des enchaînements de signifiants. Le discours est à la fois trop fluide, et polysémique, il nous semble qu'il faut le considérer seulement comme une "base de données" structurée simplement du point de vue syntaxique et sémantique par le locuteur. Cette base

pouvant seulement être interrogée par l'analyste. L'interrogation sélectionne un certain nombre de passages du discours qui ont une certaine propriété formelle ou sémantique (celle de la question). Les passages sont alors analysés pour voir en quel sens le discours a répondu à la question, et quels enchaînements cette question fait apparaître.

Nous ne donnons pas, ici, au terme "base de données", le sens courant de catalogue structuré de données statistiques ou documentaires pouvant être interrogé pour donner rapidement un chiffre, un tableau ou une liste de titres concernant un sujet. Pour un tel ensemble, le terme de "Banque d'Information" nous semble plus adéquat. C'est d'ailleurs celui employé par le département informatique de l'I.R.E.P. (1) qui est l'un des Centres de recherche le plus en pointe dans ce domaine. Pour nous, le terme "base de données" renvoie d'une part à l'idée que le discours ne devient données qu'à travers la grille des hypothèses de l'analyse, et d'autre part à l'aspect structuré de l'information (cf. infra), et à son mode d'interrogation. En disant cela, nous voulons faire remarquer qu'un même discours peut être analysé de multiples manières, nous avons souvent souligné ce point dans ce rapport. Une lecture, en terme de motivation par exemple, recherchera dans le discours des données qui soient du type : thèmes-motivés (2). Une lecture en terme de schémas culturels sous-jacents à un thème social, recherchera au contraire des données du type concepts,

(1) Institut de Recherche économique et de Planification. Université des Sciences Sociales de Grenoble, Programme RAMAS.

(2) Cf. "Analyse structurelle d'un modèle de culture normative à partir d'un Univers de thèmes motivés". M. MONTUCLARD, in *Revue Française de sociologie* XIII 1972.

ces concepts étant situés paradigmatiquement dans le réseau sémantique du thème analysé et syntagmatiquement dans le discours (1). On peut multiplier les exemples à l'infini de la psychanalyse à la sémiologie, de la sémantique à l'herméneutique, etc...

Mais l'essentiel est de prendre conscience que le même matériel linguistique est instauré comme "données" à travers la grille de lecture qui découle de chaque hypothèse, de chaque approche, de chaque visée. Il n'est pas "données" en lui-même, c'est un travail particulier qui peut en extraire des données. C'est en ce sens que nous parlons de base de données (base à partir de laquelle des données peuvent être construites)... Le terme "Base" renvoie aussi aux traitements informatiques et à l'aspect structuré de l'information : les termes du discours sont insérés dans un double réseau, celui du discours (avec ses différents éléments : syntaxe de l'énoncé minimal, relations arborescentes, ensemble de l'entretien) et celui du vocabulaire. Cet aspect sera plus longuement développé par la suite (cf. 4.2). Notre base de données n'est déjà plus tout à fait le langage brut de l'interview mais il n'est pas encore la donnée d'une interprétation.

(1) Une analyse sémantique conceptuelle. M. CANTO-KLEIN, F. LANTIER, N. RAMOGNINO in Bulletin CERP n°3 1967.

4. METHODOLOGIE ET SYSTEME DE TRANSCRIPTION

Nous avons envisagé diverses méthodes permettant de transcrire les entretiens tout en conservant leur aspect purement "signifiant" et les relations entre les termes. Ce furent les langages documentaires du type Syntol, ou les codages par constitution de chaîne du type "Terme, Relation, Terme". Nous les avons abandonnés par la suite pour les raisons que nous allons brièvement analyser. Cependant nous avons tenu à les présenter car notre système en a gardé quelques traces.

4.1. Système documentaire et codage TRT

Nous n'avons envisagé que les systèmes documentaires du type LDI (Langage pour une documentation indirecte) comme les distingue M. GOYAUD dans "Introduction à l'étude des langages documentaires" (1). Le LDI, dit-il, "doit être assez souple pour s'adapter à des questions de généralité variable et permettre au client de moduler les questions selon qu'il désire des réponses s'écartant de telle ou telle manière de la question originale". Tout autre système documentaire était, dans notre cas, trop rigide.

Nous ne décrivons que très brièvement les Eléments essentiels du Syntol, principal représentant des LDI (2). Les termes du lexique sont à la fois placés dans des classes formelles (Prédicat, Entité, Action, Etat) et reliés par des relations sémantiques (coordination, consécution, association,

(1) Documents 1 - A.T.A.L.A. Edit. C. KLINCKSTIECK
Paris 1966

(2) "L'autonomisation des recherches documentaires.
Un modèle général le syntol".
R.C. CROS J.C. GARDIN F. LEVY Edit. GAUTHIER-
VILLARS Paris 1968

prédication). L'unité de contexte du Syntol correspond à un article, rapport qui est résumé en une phrase plus ou moins longue (15-20 termes et relations en moyenne) comme par exemple la phrase : "action inhibitrice de l'acide D.Lysergique sur le Cortex" : on aura :

Acide D. Lysergique	= T1
Action inhibitrice	= T2
Cortex	= T3

Avec R1 et R2 relations reliant ces termes on aura le codage suivant : T1R1T2R2T3

Le fait que l'on veuille résumer, entraîne l'introduction d'un double système de développement : horizontal d'une part (syntagmatique) tel que T1 R1 T2 R2 T3 donne comme développement T1 R3 T3 dans certains cas; et vertical d'autre part (paradigmatique) tel que si T3 = "Cortex" on sait aussi que T1 fait partie du "Cerveau". Ce double système de référence paradigmatique et syntagmatique est propre à un domaine documentaire donné. Mais, dans notre cas, il nous est impossible de le postuler car, c'est l'objet même de la recherche. Peut-on dire par exemple que "les Allocations familiales" font partie du "Revenu" quand on sait par ailleurs que ce fait n'est pas souvent perçu par l'interviewé. Une telle relation d'appartenance est au contraire essentielle dans un système documentaire en sciences économiques.

Ne pouvant adopter un tel système de classe d'équivalence sur les termes et sur les relations successives, car ce système devrait être défini a priori, sur les signifiés, il nous fallait envisager de conserver l'idée de termes ou relations mais sans organisations préalables des uns et des autres. Nous arrivions alors à des suites de Termes (T) et de relations (R) qui devaient être écrites avec

ponctuation et parenthésage pour que l'on n'en perde pas le sens. Exemple : "L'argent est une bonne chose et une mauvaise chose" s'écrirait :

T1 R1 ((T2 R2 T3) R3 (T4 R2 T3))

avec T1 = argent, T2 = bonne, T3 = chose, T4 = mauvaise

R1 = être

R2 = relation de spécification

R3 = relation de coordination.

Toute autre forme donnerait lieu à des interprétations équivoques. On peut alors penser à bâtir des matrices de coapparition de termes liés par une relation donnée. On aurait dans l'exemple cité le tableau 1

Pour la relation R2	T1	T2	T3	T4
T1	0	0	0	0
T2	0	0	1	0
T3	0	1	0	1
T4	0	0	1	0

(Tableau 1)

Ces matrices donneraient d'une certaine manière la fréquence d'association des termes pour une relation donnée qui les lie et permettraient ainsi d'appréhender le contexte de ces termes. Mais tout de suite apparaît une difficulté : que relie dans notre exemple la relation R3 (ou R1) ? Doit-on considérer qu'elle relie des Termes composites (T2 R2 T3) et (T4 R2 T3) devenant chacun un nouveau terme Tn et Tn + 1. Mais, dans ce cas, le vocabulaire deviendrait rapidement inutilisable car tendant vers l'infini. Doit-on au contraire, considérer que le fait que la relation R2 et le Terme T3 sont apparus dans les deux ensembles entre parenthèses, entraîne que la coordination R3 se rapporte aux termes T2 et T4 seuls différents dans les deux énoncés. D'une certaine

manière on se donnerait des règles de génération de triplets T R T (ici T2 R3 T4) qui tiendraient compte des termes composites mis en relation. Mais la complexité de ces règles ne permet pas leur généralisation. Leur existence même tombe sous la critique que nous avons faite aux mécanismes de développement dans les langages documentaires. Dans notre exemple, peut-on mettre en relation "Bonne et Mauvaise" sans savoir que c'est en fait de l'argent que l'on parle ? Directement appliqué au texte, les tableaux de fréquence de coliaison par une relation donnée semble être d'un intérêt réduit, pour notre recherche.

4.2. La double situation d'un terme dans la chaîne syntaxique et dans la suite des enchaînements.

Les premiers essais de codage nous ont fait prendre conscience d'une double situation (repérable par deux systèmes relationnels) des termes. Ceux-ci sont situés syntaxiquement dans un univers limité mais qui leur donne un sens. Pour ce qui nous concerne cet univers peut être réduit à sa forme canonique. "SNS + SV + SNC" avec SNS syntagme nominal sujet, SV syntagme verbal, SNC syntagme nominal complément : Exemple " Jean parle à Jacques". Nous avons ainsi été amenés à considérer des énoncés minimaux (dit E.M.) dont la principale caractéristique est pratique, c'est celle de n'avoir qu'un verbe à la forme personnelle, et dont les caractéristiques structurelles font référence à une syntaxe élémentaire, noté par la dépendance des termes les uns vis-à-vis des autres.

La deuxième situation est plus sémantique. L'énoncé minimal pris dans l'ensemble de sa signification (et non de l'un ou l'autre de ses termes) est

situé dans un contexte plus large celui de l'ensemble des autres énoncés minimaux qui lui sont reliés par des relations de type logico-sémantique ou discursif. Par exemple "On voit la maison parce qu'elle est rouge" est réductible à deux énoncés "On voit la maison", la maison étant ici rapportée à un acte de vision, et "la maison est rouge", la maison étant ici caractérisée par la couleur rouge. Chaque énoncé seul n'est pas totalement satisfaisant puisque le sens du premier énoncé est totalement lié au sens du second énoncé. Nous avons donc introduit un système de relations dites "majeures" permettant la prise en compte d'un contexte plus large. Ici nous obtenons "(On voit la maison) parce que (la maison est rouge)".

4.2.1. Formalisation des énoncés minimaux

Nous avons dit que la forme canonique de l'Enoncé Minimal était :

"SNS + SNV + SNC"

Pour positionner chacun de ces termes par la marque de leur dépendance mutuelle et permettre les ellipses, nous allons remplacer les "+" par des relations dites "mineures". On a besoin ici de deux relations qui expriment une association orientée (AO) :

A01 de sujet à verbe

A03 de verbe à complément.

On obtient alors :

SNS. A01. SV. A03. SNC

Avant d'aller plus loin, il nous est nécessaire de noter que tout aspect du langage naturel est abandonné, en l'espèce la morphologie. Nous ne nous attachons qu'à la fonction. Ainsi c'est la fonction verbale d'un terme qui est pris en considé-

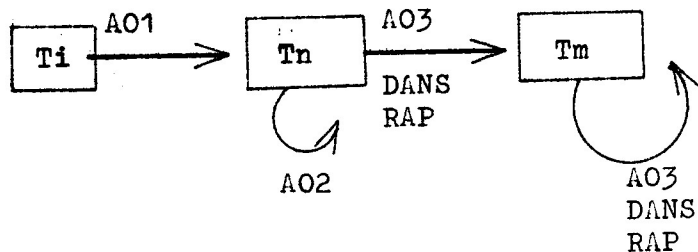
ration dans son positionnement (après A01) et non sa forme qui peut être nominale ou verbale. On peut prendre une hypothèse exactement inverse, mais c'est notre recherche d'une logique au plan de la réflexion économique qui nous fait privilégier l'aspect syntaxique sur l'aspect morphologique. Notre codage note aussi tous les SV à l'infinitif et tous les SN au masculin singulier. Comme nous marquons le SN sujet, les marques morphologiques verbales seraient redondantes. Les formes du conditionnel et du futur qui, elles ne sont pas redondantes sont exprimées par des relations ou des marques adverbiales.

Le type de discours que nous avons, l'entretien non-directif, a la particularité de postuler le plus souvent un sujet identique : le locuteur. Pour éviter des lourdeurs de codage, celui-ci est ignoré sauf s'il y a une marque d'insistance. L' E.M. se réduit alors à SV. A03. SNC. Dans des cas limites, il peut se réduire au verbe seul comme par exemple dans "je pense" du célèbre "(je pense) donc (je suis)".

Enfin nous avons voulu tenir compte de deux particularités du langage qui complexifient un peu la forme canonique. D'une part la forme verbale peut être constituée par deux verbes, l'un au mode personnel et l'autre à l'infinitif, comme dans "je pense acheter des crayons". Le deuxième verbe est alors précédé de la relation A02 : "Penser. A02. Acheter. A03. Crayons." D'autre part le complément d'objet peut être soit direct, soit indirect, il n'est pas bien nécessaire de les distinguer mais il nous est apparu important de mettre en évidence deux compléments circonstanciels, celui de lieu ou de temps noté DANS (exemple : "Acheter. DANS. futur") et celui qui désigne le "Par rapport à" le "en fonction de", le "Pour attribution", que l'on note RAP (exemple : "Acheter. A03. Télévision. RAP. Enfants").

Notre énoncé minimal peut être présenté sous forme d'un automate à états finis (cf. figure 1) qui répond aux règles de réécriture suivantes (les Termes T pouvant prendre la forme SN ou SV)

$T \longrightarrow SN$
 $T \longrightarrow \emptyset SV$
 $E.M. \longrightarrow \emptyset + SV$
 $\emptyset \longrightarrow SN + A01$
 $SV \longrightarrow SV + A02 + SV$
 $SV \longrightarrow SV + A03 + SN$
 $A03 \longrightarrow DANS$
 $A03 \longrightarrow RAP$
 $A03 \longrightarrow SN + A03 + SN$



Avec Ti : terme initial
 Tn : terme initial ou intermédiaire ou final
 Tm : terme intermédiaire ou final

Et avec Ti : forme sujet
 Tn : forme verbale
 Tm : forme complément

La syntaxe de tout énoncé minimal doit correspondre à celle de cet Automate de Vérification.

Les opérateurs syntaxiques sont alors caractérisés ainsi :

A01 (ou A1) suit le sujet d'un verbe
 A02 (ou A2) introduit le 2e verbe
 A03 (ou A3) introduit le complément

AO3 peut être réécrit DANS
 RAP
 AO3.Tm.AO3

On a aussi mis en évidence des opérateurs de termes. Ceux-ci mettent en relation deux termes. Cette relation a pour effet de donner le même positionnement syntaxique au deux termes. On peut en fait considérer ces opérateurs comme permettant la réécriture d'un terme.

T peut être réécrit T. + .T
 T. OR.T
 T. SP.T
 T.SPE.T

- "+" : L'opérateur "+" indique la coordination comme dans "l'argent est la meilleure et la pire des choses", OR indique la disjonction comme dans "l'argent c'est bon ou mauvais selon les cas".
- SP indique la spécification d'un terme : c'est l'adjectif, le complément de nom, la marque adverbiale du verbe. Sa notation se fait du spécifié au spécifiant. On a par exemple :
 "Avoir.AO3. Argent.SP.Beaucoup"
 "Etre.AO2.Payer.SP.Progressivement".
- SPE indique un petit exemple spécifiant d'un terme comme dans :
 "Pauvre.SPE.Travailleur-étranger.AO1.Vivre.Dans.Taudis".

Nous avons enfin fabriqué deux relations conventionnelles afin de coder plus rapidement : le "=" qui se retraduit dès la mise sur ordinateur par "AO1.ETRE.AO3" et le "N =" qui se retraduit de même "AO1.ETRE.SP.NON.AO3".

Notons au passage que la négation d'un ter-

me noté au codage "NON-acheter" est traduit dès sa mise sur ordinateur en "Acheter SP.NON". Ceci afin de ne pas augmenter inutilement le vocabulaire.

4.2.2. Le système de relations Inter-énoncé minimaux.

La dépendance entre Enoncés-Minimaux ou plus exactement leur enchaînement se fait à l'aide de relations dites "Majeures". Celles-ci sont, avons-nous dit de type sémantique.

Seule exception une relation d'Expansion de l'E.M. qui traduit souvent une relative mais qui est, dans le cas général, l'obligation de créer deux E.M. alors que le sens par la syntaxe est une unité. Elle est notée EXP, par exemple on a :

* (Industrie .AO1.Avoir.AO2.Capital).EXP.
(Capital.AO1.Tourner.SP.Plusieurs fois)
pour traduire "L'industrie a un capital qui tourne plusieurs fois".

* autre exemple (Prendre.AO2.Crédit).EXP.
(Amène.AO2.remboursement.SP.Contrain-
gnant)

Ici le sujet de Amène est le 1er énoncé, il ne peut pas être noté comme tel, l'Expansion nous le retransmet.

Nous pouvons ranger les relations Majeures en quatre grands types :

- Combinatoire entre E.M. : OU, ET
 - Dépendance entre E.M. : FONction de
 - Comparaison entre E.M. : COMparaison
 - Enchaînement discursif : C'est A Dire
- une dernière relation ayant un rôle un peu à part :
la Restriction : le MAIS.

Le groupe des relations exprimant un peu la logique booléenne du OU et du ET ne pose pas trop de problèmes, sauf à dire que le OU est un "ou bien" et non pas une simple marque de coordination comme dans la phrase "Il lui donnera 100 Ecus ou il l'épousera" des Noces de Figaro. Il faut enfin dire que le ET joue quelques fois le rôle d'une liaison entre énoncés successifs mais non reliés formellement dans le texte.

Le groupe des relations de dépendance a fait lui aussi à ses débuts référence à la logique, en particulier avec l'IMPLication mais en fait les implications du discours n'ont rien de très logique, et il faut beaucoup plus voir ces relations comme exprimant des liaisons orientées entre E.M. Nous avons retenu les relations suivantes :

- FONC : en fonction de, par rapport à, étant donné.
- NFONC : négation de FONC
- IMPL, PCQ : Implique et parce que. Ces deux relations sont en fait remplaçables l'une par l'autre si l'on retourne l'ordre des propositions $A.PCQ.B = B.IMPL.A$. mais on a conservé les deux pour avoir le sens du discours.
- NIMPL : négation de IMPL
- SI : relie généralement deux E.M. dont le premier est au conditionnel ou manifeste l'idée de condition et dont le deuxième est impliqué par le premier :
 $A.SI.B = \text{Si } A \text{ alors } B$.
- NSI : négation de SI
- POUR : marque la finalité. Ex. : "je demande à avoir de l'argent pour vivre normalement".
- NPOUR : négation de POUR.

Le groupe des comparaisons introduit une dimension particulière du discours. Les différents énoncés minimaux sont ici comparés et plus ou moins valorisés les uns par rapport aux autres. Quelques fois, ces relations expriment un ordre dans le discours (Premièrement, deuxièmement ...) mais notre codage y est mal adapté, car nous n'avons pas, ici, affaire à des conférences ex cathédra où l'ordre a un sens propre.

Nous avons retenu les relations suivantes :

- COMP : de la même manière que, comme
- NCOMP : négation de COMP
- SUP : mieux que
- INF : moins que
- EQ : relation d'équivalence entre les énoncés mais pas au sens de l'équivalence mathématique. C'est le plus souvent la transcription d'un verbe être ou de l'explicitation d'un énoncé par un autre.
- NEQ : négation de EQ

Enfin le discours a ses lois propres et quelques fois les explicitations sont plus complexes qu'une simple Equivalence entre E.M. On a affaire, alors, à de longs développements. Ils sont introduits par C'est A Dire : CAD, ou si le développement est un exemple, on le note SPEX.

Le MAIS marque une simple restriction. Il n'est pas catalogable facilement, peut-être parce que notre esprit s'y refuse. Mais il est essentiel et c'est souvent dans le discours un point de retournement qui a son importance et qui doit être marqué.

4.3. Les enchaînements de sens.

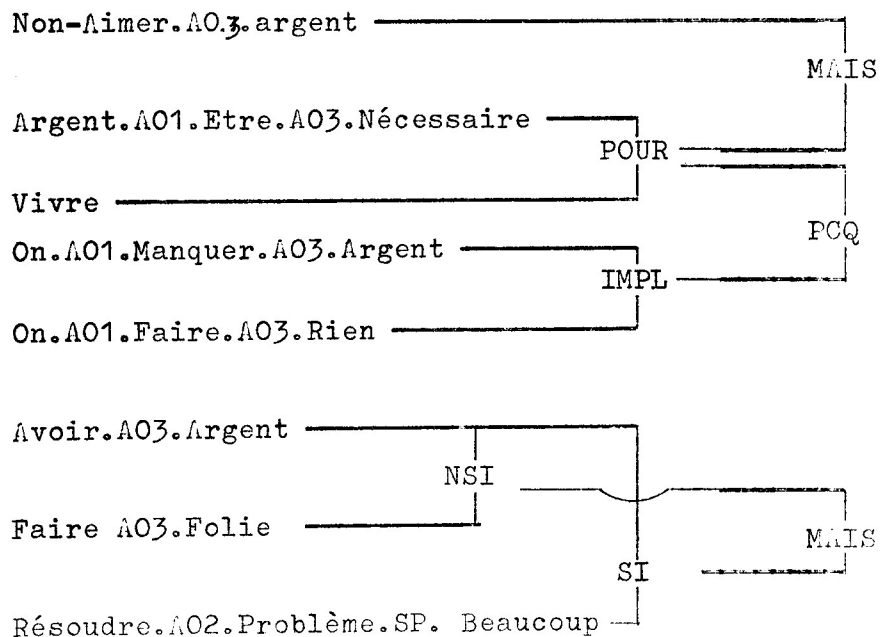
Il ne suffit pas de poser d'une part les Enoncés Minimaux et d'autre part un ensemble de relations, il nous faut voir comment ces éléments peuvent se combiner pour traduire la fluidité du discours.

Cette fluidité est déjà apparente dans le choix même des Relations qui n'ont pas la rigidité d'une logique. Même si quelques fois, elles y puisent leur inspiration, de détournement en glissement de sens, elles prennent une signification propre sans aucun correspondant logique.

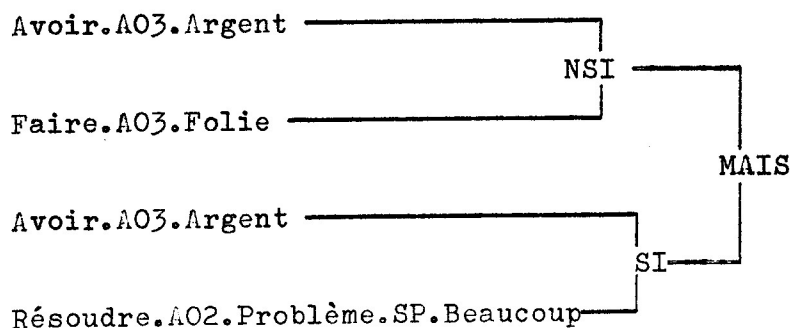
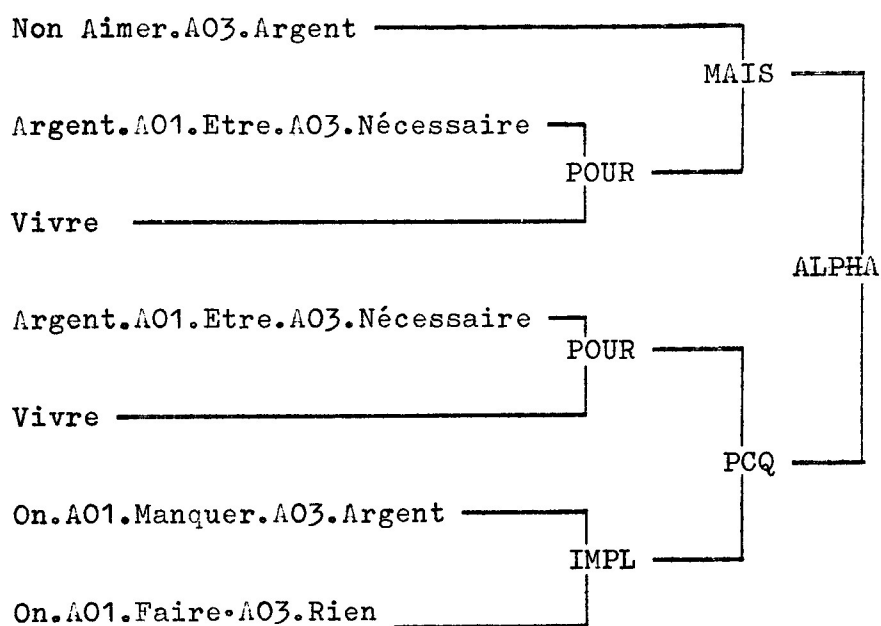
D'autre part, les emboîtements entre systèmes relationnels sont possibles.

1° - Par la règle de réécriture de tout Enoncé Minimal. EM est réécrit EM.R.EM. On a ainsi un système en arbre sans fin pouvant être écrit avec des parenthèses : (EM1.R1.EM2).R2.EM3.

2° - Par une possible traduction du discours non seulement en arbre comme ci-dessus, mais en réseau comme dans les exemples suivants :



Au niveau de codage, il suffit de mettre en relation successivement ces différents énoncés. Ensuite l'ordinateur les organisera de façon telle qu'il puisse retrouver l'ensemble de ces réseaux et les visualiser correctement. Ceci est fait en créant une relation fictive dite ALPHA et en dupliquant les Enoncés reliés à plusieurs autres. On obtient ainsi les configurations suivantes, qui ont la forme d'arbres et peuvent être traités par la suite mais qui gardent par la duplication de certains E.M. la forme du réseau que l'on peut retrouver à la main sur les éditions.



Se pose alors le problème de la clôture des enchaînements, ou inversement de la coupure entre grandes unités de sens. Ces coupures sont difficilement repérables par des marques formelles, si ce n'est l'intervention de l'enquêtrice ou un très long silence. Nous raisonnerons, ici, en disant que nous avons affaire à deux systèmes relationnels distincts si l'on ne peut pas établir une relation nouvelle entre énoncés minimaux de l'un et de l'autre des systèmes. Ce n'est pas très satisfaisant pour l'esprit, mais cette règle a pour elle d'être interne à notre système de codage et de réduire, par là même, l'interprétation de la coupure, sans pour autant l'éliminer.

Les systèmes relationnels dégagés sont repérés et mis en mémoire sous forme d'arbres comme nous les avons dessinés dans les exemples précédents. Cette forme n'étant pas perçue complètement au moment du codage, elle ne sert donc pas de guide a priori à ce codage. Ceci est important car le codeur doit seulement porter son attention sur les relations, entre Énoncés minimaux, qu'il trouve dans le texte et non sur un enchaînement arborescent logique qui ne serait que la logique du codeur à travers la contrainte de fabrication des arbres. La visualisation des entretiens se fait par l'écriture d'Arbres. Une telle visualisation permet dans un premier temps la vérification rapide du codage si on lit parallèlement le texte et les arbres correspondants. Elle doit être interprétée comme une base de données, ce qui implique en particulier que la hiérarchie des relations dans l'arbre ne soit pas conçue directement comme étant une donnée en elle-même. On ne peut pas, par exemple, tirer de cette hiérarchie, une autre hiérarchie : celle sur les relations.

4.4. Mise en oeuvre des Traitements Informatiques de création de la base de données.

La mise sur ordinateur d'un texte sous la forme d'une base de données arborescentes suppose un codage et un programme de création et d'enrichissement de la base de données. Nos options nous ont conduits à créer un codage très simple permettant de multiples mises en relations d'Enoncés Minimaux, le programme d'entrée étant alors conçu à la fois comme décodant les termes et les relations et comme transformant le réseau de relation en structure arborescente.

4.4.1. Le codage.

Le codage essaie de n'employer qu'un nombre limité de conventions. Elles concernent uniquement l'agencement syntaxique et relationnel.

Les termes sont ceux du langage courant avec une volonté de réduire les polysémies, mais sans que cette réduction ne fasse appel à autre chose qu'à la pratique quotidienne de la langue et sans qu'elle soit l'obsession du codeur. Les termes sont écrits sans marque morphologique, dans la forme où ils se trouvent dans un dictionnaire, les verbes à l'infinitif et les noms au singulier. On constitue à partir des premiers entretiens et des hypothèses de notre recherche un vocabulaire dit de base qui sera commun à tous les sujets enquêtés. Au moment de l'enregistrement de chaque entretien on relève les termes qui ne font pas partie du vocabulaire de base et on crée un vocabulaire propre à cet entretien. Les vocabulaires propres sont étudiés pour voir si certains termes peuvent être remplacés, dans le cas contraire ils sont conservés dans le vocabulaire propre.

Les Enoncés Minimaux sont considérés comme des ensembles alternés de Termes et de relations mineures. Le premier et le dernier mot devant être un Terme.

La carte perforée est un élément de base du codage, elle est composée d'un numéro et d'un texte : le texte du codage. Ce numéro peut apparaître dans un texte ultérieur pour servir à la mise en relation du texte de cette carte, pour marquer un enchaînement.

L'enchaînement entre énoncés minimaux peut être fait par une simple mise entre parenthèses si la partie de texte considérée est arborescente.

Exemple :

(Acheter.A03.Télévision).PCQ.(Enfant.A01.Réclamer.A03.Télévision).

On peut aussi repérer un Enoncé Minimal ou toutes combinaisons d'E.M. par le numéro de la carte où il se trouve et mettre en relation ces numéros. Soit pour le même exemple :

N° de carte	Texte
01001	Acheter.A03.Télévision
01002	Enfant.A01.Réclamer.A03.Télévision
01003	01001.PCQ.01002

Ce système permet l'élaboration du réseau puisque l'on peut avoir une mise en relation du même numéro de carte avec plusieurs autres cartes ou Enoncés Minimaux. Comme par exemple :

P1004 01001.EQ.(Avoir.A03.Dettes).

-- Les règles d'écriture sont alors :

avec NC = Numéro de carte perforée.

EM = Enoncé minimal (texte)

R = Relation majeure.

* NC se réécrit (EM)

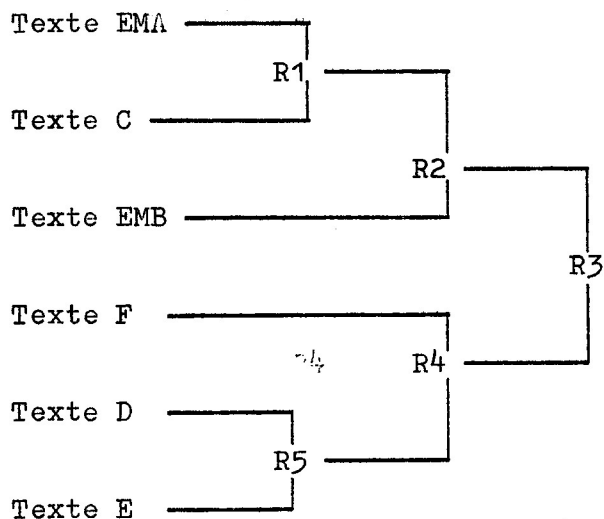
avec la convention que le début et la fin de carte peuvent tenir lieu de parenthèse ouverte et fermée respectivement.

* EM se réécrit (NC.R.NC)

On peut obtenir par exemple le discours suivant :

01001	Texte EM.A
01002	Texte EM.B
01003	((01001.R1.(Texte C)).R3.01004
01004	(01005.R4.((Texte D).R5 (Texte E)))
01005	Texte F

Le forme arborescente de ce discours étant figurée ci-dessous :



(Tableau 1)

- les règles de vérification suivantes peuvent être appliquées :

- il y a autant de parenthèses ouvertes (NbPO) que de parenthèses fermées (NbPF).

$$\text{NbPO} = \text{NbPF}$$

- il y a autant de parenthèses ouvertes que de relations majeures (NbR) et d'énoncés minimaux (NbEM)

$$\text{NbPO} = \text{NbR} + \text{NbEM}$$

Si l'on omet les parenthèses de début et de fin de carte, on a :

$$\text{NbPO} = \text{NbPF} = \text{NbR} + \text{NbEM} - 1$$

On peut voir en annexe des listings de codage.

4.4.2. Le programme d'entrée.

La structure des données (sous formes de cartes) à l'entrée de ce programme vient d'être décrite. La structure de la "Base de Données" en sortie à la forme d'une liste dont les Eléments terminaux sont les mots du vocabulaire soit propre soit de Base.

Pour passer de la forme Entrée à la forme Sortie, il faut reconnaître les Termes et relations, et Enoncés Minimaux, puis traiter leur structuration. Enfin générer la structure de Liste et de Pile.

4.4.2.1. La procédure de reconnaissance.

La première opération est celle de la délimitation.

Les délimitateurs sont :

* Pour les Termes, relations et Numéro de carte dans le texte :

- le Blanc
- le Point
- les Parenthèses ouvertes ou fermées
- le début et fin de carte (sauf convention particulière si plusieurs cartes sont nécessaires à l'écriture d'un même ensemble).

* Pour les Enoncés Minimaux :

- les parenthèses ouvertes et fermées
- le début et fin de carte.

La deuxième opération est celle de la distinction des termes, relations mineures et majeures.

- * Les relations majeures sont après un Numéro de carte ou un Enoncé Minimal.
- * Les relations mineures sont après un Terme.
- * Les Enoncés Minimaux comme les Numéros de carte suivent une parenthèse ouverte ou un début de carte, ou une relation majeure.
- * Les Termes sont soit le premier mot d'un Enoncé Minimal, soit après une relation mineure.
- * Les Numéros de carte dans le Texte sont une suite de 5 chiffres.

4.4.2.2. Constitution des différents niveaux de la structure de liste.

4.4.2.2.1. La FILE des E.M.

La première cellule de cette File indique la longueur de la File. Puis chaque énoncé est constitué d'une cellule qui indique sa longueur, d'une seconde cellule qui indique son numéro et d'autant de cellules que de termes et relations, celles-ci renvoyant à l'adresse du mot dans les vocabulaires.

Les Enoncés Minimaux sont mis l'un derrière l'autre sans blanc il forme une file ininterrompue repérant le niveau inférieur du vocabulaire et repéré au niveau supérieur par la première cellule de chaque Enoncé Minimal (cf. figure 2 p. 59)

4.4.2.2.2. La LISTE des relations et Enoncés Minimaux.

L'élément de base de cette Liste est formé de quatre cellules :

- la référence aux niveaux inférieurs (Adresse de la File des Enoncés Minimaux ou du Vocabulaire des relations majeures),

- la référence de l'élément dominant dans la liste,

- la référence de l'élément dominé en haut

- la référence de l'élément dominé en bas.

Par exemple, l'arbre ci-dessous sera repéré par le tableau 2.

	N° de l'élément	Réf. au vocabul.	Réf. à l'élément dominant	Réf. à l'élément dominé H	Réf. à l'élément dominé B
EMA					
EMB					
EMC					
	1	EMA	2	0	0
	2	R1	4	1	3
	3	EMB	2	0	0
	4	R2	0	2	5
	5	EMC	4	0	0

(Tableau 2)

La constitution de cette Liste, se fait en partie, au moment de la lecture des cartes. A ce moment là chaque E.M. génère un élément de la Liste constitué uniquement par l'adresse de sa première cellule dans la FILE des E.M.

Chaque relation majeure génère un élément de la LISTE constituée par l'adresse de la relation dans le vocabulaire des relations majeures.

Parallèlement pour chaque carte on génère une Pile constituée :

- des parenthèses qui n'encadrent pas le texte d'un Enoncé Minimal.

- des adresses dans la LISTE, des E.M.

- des adresses dans la LISTE, des relations majeures.

- des numéros de carte.

Si l'exemple ci-dessus avait été écrit ainsi,

N° Carte	Texte
01001	(EMA).R1.(EMB)
01002	01001.R2.(EMC)

on obtiendrait avant le traitement de la Pile de la carte 01001, la PILE et la LISTE du tableau 3.

PILE			LISTE				
			Référence à				
1	1	EM	1	EMA	0	0	0
2	2	R	2	R1	0	0	0
3	3	EM	3	EMB	0	0	0

↑

Caractéristique
de l'élément de la
Pile.

↑

↑

Référence à l'é-
lément de Liste.

(Tableau 3)

Après le traitement de la Pile on obtient le tableau 4.

LISTE				
1	EMA	2	0	0
2	R1	0	1	3
3	EMB	2	0	0

(Tableau 4)

Avant le traitement de la Pile de la deuxième carte 01002, on a la PILE et la LISTE du tableau 5

PILE			LISTE				
1	01001	NC	1	EMA	2	0	0
2	4	R	2	R1	0	1	3
3	5	EM	3	EMB	2	0	0
			4	R2	0	0	0
			5	EMC	0	0	0

(Tableau 5)

Après le traitement de la Pile, on obtient le tableau 6.

LISTE				
1	EMA	2	0	0
2	R1	0	1	3
3	EMB	2	0	0
4	R2	0	01001	5
5	EMC	4	0	0

(Tableau 6)

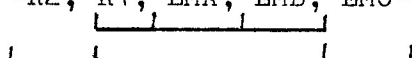
A la fin de la lecture des cartes, on rend homogène le tableau LISTE, en remplaçant les numéros de carte par l'adresse de l'élément dominant de cette carte qui a été repéré au moment du traitement de la Pile liée à cette carte. On obtient alors le tableau 2.

4.4.2.2.3 L'ARBRE en écriture Polonaise Préfixée.

Le balayage du tableau LISTE permet de repérer les éléments non dominés, ce sont les Racines des Arbres que nous allons conserver.

A partir de chacune de ces racines, on peut chercher les éléments dominés Haut et Bas, puis les éléments dominés successifs jusqu'à trouver les E.M. qui sont les feuilles terminales de l'Arbre.

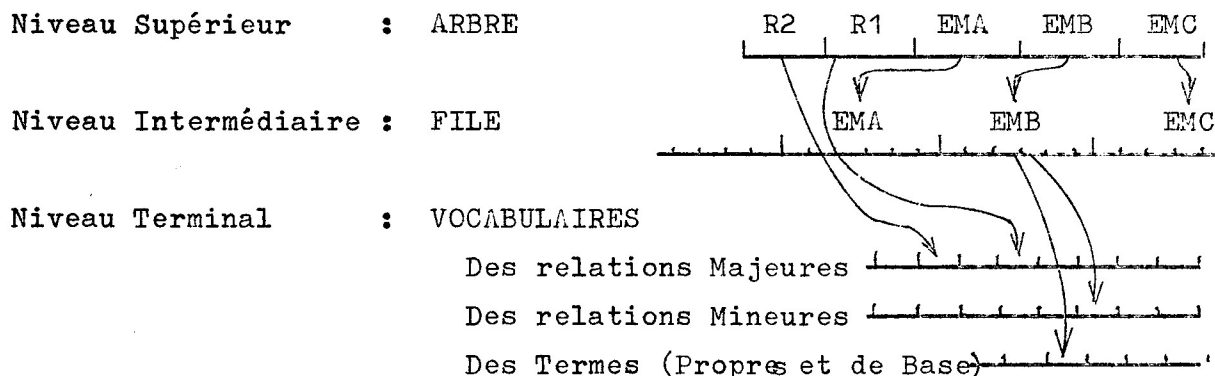
Un traitement de PILE permet cela. L'exemple que nous avons traité s'écrit alors :

"R2, R1, EMA, EMB, EMC"


Chaque relation est suivie des deux éléments qu'elle relie, ceux-ci pouvant être réécrits. Chaque élément de cet ARBRE fait référence soit au vocabulaire des Relations soit à la Première cellule de l'E.M. dans la FILE. La LISTE n'a servi que d'intermédiaire pour permettre la constitution des Arbres.

4.4.2.3. Schématisation de la Structure des Données.

On peut schématiser par la figure 2 (exemple traité) notre Base de Données les flèches indiquant que la cellule n'est qu'une adresse d'un niveau inférieur.



- figure 2 -

On peut considérer que les vocabulaires n'en forment en fait qu'un. Ceci en positionnant de 1 à 50 les relations majeures, de 51 à 100 les relations mineures, et de 101 à 999 les Termes.

L'intérêt d'une telle structure est dû à son faible encombrement et à sa structuration qui permet d'effectuer, à partir du seul traitement par Pile du niveau supérieur (ARBRE), toutes les opérations de reconnaissance et d'écriture. La référence aux niveaux inférieurs n'est qu'un simple jeu d'adressage indirect. En outre, on peut apporter des modifications à chacun des niveaux de manière indépendante des autres niveaux.

Le fichier de sortie est constitué de trois ensembles pour chaque entretien :

- le vocabulaire propre sous forme d'un fichier séquentiel.

- l'ensemble des Enoncés Minimaux sous la forme de la FILE des E.M. (renvoyant aux vocabulaires).

- L'ensemble du système relationnel arborescent sous la forme de la suite des ARBRES écrit en notation Polonaise Préfixée : les relations majeures jouant le rôle d'opérateur et les Enoncés Minimaux, repérés par leur premier terme dans la File, jouant le rôle des Termes de l'Opération.

Le programme d'entrée a été écrit en FORTRAN et fait environ 1250 cartes. Il permet de traiter 3000 cartes à la minute sur l'Ordinateur 370/165 du CIRCE à ORSAY.

4.4.3. Le Programme de mise à jour

La phase de codage entraîne deux types d'erreur difficilement évitables. D'une part des erreurs portant sur la syntaxe des Enoncés Minimaux, ou sur le système de relation entre Enoncés, d'autre part des erreurs de codage proprement dit, celles-ci introduisant des distorsions entre le texte et la transcription.

Le programme d'entrée corrige lui-même les erreurs les plus simples (tel que oubli d'une parenthèse) et indique les erreurs du premier type. On peut en voir un exemple en annexe.

Pour corriger les erreurs du deuxième type, il a été effectué systématiquement une "lecture à deux", l'un ayant en face le texte dactylographié, et l'autre les arbres sortis lors d'un premier passage sur ordinateur.

Dans l'un et l'autre cas, il est établi des bordereaux de perforation permettant de modifier le fichier bande tel qu'il avait été constitué lors du premier passage. Ces cartes modifications peuvent soit annuler ou remplacer un enregistrement du fi-

chier bande, soit modifier l'une ou l'autre des colonnes d'un enregistrement, soit enfin créer de nouveaux enregistrements.

Le fichier obtenu en sortie devient le nouveau fichier entrée du programme de création de la base de données.

5 - Méthodes et Programmes d'interrogation de la base de données liés aux premières hypothèses

5.I Traitement du vocabulaire

L'analyse d'un terme consiste en l'étude de son contexte : contexte minimal d'une part, au sein des énoncés minimaux, contexte plus large d'autre part, au sein des arbres où il apparaît. La première analyse est distributionnelle, la dernière conduit à une étude en champs sémantiques.

5.I.I. Distribution d'un terme

Passant en revue l'ensemble des Énoncés Minimaux d'un entretien, ou d'un sous-ensemble quelconque du corpus, on peut repérer pour chaque terme à la fois les Énoncés où il se trouve et la place qu'il occupe dans ces Énoncés. Il suffit ensuite, de trier par numéro de terme croissant les informations concernant chaque terme. Ce tri effectué, on peut éditer (sur rupture) pour chaque terme les Énoncés où il se trouve en marquant ce terme par une astérisque. On peut en même temps compter les occurrences de ce terme aux différentes places syntaxiques qu'il occupe dans ces Énoncés. On obtient alors la fréquence des places de ce terme dans les Énoncés. Ces places sont notées :

Notation	Place du terme
S	Sujet
V-1	Verbe au mode actif
V-2	Verbe au mode infinitif
C-1	Premier complément
C-2	Deuxième complément
RAP	Le terme est un référant
DANS	Complément circonstanciel de temps ou de lieu
SP	Le terme est spécifiant d'un autre terme qui est spécifié

Le tableau de fréquences des places n'a qu'un intérêt réduit tout au moins de notre point de vue. Il est surtout utilisé comme garde-fou pour les réductions polysémiques du vocabulaire particulier à chaque entretien. Il est aussi utilisé pour repérer des termes jouant un rôle particulier du point de vue syntaxique si ce rôle n'est pas trivial.

L'analyse distributionnelle du terme a un plus grand intérêt. Elle consiste en l'observation du contexte dans lequel se trouve ce terme. Pour cela, on écrit suivant la grille du tableau de fréquence tous les énoncés où se trouve ce terme. L'analyse de tous les Énoncés Minimaux formant le contexte d'un terme permet l'analyse de la signification de ce terme. Elle permet aussi de repérer les termes qui apparaissent ensemble et ceux qui ont des significations proches ou synonymiques. En outre l'analyse entre plusieurs entretiens est particulièrement fructueuse, car elle permet de voir les significations différentes que prend un même terme dans des groupes sociaux différents.

Un autre intérêt de l'analyse distributionnelle se trouve dans la sélection d'Énoncés Minimaux souvent très proches les uns des autres. Ils peuvent suggérer l'étude d'une classe de termes commutants au regard du terme dont on a la distribution. Il suffit ensuite de regarder tous les Énoncés Minimaux, où se trouvent ces termes; c'est-à-dire qu'il suffit de se reporter aux distributions de chacun des termes qui commutent. On observera alors la signification de ces commutations en particulier si l'on peut trouver un Énoncé qui relie syntagmatiquement deux des termes commutants.

On peut ici décrire un exemple à partir du listing joint : En observant la distribution du terme "Prévoir" on entrevoit une commutation possible entre

"gain" et "contrainte naturelle" mais la liaison entre ces deux termes ne nous est pas donnée. En se reportant à "gain" on n'obtient aucune information nouvelle, il n'apparaît que dans un seul Enoncé Minimal. Par contre en se reportant à "contrainte" on voit que ce terme commute avec "Production" (SPE - récolte annuelle) et que cette production renvoie à la référence "contrainte naturelle".

On a ainsi une figure à quatre positions :

Gain Contrainte naturelle

Production RAP Contrainte naturelle

qui fait penser à la synonymie Gain - Production, et à une liaison "RAP" entre Gain et Contrainte naturelle, même si elle n'est pas dite en ces termes dans le discours.

5.1.2. Champs sémantiques

L'analyse du contexte élargie au niveau de l'ensemble des arbres où un terme apparaît (une ou plusieurs fois) est lourde si l'on veut en faire une exploitation systématique aussi préfère-t-on analyser le champ sémantique de certains termes, choisis pour leur intérêt particulier eu égard aux demandes des formateurs.

Ce traitement fait appel à la constitution d'un fichier annexe dit "fichier inverse des arbres", qui, adressé par un terme, permet de connaître l'ensemble des arbres où se trouve ce terme. Pour effectuer une telle opération sur un vocabulaire de 500 termes, et un fichier de 25.000 Enoncés Minimaux, en utilisant l'extrême rapidité d'une mémoire centrale d'ordinateur mais sans utiliser une trop grande place

0	I	C	I	6	I	0	I	0	I	C	I	0	*	0
	I		I		I		I		I		I		*	
	I		I		I		I		I		I		*	
6217		IMPOSSIBILITE	=PREVOIR		GAIN					AGRICULTURE				
6218		POSSIBILITE	=PREVOIR							VITICULTURE				
6218			*FACILEMENT										***	
6218			*PLUS										***	
6218		IMPOSSIBILITE	=PREVOIR		CONTRAINTE									
6218					*NATUREL								***	
6219		IMPOSSIBILITE	=PREVOIR											
6220		POSSIBILITE	=PREVOIR							NATION				
6220										EXPLOITATION				
6225		IMPOSSIBILITE	=PREVOIR											

**
*** GAIN
**

S	I	V-1	I	V-2	I	C-1	I	C-2	I	DANS	I	RAP	*	SP
*****	I	*****	I	*****	I	*****	I	*****	I	*****	I	*****	*****	
	I		I		I		I		I		I		*	
C	I	C	I	0	I	1	I	0	I	C	I	0	*	0
	I		I		I		I		I		I		*	

6217		IMPOSSIBILITE	PREVOIR	=GAIN		AGRICULTURE
------	--	---------------	---------	-------	--	-------------

**
*** CONTRAINTE
**

S	I	V-1	I	V-2	I	C-1	I	C-2	I	DANS	I	RAP	*	SP
*****	I	*****	I	*****	I	*****	I	*****	I	*****	I	*****	*****	
	I		I		I		I		I		I		*	
0	I	C	I	0	I	1	I	0	I	C	I	1	*	0
	I		I		I		I		I		I		*	

6218		IMPOSSIBILITE	PREVOIR	=CONTRAINTE		
6218				*NATUREL		***

6335		DETERMINER		PRODUCTION		DELAI-SP-RIGIDE
6335				*RECOLTE		=CONTRAINTE
6335				*ANNUEL		*NATUREL

**
*** PRODUCTION
**

S	I	V-1	I	V-2	I	C-1	I	C-2	I	DANS	I	RAP	*	SP
*****	I	*****	I	*****	I	*****	I	*****	I	*****	I	*****	*****	
	I		I		I		I		I		I		*	
1	I	0	I	0	I	2	I	0	I	C	I	0	*	0
	I		I		I		I		I		I		*	

6220 = PRODUCTION	SUPPORTER	CHANGEMENT		
6220		*PEU		***

6335 INDUSTRIEL	AUGMENTER	=PRODUCTION		
6335	DETERMINER	=PRODUCTION		
6335		*RECOLTE		DELAI-SP-RIGIDE
6335		*ANNUEL		CONTRAINTE
				*NATUREL

66

dans cette mémoire, il a fallu mettre au point un chaînage dans le tableau Terme x arbre. Ce tableau ($T \times A$) devient alors un tableau unidimensionnel à utilisation dynamique, formé de cellules doubles où l'on trouve : en première position le numéro de l'arbre et de l'interview qui contient le terme x, et en seconde position l'adresse de la prochaine cellule du tableau ($T \times A$) où l'on trouvera le prochain arbre qui contient le même terme x.

Pour suivre ce cheminement dans le tableau, il nous faut une série de pointeurs qui indiquent pour chaque terme le numéro de la première cellule où l'on trouve le premier arbre qui contient ce terme, et le numéro de la dernière cellule où l'on trouve le dernier arbre qui contient ce terme.

Lors de la création du tableau ($T \times A$) chaque terme de l'arbre que l'on dépouille donnera lieu à la création d'une nouvelle cellule; puis à la sélection d'un pointeur; puis à la création du chaînage entre la précédente cellule, marquée par le pointeur, et la nouvelle cellule; enfin du déplacement du pointeur sur la nouvelle cellule. On peut indiquer ces opérations sur la figure suivante. Soit l'arbre A_0 et le terme x.

1) On indique en m l'arbre A_0

A1				A_0	
n	n+1		m-1	m	



2) Le pointeur P (x) contient n donc le précédent arbre où l'on trouve x est A_1 .

3) On donne à $n + 1$ la valeur m d'où chaînage.

A1	m			A_0	
n	n+1			m	m+1

4) On modifie le pointeur $P(x)$ en lui donnant la valeur m , le tableau (TxA) est alors utilisé jusqu'à la position $m+1$ laissée vide pour l'instant dans l'attente d'un prochain chaînage.

5) On passe au terme suivant.
Cas particulier si la précédente cellule de x contient le numéro de l'arbre que l'on est en train de dépouiller, on ne recrée pas une nouvelle cellule, le pointeur du tableau (TxA) est remis à la valeur $m - 1$ et on ne modifie pas la valeur du pointeur $P(x)$ qui reste n .

L'ensemble des arbres étant passé en revue on édite le tableau en reprenant terme par terme le chaînage qui nous amène du premier arbre où ce mot est apparu à son dernier arbre. On obtient alors un enregistrement dont la première zone contient le numéro du terme, et les suivantes les numéros d'arbre et entretiens où l'on trouve ce terme.

A partir de ce dernier fichier, il est facile d'étudier le champ sémantique d'un thème. Le questionnement n'est qu'un problème d'union et d'intersection des sous-ensembles d'arbres contenant les termes proches du thème que l'on veut analyser. Pour ce faire, le programme a en entrée une carte lui indiquant les termes qui cernent un champ sémantique. Comme par exemple : "Entreprise ET Investissement". Les seules relations entre les termes sont booléennes (ET, OU).

L'expression formée par les termes et les relations est logique, elle est enregistrée sous forme Polonaise. Le programme lit ensuite le fichier inverse et sélectionne les enregistrements correspondants aux termes de la liste. Elle traite l'expression booléenne des termes en considérant chaque enregistrement comme un "ensemble" d'arbre (au sens

66

mathématique du terme). Les opérations effectuées par le programme sont alors l'intersection ou l'union de ces ensembles. Le programme sort ensuite la liste des arbres, qui recouvre le champ analysé. Il peut sur option éditer ces arbres.

5.1.3. Programme d'analyse statistique

Pour avoir les éléments statistiques permettant d'orienter des opérations de sélection du vocabulaire ou d'importance des relations, nous avons créé deux programmes d'analyse statistique. L'un est relatif aux termes, l'autre relatif aux relations. Le programme d'analyse statistique des termes a en entrée le fichier inverse. Il édite pour chaque terme, la fréquence d'apparition du terme chez chacun des groupes d'interviewé :

- cadre autodidacte
- cadre diplômé
- ouvrier C.G.T.
- ouvrier C.F.D.T.
- ouvrier entreprise
- agriculteur Paris
Bouches-du-Rhône
Drôme
- Travailleurs sociaux.

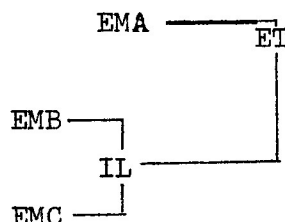
Il édite ensuite les numéros d'entretien où apparaît ce terme ainsi que sa fréquence d'apparition.

Le programme d'analyse statistique des relations repère l'occurrence d'une relation et donne deux matrices indiquant la fréquence de succession des relations deux à deux. Pour cette analyse, il nous faut repérer deux informations distinctes :

- la succession des relations dans le discours du type : EMA.ET.(EMP.IL.EMC), on dit ici que

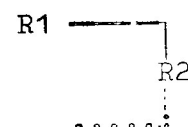
l'implication (IL) suit le ET.

- la supériorité d'une relation sur une autre, supériorité au sens où l'une des relations est dans l'arborescence à un niveau supérieur, c'est le cas de ET, dans notre exemple.

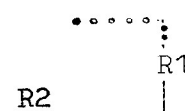


Le croisement de ces deux informations donne les configurations suivantes :

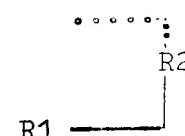
- a) R1 est suivi par R2
qui lui est supérieur



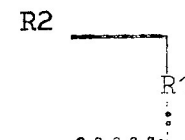
- b) R1 est suivi par R2
qui lui est inférieur



- c) R1 suit R2 qui lui est
supérieur



- d) R1 suit R2 qui lui est
inférieur



On voit que les cas a et d sont semblables à une permutation de R1, R2 près. De même pour les cas b et c.

Les deux informations que nous voulons obtenir peuvent donc être condensées en deux matrices: l'une indiquant la relation de succession et la re-

lation supériorité (S); l'autre la relation de succession et d'infériorité (I).

(S)

	R1	R2
R1	a	
R2	d.....	

(I)

	R1	R2
R1	b	
R2	c.....	

Les matrices mettent en évidence des phénomènes fins du langage mais sont fortement contaminées par les règles de la procédure du codage. Aussi, nous éditons ces deux matrices pour chaque interview afin d'avoir les moyens de vérification et d'interprétation des éléments apparus. L'entrée du programme est constituée par la base de données.

5.2. Traitements des arbres.

Avec le traitement des arbres, nous entrons vraiment dans la procédure d'interrogation de la base de données qui répond aux principes généraux suivants : On pose, sur les termes du vocabulaire, à partir des hypothèses et des résultats del'analyse contextuelle des E.M., un certain nombre de catégories a priori qui n'ont qu'un rôle opératoire de questionnement.

- La base de données est conçue comme constitution d'un ensemble d'arbres qui forment autant de chaînes de signifiants.
- Les questions sont conçues comme des "Formes", c'est-à-dire comme des enchaînements particuliers de signifiants assortis de règles de réécriture.
- Le questionnement consiste à sélectionner les arbres qui ont même forme que la question ou qui sont compatibles avec cette forme.

La forme des questions procède d'une part d'hypothèses théoriques, d'autre part de leur tra-

duction opératoire provisoire en terme de classes, de relations et de termes. Les hypothèses a priori sont nécessaires pour analyser les arbres sélectionnés, c'est-à-dire pour les replacer dans un cadre explicatif, ces arbres infirmant ou vérifiant ce cadre. Leur traduction opératoire est surtout nécessaire au niveau du processus de reconnaissance (de questionnement) aussi est-elle toute provisoire et guidée aussi par des impératifs pratiques. On analyse l'ensemble des arbres sélectionnés pour vérifier la validité du questionnement (et non la validité des catégories a priori).

Les programmes que l'on peut envisager à partir de ces principes d'interrogation de la base de données sont nombreux et n'ont pas tous reçu leur plein développement. Nous ne traiterons ici, que ceux qui servent à l'analyse des arbres suivant les niveaux de discours, les champs dont on parle et les relations entre ces champs. (Pour la définition de ces termes se reporter au rapport d'étude).

5.2.1. Recodage des arbres à partir de quatre catégories.

On définit quatre catégories a priori que l'on appelle champs.

- Economique (acte de Production et de consommation des biens, circulation monétaire, acte supposant l'échange monétaire).

- Politique (acte de pouvoir, institution politique, et rapport de force).

- Social (acte quotidien, description sociale ...)

Puis on analyse chaque Enoncé Minimal en fonction des termes qu'il contient et du contexte minimal qu'il forme.

En même temps on code chaque Enoncé Minimal comme formé d'une suite de catégories et de relations. Ces relations sont très simplifiées : le "=", le "+", représentent la finalité et l'implication.

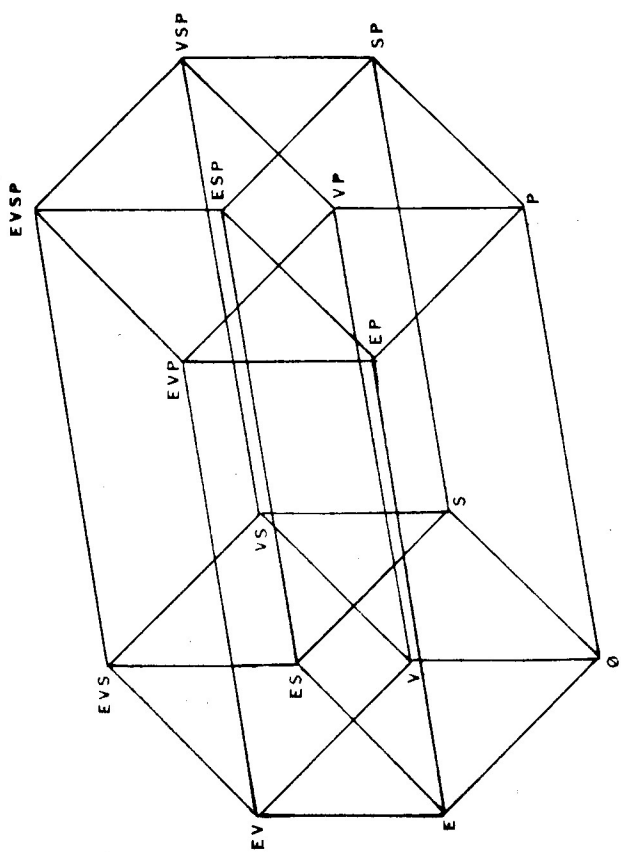
En dernier lieu, on repère pour chaque énoncé le niveau du discours (particulier, général, notionnel).

On obtient après traitement un enregistrement par Enoncé Minimal qui contient :

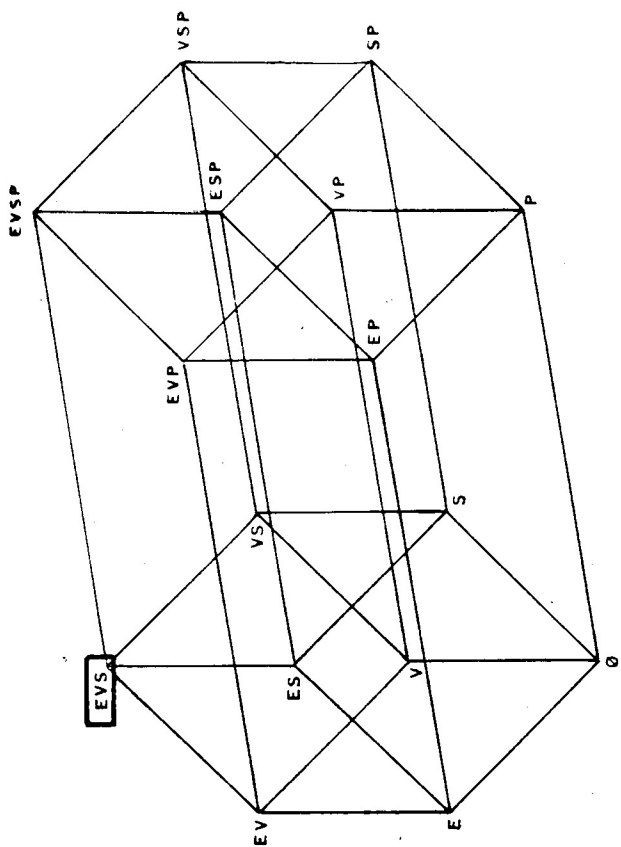
- le numéro d'interview
- le numéro d'arbre
- un quadruplet booléen qui indique l'apparition de E, V, P, S dans l'énoncé minimal.
- le niveau du discours
- le thème (ou question de l'entretien) où se trouve l'arbre
- la variable Distribution
- le recodage de l'énoncé (comme par exemple : $SFE = V$)
- la décomposition en triplet de cet énoncé (ici $SFE, E = V$).

5.2.2. Traitement en treillis

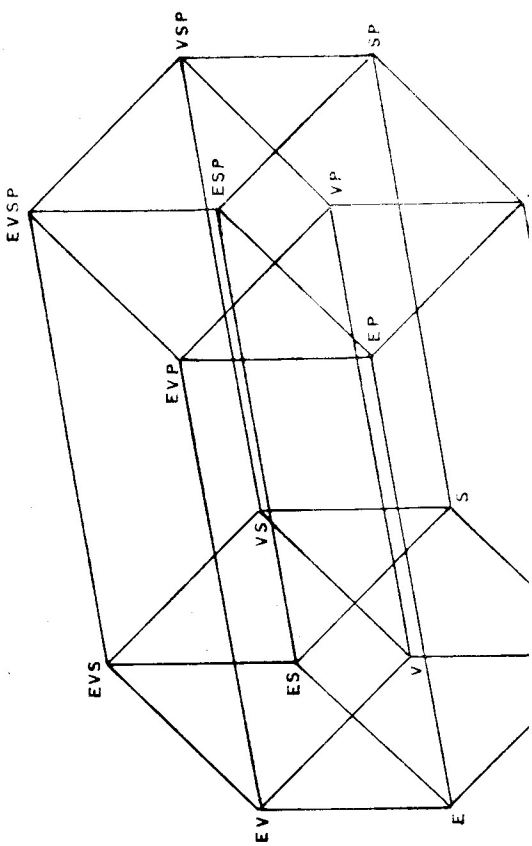
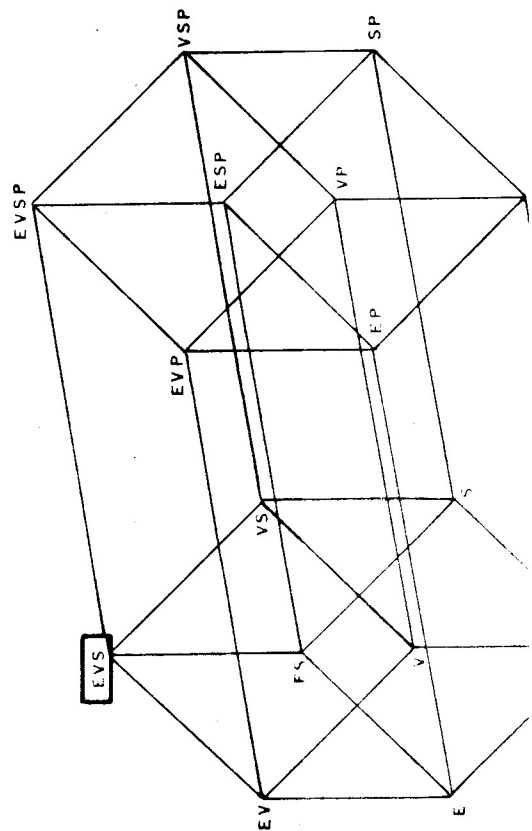
A partir des enregistrements précédents, on veut obtenir le treillis des champs. Il est en effet possible d'utiliser le quadruplet booléen indiquant l'apparition des champs pour remplir un tableau que l'on place ensuite sur un treillis (ici l'exemple : $SFE = V$ serait situé en EVS voir figure). Notre programme d'édition des treillis permet la



NOTIONNEL



ENSEMBLE



permutation de deux ou plusieurs colonnes de notre quadruplet booléen ainsi que l'inversion de l'une des variables booléennes (E transformé en Non E). Chaque sommet du treillis est affecté de sa fréquence d'apparition. On peut alors noter ceux qui ont les fréquences d'apparition les plus élevées et l'édition directe du treillis fait apparaître par graphique les phénomènes d'implication logique entre champs, et les phénomènes d'échelle ou de tresse entre ces champs.

Nous effectuons un traitement par entretien qui comprend une série de treillis : à partir du treillis d'ensemble, nous pouvons le décomposer en treillis apparaissant dans chaque niveau de discours (cf. figure). La série des quatre treillis : Particulier, général, notionnel, ensemble, peut être éditée pour chaque thème, et pour chaque variable Distribution.

Le même traitement peut être effectué pour un groupe d'entretien afin d'opposer synthétiquement ces groupes.

5.2.3. Traitement complémentaire au niveau de l'Enoncé Minimal.

Un premier traitement complémentaire consiste en l'analyse des relations entre champs au sein de l'E.M. Il se fait en traitant les triplets : champs-relation-champs. Il consiste en l'édition d'une matrice de relation entre les champs pour une relation donnée : il y a autant de matrices que de relation. Cette série de matrices de transition est éditée pour l'ensemble d'un entretien puis pour chaque niveau du discours et chaque variable distribution.

Les matrices de transition ont été envisagées pour compléter l'analyse du treillis. Elles permettent une analyse fine du type de relations qui peuvent exister entre certains champs qui apparaissent souvent ensemble.

Un deuxième traitement complémentaire consiste en la sélection d'arbres satisfaisant à certains critères. On recherche en particulier les arbres pour lesquels il y a eu changement de niveaux de discours comme le passage de l'élémentaire au notionnel. On repère les arbres qui comprennent :

- les trois niveaux
- deux de ces niveaux
- un seul niveau

Ils peuvent être édités sur option et l'on peut les croiser avec les variables thèmes et distribution.

On recherche enfin les arbres homogènes, pour un champ donné. Le traitement sert à repérer l'autonomisation de l'économie. Il suffit de se donner un critère tel que "on désire les arbres pour lesquels les énoncés minimaux comportent tous le champ économique", ou encore "le comporte dans la proportion de 80 %". On peut avec le même programme repérer les arbres articulant deux champs particuliers comme Economique et Social ou Economique et Politique.

5.2.4. Traitement des enchaînements au niveau des ARBRES.

L'automatisation de ces traitements se heurte encore actuellement à de nombreuses difficultés. Il est possible d'interroger la base de données

avec des questions assez précises telles que "arbre ayant les termes x, y, z, reliés ainsi ((x.II.y). PR.z), mais au point où nous en sommes il est encore difficile de mettre au point un dépouillement systématique. Nous connaissons bien certaine forme, telle que la suite de ET, ET, ET ... qui indique le plus souvent si l'on est en notionnel, une liste de termes formant une certaine catégorie pour l'interviewé. Mais la colonne de ces formes n'est pas encore totalement connue.

Nous envisageons actuellement un premier traitement manuel permettant la mise à jour de raisonnement type. Ceux-ci étant repérés ils seront programmés pour permettre leur validation sur la base de données.

5.3. Conclusion sur le traitement informatique.

Les programmes que nous avons décrits ici ont une certaine généralité. Ils sont adaptés à tout codage qui distingue des éléments de texte mis en relation. Le codage par ailleurs peut affiner l'analyse de ces éléments de texte en distinguant des termes et des relations mineures ou toute autre forme de positionnement des termes. Enfin le vocabulaire des termes peut être constamment enrichi.

Ces programmes permettent l'écriture des textes sous une forme arborescente ou sous forme de réseau (avec des conventions particulières). Ils permettent enfin leur questionnement grâce à la création d'une base de données à structure tout à la fois de pile et de liste; la base de données pouvant être enrichie d'éléments extérieurs au texte, tels que des notations sur le déroulement de l'entretien ou sur les caractéristiques du sujet enquêté, ces renseignements pouvant se situer indifféremment au

niveau des arbres ou des E.M.

Dans notre cas, les termes et relations sont contingents à la spécificité des textes traités et à la problématique de notre étude. Mais on peut dire que les traitements, bien que liés à nos hypothèses générales, ont acquis une certaine indépendance, qui leur permet de ne pas être liés à cette contingence.

Les traitements que nous avons envisagés ne sont pas limitatifs, on peut se poser d'autres questions, envisager un recodage différent à partir du vocabulaire et des Enoncés Minimaux. Nous avons essayé de bâtir une structure de données la plus simple possible afin de laisser libre cours aux hypothèses multiples que l'on peut avoir à propos des représentations économiques.

ANNEXEGUIDE D'ENTRETIEN1 - GRILLE PRINCIPALE

- 1 - Pouvez-vous me parler de l'argent ?
- 2 - Pouvez-vous me parler de l'organisation de votre budget ?
- 3 - Pouvez-vous me parler de vos économies ?
- 4 - Pouvez-vous me parler de l'achat à crédit ?
- 5 - L'argent que vous dépensez, c'est l'argent que d'autres gagnent, réciproquement l'argent que vous gagnez, c'est l'argent que d'autres ont dépensé. Est-ce que vous pouvez me parler de l'argent (pour l'ensemble des gens du pays), pour l'ensemble du pays ?

Le temps approximatif proposé à l'enquêteur pour l'exploration des questions est de 30 à 45 minutes pour la première question et un temps équivalent pour les quatre dernières.

Si l'un des thèmes avait été abordé longuement dans le cours d'une précédente question, il n'était, évidemment pas nécessaire de poser la question correspondante, sinon sous forme de reformulation de ce qui avait été dit par l'interviewé.

2 - EXPLICATION COMPLEMENTAIRE

A la fin de l'entretien, si certains sous-thèmes n'avaient pas été abordés, ils l'étaient sous forme de petites questions :

- BUDGET : sous-thèmes : Prévision, choix

72

+ Dans Prévision faire préciser :

- Quelles sont les dépenses qui font l'objet d'une prévision ?
- Quels sont les critères de détermination de ces dépenses ?

+ Dans Choix, faire préciser :

- Quelle est la hiérarchie des dépenses ?
- Quels sont les critères de différenciation des dépenses (fondamentales, superflues ...). Ces critères interfèrent avec les précédents (prévision).

- REVENU : sous-thèmes : affectation de certaines sources de revenus, avantage en nature, changement dans l'organisation du budget.

- Y a-t-il des revenus particuliers (au sens de sources de revenus).
- Ces revenus particuliers sont-ils affectés à des dépenses particulières ? (critères de ces affectations).
- Y a-t-il des avantages en nature (ex. : logement gratuit) et les prend-il en considération dans son revenu ?

- ECONOMIES : sous-thèmes : Prévoyance, Placement, objectif particulier.

- Vous est-il arrivé de faire des économies en vue d'assurer votre sécurité ou de prévenir une difficulté ?
- Vous est-il arrivé de faire un placement d'argent ? (dans le cas où il n'y a pas compréhension ajouter : à la caisse d'épargne par exemple).

20

- Vous est-il arrivé de faire des économies en vue d'un achat particulier ?

- ACHAT A CREDIT : sous-thème : ce qu'en disent les autres ?

- Dans le cas où le sujet n'aurait cité que des exemples le concernant, et ne serait jamais situé par rapport aux autres, lui demander ce que d'autres en situations différentes (familiale, ou de revenu) disent de l'achat à crédit.