

[Saadi Lahlou](#)

## La modélisation de représentations sociales à partir de l'analyse d'un corpus de définitions

### Book section

**Original citation:**

Lahlou, Saadi (1996) *La modélisation de représentations sociales à partir de l'analyse d'un corpus de définitions*. In: Martin, Eveline, (ed.) *Informatique textuelle. Etudes de sémantique lexicale* (3). Didier Erudition, Paris, France, pp. 55-98. ISBN 9782864602774

© 1996 [Didier Erudition](#)

This version available at: <http://eprints.lse.ac.uk/28372/>

Available in LSE Research Online: March 2011

LSE has developed LSE Research Online so that users may access research output of the School. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Users may download and/or print one copy of any article(s) in LSE Research Online to facilitate their private study or for non-commercial research. You may not engage in further distribution of the material or use it for any profit-making activities or any commercial gain. You may freely distribute the URL (<http://eprints.lse.ac.uk>) of the LSE Research Online website.

This document is the author's submitted version of the book section. There may be differences between this version and the published version. You are advised to consult the publisher's version if you wish to cite from it.

(1996) - LAHLOU, Saadi. La modélisation de représentations sociales à partir de l'analyse d'un corpus de définitions. In : Eveline Martin (éd.). *Informatique textuelle*. Coll. Etudes de Sémantique Lexicale. Institut National de la Langue Française. Paris : Didier Erudition. pp. 55-98.

## La modélisation de représentations sociales à partir de l'analyse d'un corpus de définitions

Saadi LAHLOU, EDF - Direction des Etudes et Recherches

Malheureux lecteur, je vous dois une explication sur la construction peu académique de ce texte. Il est tard, Je suis en retard pour livrer mon papier et, l'ayant relu encore, je trouve toujours la partie théorique absconse, prétentieuse, et insuffisamment aboutie. Je commencerai donc par vous livrer les résultats, pour ne pas vous décourager. Il s'agit d'un essai empirique pour décrire le sens des mots.

La recherche du sens est une quête ardue. Si, au plan théorique, son aboutissement absolu paraît, pour le moins, lointain, il semble que, localement, des méthodes d'analyse lexicale permettent d'obtenir une représentation approximative du sens commun, c'est-à-dire, pour dire vite, de ce qu'une culture met derrière un mot particulier. Je présente ici un protocole expérimental qui permet, me semble-t-il, d'obtenir une représentation du sens d'un mot.

Cette méthode peut s'avérer utile pour repérer les différentes dimensions sémantiques d'un terme, à des fins de recherche exploratoire. Deux exemples : les mots "manger" et "information", fournissent une illustration empirique qui valide la méthode.

Je présente à cette occasion la théorie de l'interprétation des corpus d'énoncés en langue naturelle qui sous-tend la méthode. Celle-ci est encore, comme je l'ai dit, incomplète. Peut-être des chercheurs plus compétents pourront-ils en faire quelque chose.

# 1. Un premier exemple : manger

C'est un des rôles de la langue que de servir de catalogue public des objets partagés par les membres d'une collectivité. Ce grand catalogue se présente sous la forme d'un réseau dans lequel les définitions, les connotations de chaque mot, en renvoyant à d'autres mots, fournissent une image du rapport entre les objets auxquels les mots réfèrent, dans le monde tel qu'il est vécu par les membres de cette collectivité. Le sens d'un mot, c'est le rapport que l'objet qu'il désigne entretient avec les autres objets du monde. Par l'analyse du rapport entre mots dans la langue, on espère récupérer une représentation de ce sens.

Pratiquement, on recueille auprès d'une source informée, et capable de s'exprimer en langue, un grand nombre d'énoncés "à propos de" la chose, puis on construit, par analyse des cooccurrences lexicales, un champ de connotations. On partitionne celui-ci en classes, dont les représentants canoniques sont alors assimilés à des noyaux de sens, dont le réseau constituerait le sens du mot.

Je présenterai une application de la méthode utilisant comme source de connaissance le dictionnaire.

Pourquoi le dictionnaire ? Car il contient du savoir social sur le monde, sédimenté en langue. Notre société se construit en certains lieux des modèles d'elle-même, où elle formalise sa vision du monde. Ces lieux sont les encyclopédies et les dictionnaires ; et ces modèles sont aux représentations sociales (Moscovici, 1961) que chacun de nous partage ce que le mètre-étalon en platine iridié conservé au pavillon de l'Institut International des Poids et Mesures, à Sèvres, était à toutes les règles graduées et aux mètres pliants du monde : la norme de référence.

L'encyclopédie se réfère aux objets du monde, et le dictionnaire aux mots. Mais ils sont finalement assez proches comme le montre Meschonnic (1991)<sup>1</sup>. On peut à bon droit considérer les deux comme des tentatives de décrire l'Umwelt (l'entour) de l'homme de manière systématique.

Les encyclopédies et les dictionnaires contiennent de la représentation publique au sens de Sperber (1989). Publique pour tous, elle est notamment d'accès facile au chercheur. Il y a là un extraordinaire gisement de connaissance sociale qui ne demande qu'à s'ouvrir à l'investigation systématique.

Par rapport à l'encyclopédie, le dictionnaire offre théoriquement l'avantage d'être un système fermé et entièrement auto-défini, sans référence au monde extérieur : ce sont les mots qui sont les objets désignés ; nous nous situons alors dans un

univers discursif cohérent. Pratiquement, l'avantage est que tous les mots utilisés comme traits de description y figurent également comme entrée de définition. On a donc un réseau dans lequel tous les traits ont le même statut, contrairement à l'encyclopédie où seuls les mots "importants" ont une entrée.

Le grand Robert est un assez gros dictionnaire, puisqu'il représente, sous forme imprimée, 9 volumes, totalisant 9 440 pages. Il contient 100 000 entrées, soit 80 000 formes canoniques et des formes fléchies, 160 000 citations et 1000 000 renvois analogiques.

Pour lui appliquer la méthode d'association libre, nous allons utiliser son aspect analogique. Chaque entrée comporte, après la définition, une liste des termes associés : synonymes, analogues homonymes et dérivés. Nous allons recueillir ces associations, comme si le Grand Robert était un *porte-parole* de notre culture.

Pour "manger", nous disposons ainsi d'une première liste d'associés, et chacun d'eux va nous fournir, à son tour, une liste d'associés. Nous nous arrêterons aux mots de ces deux premières listes. Sinon, on risque fort de se retrouver avec l'ensemble des 100 000 entrées du dictionnaire, car la langue est un réseau connexe.

Les définitions ont été extraites manuellement du Grand Robert Électronique (édition 1991). Nous obtenons ainsi un corpus constitué des définitions de tous les synonymes, analogues, et dérivés du mot (de l'ordre du Mega-octet), soit 544 définitions. À l'aide du logiciel ALCESTE (Reinert, 1987a et b, 1990), le corpus est segmenté en énoncés, ou plus exactement en unités de contexte élémentaires (uce) d'une ou deux lignes, respectant la ponctuation forte. Pour notre corpus, on obtient ainsi quelques milliers d'uce. Une classification descendante hiérarchique de ces énoncés, sur la base de leur contenu lexical, fournit des classes d'énoncés qui correspondent à des champs de connotation du mot initial.

On procède de la manière suivante. On scinde d'abord le corpus en deux sous-ensembles d'uce les plus éloignés possibles C1 et C2 de manière à ce que les uce à l'intérieur d'une classe soient très proches en termes de vocabulaire utilisé, et le plus différentes possible de celles de l'autre classe<sup>2</sup>. Ensuite, on réitère, à chaque étape avec la plus grosse classe restante, créant ainsi une arborescence descendante. On finit ainsi par obtenir des classes de plus en plus petites, chacune séparée *par contraste* avec les autres. Ces classes rassemblent des énoncés analogues en ce qu'ils contiennent les mêmes unités lexicales, en l'occurrence des mots "stemmatisés" (réduction à une forme canonique, ou à la racine, par lemmatisation et/ou troncature à droite), ou des locutions reconnues.

Je considère les classes obtenues comme étant les éléments de base constitutifs de la représentation sociale, autrement dit du “ sens commun ” auquel renvoie le mot.

Voici un exemple simplifié. Il est clair intuitivement que les deux énoncés :

- (1) fruit légume viande ou poisson céréales laitages
- (2) viandes fruits légumes pain et un coup de rouge ! et du fromage

seront proches l'une de l'autre, mais éloignées de :

- (3) à la bonne cuisine à la maison avec des produits naturels

qui sera lui-même plus proche de :

- (4) une bonne bouffe à la française plats cuisinés maison

Le logiciel retrouve ces proximités en calculant une distance lexicale. Nous calculons les distances sur les réponses considérées comme "sacs de racines de mots". Les deux derniers énoncés deviennent :

- (3') bon+ cuisin+ maison avec produit naturel+
- (4') bon+ bouffe+ français+ plat+ cuisin+ maison

(les "+" et autres caractères spéciaux renvoient au type de reconnaissance de la racine par le logiciel). Dans ce cas particulier, les deux énoncés ont en commun 3 formes (*bon+*, *cuisin+*, *maison*) sur les 6 que chacun contient. Compte tenu de la fréquence d'apparition des mots dans le corpus, la probabilité d'une telle cooccurrence est très faible, et la distance lexicale entre les énoncés est donc très petite. On comprend bien que ces deux énoncés vont se retrouver dans la même classe.

Mais examinons maintenant le cas suivant :

- (w) *repas copieux avec entrée plat\_résistance et dessert*

se retrouve dans la même classe que :

- (z) **des crudités de la viande et des légumes de la salade et un fruit**

ce qui est sémantiquement satisfaisant, mais à première vue surprenant, puisque ces deux énoncés n'ont aucun mot plein (actif) en commun. Cela s'explique par la présence dans la même classe d'autres énoncés qui sont des formulations intermédiaires du même paradigme, ayant des mots communs avec ces deux formes extrêmes, comme :

- (x) à un bon *repas entrée plat\_résistance* fromage **salade**  
*dessert café*

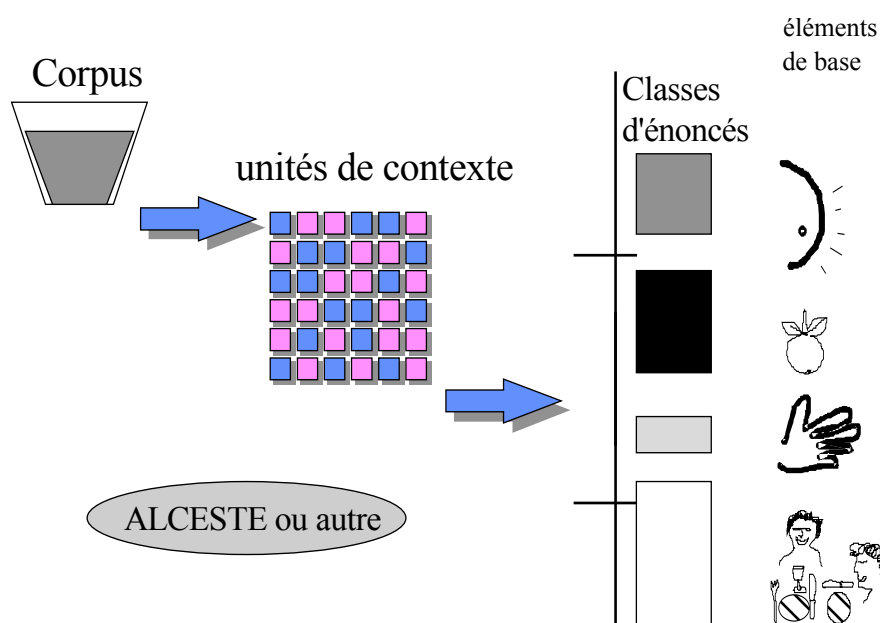
(y) viande les légumes dessert les fruits le pain le fromage

qui sont proches des deux formes extrêmes, et également proches entre elles par cooccurrence de *fromage* et *dessert*.

La classification qui combine analogie et contraste, permet donc de rassembler des énoncés de même "sens".

La figure suivante résume la démarche adoptée pour l'analyse d'un corpus.

Fig. 1 : Procédure d'analyse du corpus



## 1.1 Structure du lexique

Le corpus complet est constitué d'un lexique d'environ 17 000 mots différents. Après lemmatisation et élimination des mots grammaticaux, nous n'avons conservé que les 10 000 racines les plus fréquentes, et pour la construction des classes, n'ont été actives que les 828 racines apparaissant plus de 22 fois dans le corpus.

On présente ci-dessous les principaux indicateurs lexicaux, avec, à titre illustratif, une comparaison avec un corpus d'évocations libres d'un "petit déjeuner idéal" par des consommateurs (Beaudouin, Lahlou et Yvon, 1993), et avec les oeuvres complètes de Racine et Corneille. Cette comparaison met en avant l'extrême richesse lexicale du corpus (Beaudouin et Hébel, 1994)

Fig. 2 : Indicateurs statistiques de caractérisation des corpus utilisés

Nom du corpus	V	N	N/V	Fréquence médiane	V1 Hapax	V1/V	% V (freq. <45)	% N (freq. <45)	% N 10 freq. max.
Manger	16 896	137 576	8,1	118	8734	52	98,1	40,8	19,9
Corneille	14 055	547 297	39	249	5060	36	91,7	13,8	20,0
Racine	9 288	164 845	17,7	151	3805	41	95,5	26,4	19,6
Petit déjeuner	879	11 292	12,8	49	440	50	94,5	29	34,3

Légende :

**V** : Nombre de vocables (étendue du vocabulaire)

**N** : Nombre de mots (étendue du texte)

**N / V** : Fréquence moyenne (Nombre moyen d'occurrences)

**V1** : Nombre d'hapax : nombre de mots n'apparaissant qu'une seule fois

**V1 / V** : proportion en vocables de fréquence 1

**% V (freq<45)** : Pourcentage de vocables représentés par les fréquences de 1 à 45

**% N (freq<45)** : Pourcentage d'occurrences représentées par les fréquences de 1 à 45

**% N 10 freq max** : proportion des mots du corpus représenté par les dix fréquences les plus élevées

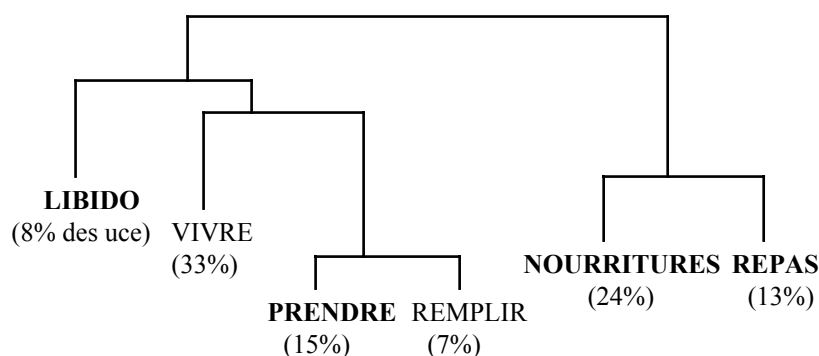
La distribution des fréquences d'occurrence des formes réduites a une allure conforme à ce que l'on peut attendre d'un texte en langue naturelle (courbe de fréquences de type Zipf "en baignoire").

## 1.2 Les noyaux de base du manger

Je donnerai les noms des classes avant d'avoir décrit leurs traits. Cette procédure anticipe un peu sur l'analyse (que j'ai, évidemment, faite avant le lecteur) dans la présentation des résultats. Mais il faut bien garder à l'esprit que ces noms résultent de mon interprétation des traits typiques, et ne sont pas fournis par le logiciel.

L'arbre de classification a la forme suivante :

Fig. 3 : Arbre de classification de l'analyse du corpus "Manger" du Grand Robert



### 1.2.1 LIBIDO

La description de cette classe sera l'occasion de montrer en détail les aides à l'interprétation fournies par le logiciel.

Cette classe, comme les suivantes, est caractérisée par des traits très typiques, au sens statistique, comme on le voit sur le tableau suivant.

occ. cl.	occ. tot.	%	chi2	trait
87	101	86	869	désir+
62	75	83	587	faim
54	67	81	495	appétit+
47	52	90	494	soif
44	55	80	399	satisfaire.
52	77	68	384	envie+
32	32	100	377	convoit+
31	32	97	352	assouvi+
31	33	94	340	rassasi+
24	27	89	246	avidité
24	28	86	236	apais+
32	48	67	231	dévor+
26	35	74	215	avide+
19	21	90	199	affame+
16	18	89	164	cupid+

Pour les traits (dernière colonne), classés par ordre décroissant de typicité (chi2 d'appartenance à la classe, avant-dernière colonne), on lit en première colonne le nombre d'énoncés de la classe contenant le trait, comparé (seconde colonne) au nombre d'énoncés contenant le trait dans le corpus entier. Par exemple, 86% des 101 uce contenant le trait désir+ sont dans la classe, tous les énoncés contenant convoit+ sont dans la classe<sup>3</sup>.

L'interprétation des traits montre que cette classe renvoie d'abord à une pulsion, à une appétence, la *faim*. Mais elle est plus générale dans son contenu ; c'est une



sorte de généralisation abstraite, en ce qu'elle inclut des *désirs* non alimentaires. Le caractère intense et primitif apparaît nettement avec des traits comme *avidité*, *dévor+*. Je la baptise donc LIBIDO.

Pour dire vite, les classes sont extrêmement typées ; il s'agit là de très beaux résultats d'un point de vue statistique. C'est d'autant plus remarquable qu'aucun effort n'a été fait pour optimiser les résultats de l'analyse, puisque j'ai ici utilisé, pour rendre la démonstration plus convaincante, le plan standard proposé par le logiciel, et qui n'est sans doute pas le mieux adapté à ce corpus particulier. La réaction naturelle du statisticien devant ce type de résultat est le sentiment profond que la méthode "a touché quelque chose de solide" dans la structure naturelle des données.

Voici une liste plus complète des traits typiques de la classe LIBIDO classés par ordre de significativité décroissant:

*désir+*, *faim*, *appétit+*, *soif*, *satisfaire.*, *envie+*, *convoit+*, *assouvi+*, *rassasi+*, *avidité*, *apais+*, *dévor+*, *avide+*, *affame+*, *cupid+*, *content+*, *besoin+*, *mourir.*, *arde+*, *curiosité+*, *excit+*, *yeux*, *passion<*, *regard+*, *tendance+*, *attrait+*, *glouton+*, *éprouve+*, *amour*, *sexuel+*, *instinct+*, *honneur+*, *avoir*, *brul+*, *FIG*, *moder+*, *dévor*, *soul+*, *être*, *aval+*, *inclination*

On notera ici des synonymes ou analogues du désir (essentiellement alimentaire, mais également sexuel, ou sans connotation fonctionnelle particulière) et de la violence (*dévor*, *ardent*, *extrême*, *crever*, *force*, *avide*, *mourir...*). Notons que le sens du désir est essentiellement la vue : c'est une perception de l'objet un peu à distance, contrairement à la satisfaction qui implique une prise de l'objet.

## 1.2.2 PRENDRE

Cette classe renvoie à un opérateur d'action. Ses traits caractéristiques sont essentiellement verbaux. C'est d'autant plus remarquable que le logiciel est "sourde" aux catégories syntaxiques lorsqu'il compose les classes. Cette classe d'appropriation est chargée de connotations violentes, agonistiques, agressives. On retrouve ici le second terme de l'ambivalence repérée par Fischler (1990, p. 134) : "manger c'est à la fois aimer et tuer".

Voici les traits caractéristiques principaux ( $\chi^2$  supérieur à 20) :

*touch+*, *attrape+*, *prendre.*, *main+*, *nez*, *attaqu+*, *embrass+*, *baise+*, *joue<*, *mordre.*, *ventr+*, *gonfl+*, *qqn*, *saisir.*, *parole+*, *battre.*, *lèvre+*, *ouverture+*, *doigt+*, *bras+*, *ball+*, *navire+*, *pied+*, *serre+*, *sur*, *fondre.*, *entrer.*, *avec*, *ouvrir.*

*se, PRISE, tirer., claque+, coup<, aspir+, tomb+, lui, langue<, visage, LOC, jeter., partie<, tenir., passer., 12, laiss+, dent+, mouvement+, bouche+, voir., devant, ferme+, FIGURE, arme+, porte<, bout<, contre, arrêt+, gueul+.*

*Prendre* n'est pas le mot le plus typique de la classe, mais il est de loin le trait le plus fréquent de la classe. *Toucher* et *attraper* le devancent, et sont plus typiques d'un caractère primitif et animal qui sous-tend la classe. La présence de *main, nez, joue, lèvres, bras, langue, bouche, dent, gueule, mouvement, devant, contre* peut être interprétée dans le même sens d'un pragmatisme accentué. Il est évident qu'on a affaire ici à une classe actionnelle, avec une connotation agonistique (*attaquer, embrasser, mordre, saisir, battre, serrer, fondre (sur), tirer, coup, arme ...*).

L'interprétation est presque superflue, puisque le texte parle de lui-même.

### 1.2.3 NOURRITURES

Cette classe renvoie aux objets mangés, compléments directs de l'acte de prise : les objets comestibles ; comme en témoigne la liste des traits les plus typiques.

*viande+, pain+, aliment+, fruit+, pat+, légum+, animal<, cuire., tranch+, bouill+, plant+, couper., salad+, lait+, dent+, morceau+, coût+, conserv+, rat+, digest+, porc, mange+, poisson+, sec+, fromage+, gras, végétal+, soup+, boeuf+, nourrit+, sucre+, comestible+, suc+, beurre+, tartine+, liquide+, herb+, boîte+, saucisson+, trempe+, nourr+, bouche+, maigre+, épais+, boire., gibier+, fourr+, rumin+, TECHN, oiseau+, frais+, grain+, grill+, chair+, vert+, chien+, petit+, boul+, prépar+, cuiss+, feuil+, gâteau+, croût+, potage, aval+, croqu+, fleur+, fine+, boisson+, chaud+, seche+, substance+, froid+, garni+, oeuf+, orifice+, taille+, point+, gros, produit+, arbre+.*

*Viande* (103 occurrences dans la classe) est le trait le plus typique avec un  $\chi^2$  deux fois plus élevé que son successeur immédiat, *pain* ; même si les traits *aliment* (125 occurrences), *mang+* (122 occurrences), et *pain* (102 occurrences) sont quantitativement aussi importants. On ne sera pas surpris de constater que l'aliment le plus typique, et de loin, est *viande*. Ceci appelle deux remarques. D'abord, le terme "viandes" avait, historiquement, le sens général de "nourriture". Ensuite, il ne fait pas de doute que le statut d'omnivore comporte d'abord, culturellement, celui de carnivore. Toutes les études anthropologiques confirment la valorisation de la viande comme aliment par excellence, les autres substances n'ayant parfois que le statut d'accompagnement. Il existe même dans certaines cultures une "faim de viande spécifique". Ensuite, ceci est à rapprocher

de la connotation fortement agressive de la classe PRENDRE. On comprend mieux celle-là si la nourriture est quelque chose qui se chasse, s'attrape et se tue.

Or, malgré son statut prestigieux dans la plupart des cultures, et en particulier sous forme de gibier, la viande reste en général beaucoup moins consommée que le féculent de base de la cuisine locale (pain, patate, riz, mil, igname, etc., selon le lieu). Ce sont ces substances moins prestigieuses qui ont en réalité le statut de prototype de nourriture. Dans certaines régions d'Asie, *riz* et *nourriture* sont désignés par le même mot, en France on "gagne son pain". A titre indicatif, la consommation par tête de viande de boucherie en France était de 22 kg par an en 1985, contre 46 kg de pain (INSEE). Par contre le prix de vente de la viande, qui reflète sa valeur dans notre société, est nettement plus élevé que celui du pain.

Nous pouvons donc penser que le dictionnaire nous fournit un indicateur de la *saillance* des traits des représentations plutôt que des comportements réels.

Les traits caractéristiques de la classe *NOURRITURES* décrivent des aliments de base plutôt que des préparations cuisinées. C'est sans doute dû à un effet mécanique de la maille sémantique (plus un trait décrit des objets du monde à un niveau de détail fin, plus il voit son usage en concurrence avec les autres traits qui décrivent la maille à un même niveau, et sa fréquence d'utilisation est d'autant plus faible). *Saucisson*, *gâteau* et *potage* échappent ici à cette règle, mais c'est probablement par accident. Il est à noter cependant que *plat*, qui aurait pu apparaître typique de cette classe, a sans doute été victime de sa polysémie et s'est trouvé entraîné dans la classe *REPAS*, par les arts de la table.

#### 1.2.4 REPAS

Cette classe décrit des aspects circonstanciels, instrumentaux, de la prise. *Repas* et *table* dominant largement les traits typiques (ici, ceux dont le  $\chi^2 > 20$ ).

*repas+*, *table+*, *restaur+*, *plat+*, *dîne+*, *cuisin+*, *déjeuner*, *invit+*, *serv+*, *buffet+*, *vaissel+*, *servir.*, *cantin+*, *festin+*, *couvert<*, *fête+*, *café+*, *menu+*, *noce+*, *nappe+*, *hosti+*, *gastronom+*, *soir+*, *heure<*, *assiette+*, *communi+*, *coll+*, *mange+*, *ensemble*, *thé+*, *convive+*, *soup+*, *tasse+*, *gala+*, *récept+*, *jour+*, *entrée+*, *dessert+*, *offic+*, *serviette+*, *cher+*, *spécial+*, *paye<*, *prépar+*, *léger+*, *carte+*, *pièce+*, *on*, *frugal+*, *général+*, *nuit*, *mettre.*, *boisson+*, *région<*, *cours*, *verre+*, *METONYMIE*, *hôte+*, *milieu+*, *grand+*, *après*, *chez*, *réceptant+*.

Le repas apparaît caractérisé par des marqueurs temporels dans la journée (*dîner*, *déjeuner*, *soir*, *heure*, *jour*, *nuit*, *milieu*, *après*,...), par son contenu social (*invités*, *coll+*, *convive*, *hôte*, *chez*.), et technique : arts de la table (*plat*, *buffet*,

*vaisselle, couvert, nappe, assiette, tasse, serviette, pièce, verre*), les mets, et le lieu. La classe a des connotations collectives (*ensemble, coll+*), qui apparaissent notamment par de nombreux noms d'occasions rituelles (*fête, café, noce, communion, réception, mariage...*).

On remarquera le caractère métonymique important des traits : le repas est souvent décrit par l'un de ses composants : *café, thé, buffet, ...* Ce caractère métonymique apparaît d'ailleurs en clair avec la présence du méta-trait METONYMIE qui marque ces emplois dans le dictionnaire.

D'une manière plus générale, on remarquera la présence de traits polysémiques comme *plat, cuisine, souper, communion*. La polysémie de ces traits provient souvent de la consécration par l'usage d'une métonymie construite dans le contexte du noyau sémantique de cette classe (le repas). Par exemple, c'est dans le cadre du repas que l'assiette (*plat*) peut en venir à désigner son contenu alimentaire. On ne sera donc pas surpris de constater que les divers sous sens de ces mots polysémiques appartiennent bien chacun de plein droit au paradigme sémantique du repas. Ceci nous montre encore à quel point la langue est la partie discursive d'un corpus cohérent : l'univers représentationnel du sujet, qui se construit dans les pratiques.

Avec les deux dernières classes, nous abordons les aspects périphériques de la représentation, ou du "sens commun" de *manger*. L'interprétation est beaucoup plus délicate, car ces parties n'appartiennent pas au "noyau central" de la représentation<sup>4</sup>. D'abord elles sont moins stables statistiquement (des variantes jouant sur les paramètres techniques de l'analyse en modifient le découpage). Il faut rester très prudent, d'autant que les sorties informatiques sous forme de texte sont un matériau très projectif quand il est ambigu. Ensuite, ces classes contiennent un mélange de paradigmes locaux, qui ne sont ici liés entre eux que par le lien commun qu'ils ont au "noyau central".

### **1.2.5 VIVRE (problématique existentielle)**

On récupère ici des extensions du "manger" à la *vie de relation* en général. Le paradigme d'incorporation sert de modèle à une grande quantité d'actes de la vie quotidienne, parmi lesquels les plus importants sont l'apprentissage, le jugement, et la régulation des relations en groupe. On peut distinguer deux actualisations actives dans la construction de cette classe. La première est simplement le principe d'incorporation (l'objet consommé s'intègre à l'être) que l'on pourrait appeler paradigme d'assimilation (en ayant soin de n'entendre par là que la

généralisation de l'incorporation à des traits ou valeurs non comestibles). Le champ d'application le plus saillant de ce paradigme est l'assimilation de la connaissance. La seconde actualisation lie, assimile, le bon à l'utile. Elle est fortement chargée d'affects hédoniques, actifs dans les mécanismes d'apprentissage, et son champ d'application se trouve dans le mécanisme de jugement et de choix. Notons que Chiva (1979, 1985) avait déjà émis l'hypothèse que le jugement moral pourrait prendre son origine dans le jugement gustatif ; comme le suggère le fait que les mimiques faciales élicitées par le sucré et l'amer correspondent à celles qui accompagnent les jugements moraux positifs ou négatifs.

La vie de relation étant par construction un tout complexe dans lequel interagissent tous les niveaux, il est extrêmement difficile de démêler dans le corpus discursif, comme dans le corpus praxéo-discursif qui le sous-tend, ces différentes articulations (connaissance, jugement, décision). C'est précisément la façon de les assembler qui constitue un ensemble de règles de vie ; et c'est bien cet ensemble complexe qui sous-tend la construction de cette classe rassemblant la connaissance et l'éthique. La généralité de la problématique existentielle qui sous-tend cette classe me fait prédire qu'on retrouvera cette problématique sous une forme ou sous une autre dans la plupart des analyses de représentations de sujets vitaux à travers des corpus issus de dictionnaire.

Les traits les plus saillants de la classe ( $\chi^2 > 20$ ) sont :

*connaître., bon+, sentir., aim+, agréable+, emploi+, goût+, posséd+, vivre., est, je, vie, éducation, appréci+, ne, idée+, joui<, femme+, esprit+, juge<, il, beau+, âme+, amer, apprendre., suivre., moral+, riche+, social+, vit, pas, caracté+, que, ressentir., savoir, société+, cet, charmant+, corromp+, accueil, dieu+, exquis+, coeur, ouvrage+, valeur<, sentiment+, c'est, vous, assimil+, adopt+, passion<, affect+, ai, musique, sens, ador+, habit+, intéré<, nature, instinct+, homme+, qualité+, travail+, verbe, inf+, fill<, ce, humain+, recevoir., enfant+, me, chose+.*

D'une manière générale, les  $\chi^2$ , plus faibles, des traits principaux montrent que la classe est moins typée que les précédentes ; c'est en partie dû à la taille importante de la classe, mais ce facteur n'explique pas tout.

Notez le caractère général, souvent abstrait, des traits (*connaître., sentir., emploi+, vie, éducation, idée+, esprit+, apprendre. moral+, social+, savoir, société+, dieu+, ouvrage+, valeur<, sentiment+, sens, habit+, intéré<, nature, instinct+, homme+, qualité+, travail+, verbe, humain+, recevoir., chose+, monde, mot+, absolu+, langage, formule+, abstrait+, savoir., pouvoir*) mais

également leur connotation positive (*bon*, *aim+*, *agréable+*, *appréci+*, *joui<*, *beau+*, *charmant+*, *exquis+*, *ador*, *bonheur*, *plaisant<*, *bien+*, *plaisir<*, *avantage+*, *mieux...*)

Les deux domaines (connaissance et éthique) qui ont été évoqués apparaissent assez clairement. La connaissance apparaît essentiellement avec la connotation d'apprentissage : *éducation*, *apprendre.*, *assimil+*, *adopt+*, *recevoir.*, *enfant+*, *étud+*, *maîtr+*, *exerc+...*

Les aspects éthiques sont centrés sur la vie en société : *moral+*, *social+*, *société+*, *dieu+*, *travail+*, *nous*, *pays...* On notera ici la présence de la *femme*, de *l'enfant* et de *l'homme*, qui apparaissent dans des contextes généraux d'éducation, d'espèce, et non pas avec la connotation sociale qui, elle, est plus caractéristique de la classe REPAS, comme le montre une analyse détaillée de l'usage de la classe.

Le lien organisateur semble celui de la vie des êtres humains : *vivre.*, *est*, *je*, *vie*, *femme+*, *société+*, *nature*, *instinct+*, *homme+*, *fill<*, *humain+*, *recevoir.*, *enfant+*, *monde*, *exist+...* Il s'agit d'être-au-monde, et l'on conçoit qu'il s'agit d'une problématique existentielle à la fois vaste, complexe, et difficile à résumer. Elle est notamment liée à l'aspect socialisant de l'alimentation dans sa fonction de partage régulateur (Fischler, 1990, p. 372). Qu'il suffise de retenir ici que les noyaux périphériques mettent le manger en connexion avec une problématique existentielle plus vaste, qui nous rappelle "qu'il faut manger pour vivre et non pas vivre pour manger" (Molière) : manger n'est qu'une pratique locale qui doit être resituée dans le cadre plus large de la vie de relation.

### **1.2.6 REMPLIR**

Cette classe est, comme la précédente, périphérique. On ne s'étendra pas longuement ici sur son commentaire. Disons simplement qu'elle tourne autour de la notion de réplétion et de dépense. La réplétion est un facteur critique dans l'acte alimentaire, et ce d'abord parce qu'elle apporte la satisfaction, avec des connotations positives, qui ont été bien développées par les psychanalystes (on consultera avec profit à ce propos Anzieu, 1974). Ensuite, parce qu'elle a des fonctions physiologiques largement décrites par les nutritionnistes, et les neurophysiologistes. On ne s'étonnera donc pas de retrouver ce noyau de sens dans l'analyse du Robert.

La classe avec ses deux aspects de réplétion et de dépense, pose une problématique de gestion des flux, qui oppose le remplissage (*remplir*,

*approvisionner, emplir, imbiber, plein, fortifier, imprégner...*), et l'acte inverse, dépense, consommation et destruction. Tout se passe autour du point d'équilibre, signalé par une saturation et un sentiment de plénitude. Je l'interprète comme une généralisation abstraite de la sensation de réplétion alimentaire, qui marque le début d'un arrêt, puis d'une inversion du processus de prise alimentaire.

On comprend que ce noyau puisse engendrer des problématiques morales, dans la mesure où il débouche sur la construction de règles du type "pas trop". Il fixe une limite, un équilibre, dont le dépassement peut être destructeur de l'enveloppe de l'être, par *défaut*, ou par *excès*. Il est même probable que l'on a ici le germe de la notion morale d'*excès* :

" (...) le "corps" même de la notion d'excès se nourrit en priorité des images liées aux consommations alimentaires pour être parlante, persuasive : de façon implicite tout excès a à voir avec le plaisir du manger et du boire, dont la placide obscénité enfantine menace l'exercice vigile de la conscience" (Nahoum-Grappe, 1991)

On retrouve ici un archétype ancien, celui de la mesure, du juste milieu, qui évoque aussi la problématique que Freud avait associée au stade anal.

Voici les traits typiques :

***rempl+*, *épui+*, *encombr+*, *ronge+*, *sature+*, *consum+*, *détruire.*, *approvisionn+*, *sujet+*, *absorb+*, *empli+*, *imbibe+*, *vide+*, *feu*, *plein+*, *farc+*, *bourr+*, *chose+*, *complet+*, *abreuv+*, *fatigue+*, *fortifi+*, *imprégn+*, *voiture+*, *dévorer*, *consomm+*, *n*, *placer.*, *garni+*, *occup+*, *eau*, *entier+*, *gonfl+*, *rendre.*, *fonction+*, *brûl+*, *pron+*, *jusque*, *gorge+*, *discours*, *temps*, *user.*, *article+*, *accompli+*, *trop*, *air*, *abstrait+*, *ses*, *marchand+*, *ADJ*, *soul+*, *pass+*, *dissip+*, *compl+*, *son*, *force+*, *sang+*, *truffe+*, *dépenser*, *vill+*, *entam+*, *quantité+*, *se*, *CONTR*, *fer*, *dévor+*, *muni+*, *dépense+*, *perdre.*, *tout*, *tête*, *désign+*, *coeur*, *fort+*, *livre+*, *marche+*, *racine+*, *ravitaille+*, *charge+*, *bouff+*, *anal*.**

La présence de traits opposés, d'ailleurs marquée par le méta-trait CONTR. est caractéristique d'une telle *problématique*.

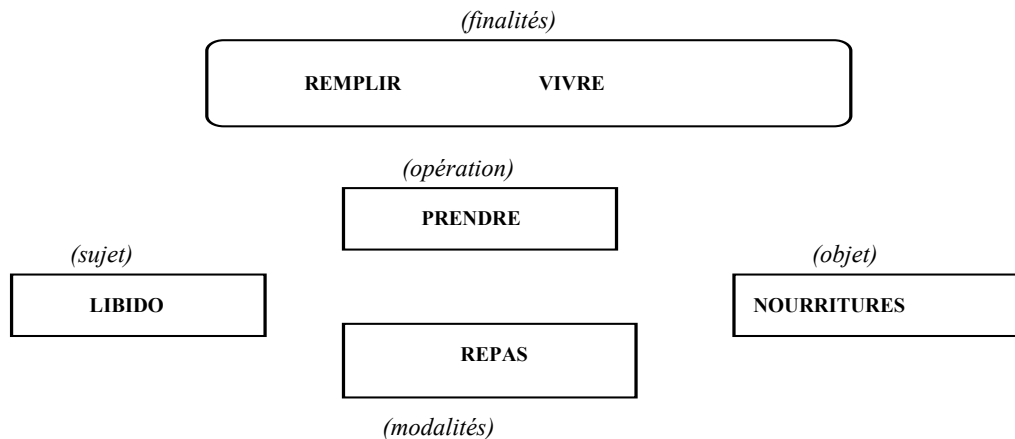
### 1.2.7 Un schéma pragmatique

Pour conclure, on peut remarquer que l'analyse complète livre une structure proche du modèle linguistique sujet/verbe/objet/complément. Celui-ci s'articule d'abord dans un cadre pragmatique en reliant sujet (ou, plus exactement, pulsion), objet, opération, et modalités. Nous avons ici un schéma "agent-action-patient" qui serait

"une représentation implicite, présente dans la mémoire sémantique de tout locuteur, de la causalité humaine - c'est-à-dire de la causalité dans laquelle c'est un être humain qui joue le rôle de cause" (Le Ny, 1989, cité par Cordier, 1991).

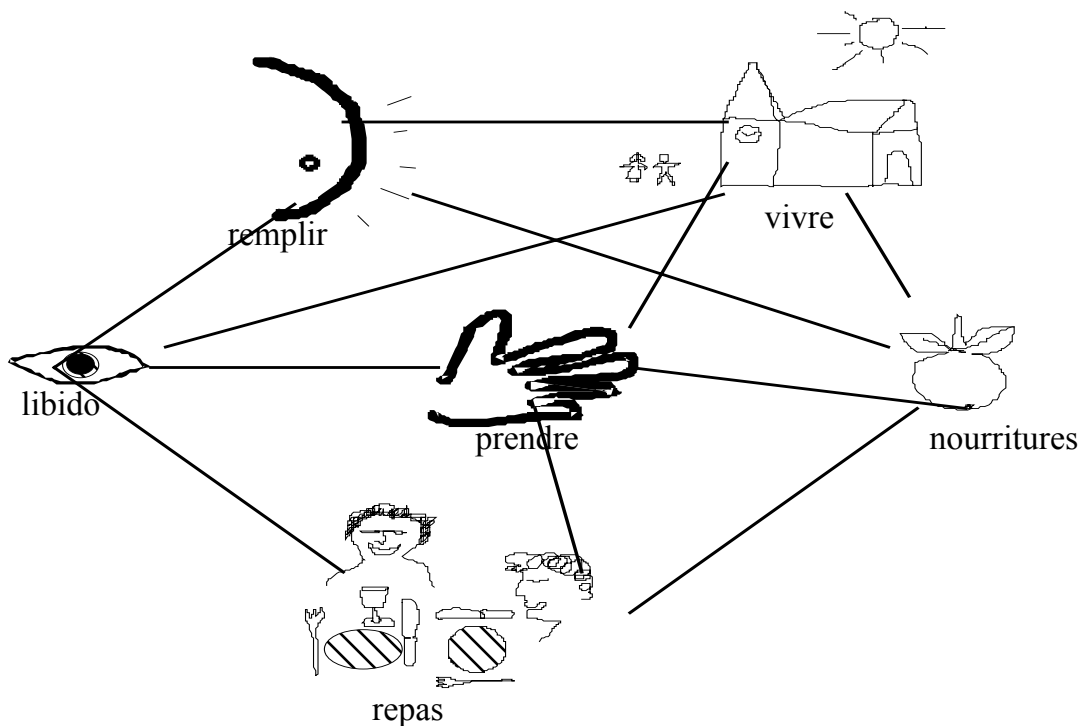
Mais l'analyse nous livre en plus des finalités (REEMPLIR, VIVRE).

Fig. 4 : "Manger" d'après le Robert : les classes replacées dans un cadre conceptuel



Comme les classes correspondent à des éléments vécus plus qu'à des mots, nous pouvons imaginer que, subjectivement, l'articulation entre ces éléments du point de vue du sujet culturel ressemble à quelque chose comme :

Fig. 5 : "Manger" d'après le Grand Robert : les classes en vue subjective





Cette première application de la méthode à un corpus d'associations libres montre qu'elle permet de faire émerger, par analogie et contraste, les noyaux de base de la représentation. Ceux-ci sont assez clairs, et leur articulation fait sens.

Car le résultat le plus intéressant et novateur sur le plan technique est que l'on a obtenu la construction d'un sens. En effet, le sens n'est, finalement, que l'association entre des objets, des noyaux de représentation. Et l'analyse nous livre, "brut de décoffrage", que *manger*, c'est l'articulation entre les noyaux APPETIT/ PRENDRE/ NOURRITURE/ REPAS/ REMPLIR/ VIVRE. Que ceux-ci s'articulent dans un cadre pragmatique (reliant sujet -ou, plus exactement, pulsion-, objet, opération, modalités et buts), qui est le décalque de l'acte physiologique de manger nous permet de dire qu'il s'agit bien là de "bon sens".

Ceci peut paraître trivial, si nous le savions déjà. Tout l'intérêt réside dans le fait que cette articulation n'est pas une interprétation, ni une définition, mais dans une grande mesure (nous reviendrons sur ce bémol) le *résultat empirique* d'une méthode aveugle, et sourde au sens, qui est rigoureuse et reproductible. Que l'on retombe ici sur du bon sens est encourageant, et nous incite à penser que, appliquée à des corpus dont le contenu n'est pas connu au préalable, la méthode nous livrera les grandes dimensions du sens sous forme d'une articulation des noyaux de base.

Notons par ailleurs que les résultats ne sont absolument pas triviaux pour les spécialistes du domaine. On retrouve notamment des éléments intéressants sur le plan de la recherche. Ainsi la présence d'agressivité dans le noyau "prendre" confirme des intuitions anciennes de Freud sur le processus "d'incorporation" de la phase orale qui n'avaient pas jusqu'ici reçu beaucoup de confirmations expérimentales :

"En fait trois significations sont bien présentes dans l'incorporation : se donner un plaisir en faisant pénétrer un objet en soi ; détruire cet objet ; s'assimiler les qualités de cet objet en le conservant en dedans de soi. C'est ce dernier aspect qui fait de l'incorporation la matrice de l'introjection et de l'identification. "(Laplanche et Pontalis, dans l'article "incorporation" du Vocabulaire de Psychanalyse, 1967, 1990).

Enfin, dernier aspect non négligeable, une analyse des associations libres sur l'inducteur "manger" auprès d'une population de 2000 adultes représentatifs de la population française a fourni des noyaux de sens pratiquement identiques. La seule différence majeure résulte de la quasi disparition du noyau PRENDRE, ce qui s'explique probablement par le fait qu'il s'agissait d'un questionnaire en face à face. Cet aspect de la représentation, évident et trivial, aurait selon moi

disparu car la situation d'interlocution ferait jouer les maximes de conversation de Grice (1975). Le trait PRENDRE, évident, serait implicite dans la conversation. Ceci incline à penser que la méthode permet bien de faire parler l'Esprit social. Je renvoie le lecteur à (Beaudouin et Lahlou, 1993, Lahlou, 1994, 1995b) pour une analyse plus détaillée de ces questions.

## **2. Du sens partagé des objets**

Après cette première plongée dans les résultats, je souhaite discuter deux points. Le premier, qui fait l'objet de cette section, est que le "sens" qui ressort de cette analyse sur la langue ne provient pas de la langue elle-même, mais de ce que celle-ci contient une image du monde vécu. Le second (section suivante) est un essai de théorisation du processus d'interprétation.

Je ne suis pas linguiste, mon objet d'étude, c'est les représentations sociales (Moscovici, 1961). Pour les étudier, j'ai choisi de me placer, de manière radicale, dans le monde subjectif des observateurs. De là, délibérément placé du point de vue du sujet, je considère tous les objets, matériels ou non, comme des représentations. Il n'y a alors pas d'objets en soi, il n'y a que des objets tels que se les représente l'observateur.

Ces éléments peuvent être des objets qui ont une existence propre. Par exemple, un pot de café est composé de l'objet "pot" et de l'objet "café". En dernière analyse, les objets sont composés de traits sémiotiques qui correspondent à des éléments (perceptions, émotions, sensations...) du monde vécu par le sujet. Par exemple, le café est liquide, noir, chaud, excitant, buvable... Dans l'univers subjectif de l'individu, une représentation d'objet serait finalement constituée d'un arrangement de traits sémiotiques, des traces "élémentaires" d'états vécus stockés en mémoire<sup>5</sup>, que j'appellerai ici faute de mieux "éléments de base". La représentation est multimodale, elle a des dimensions sensorielles, motrices, émotionnelles, cognitives, linguistiques... Nous ne n'avons étudié ici que les aspects linguistiques, ce qui est très réducteur mais bien pratique.

Le contour exact des éléments de base ne nous est pas connu a priori. La représentation est, pour paraphraser une ancienne métaphore, un réseau infini dont le centre est partout et la circonférence nulle part ; puisque chaque trait ne prend sens que dans un contexte plus général, celui de la connaissance globale du monde, de même qu'un mot ne prend sens que par sa situation dans

l'ensemble de la langue<sup>6</sup>. Il n'y aurait donc pas d'"éléments", mais seulement des portions, plus ou moins étendues, de ce réseau. Néanmoins, comme le montre l'analyse qui précède, il semble qu'il soit possible localement de distinguer des noyaux relativement stables, qui font "sens" pour nous, à des niveaux plus ou moins fins de la maille sémiotique<sup>7</sup>. J'ai considéré qu'ils représentent les éléments que je recherche.

Comment, pratiquement, repérer ces éléments ? Par des méthodes distributionnelles. Pour dire vite, une représentation de l'objet serait constituée de l'association des principales perceptions formant les contextes dans lesquels cet objet "apparaît". Ce qu'il y a de constant dans ces apparitions, par contraste avec ce qui est variable, fonde, selon moi, l'existence de l'objet en tant que tel par rapport à ce qui est "le reste du monde" du point de vue de l'observateur. Les éléments de base, ou leurs combinaisons, se produisent donc, "émergent" par la constance du "même" par contraste avec des "différents".

Etendons maintenant ce principe au niveau social. De même qu'au niveau d'un observateur individuel ce qui définit l'objet est le contraste entre le même, qui se reconnaît, et le différent qui sert de contexte, de même dans une culture donnée les objets "réels", c'est-à-dire ceux qui existent pour tous, seront définis précisément par leur constance dans les différentes vues, inter-individuellement. Et les objets communs à tous forment notre monde du sens commun.

Or, derrière toute société digne de ce nom, il existe un tel monde commun, nécessairement. Car, pour que la vie soit possible en société, il faut un minimum de recouvrement des "entours" sémantiques individuels, qui permette des interactions pragmatiques. C'est la nécessité sociale du "concept" (Durkheim, 1912). Pour une "chose" particulière, c'est cette intersection floue, et multimodale, à laquelle renverrait le nom de la chose. C'est cette intersection floue, à laquelle réfère le nom, que je cherche à expliciter. Je postule donc qu'elle est constituée d'une association relativement solide, partagée, permanente, entre éléments de base.

Whorf (1927) a fourni une approche de cette perspective dans la langue avec la notion de "connexion". Il oppose les "associations", traces mémorielles idiosyncrasiques, issues de l'histoire personnelle du sujet, à celles qui sont partagées par la communauté sociale du locuteur, et constituent les véritables "connexions". La notion de connexion permet de faire le lien logique entre la vision encyclopédique (le langage comme réseau continu et insécable de mots réciproquement définis), et le postulat de Bloomfield<sup>8</sup>, qui considère que le sens des mots est partagé par tous, au moins jusqu'à un certain point. La distinction

entre *associations* et *connexions* recouvre celle faite par Durkheim ou Halbwachs entre les *images individuelles* et les *représentations collectives*. Il est clair que ce que nous cherchons ici, ce sont les "connexions".

Le sens, certes, ne saurait se trouver dans la langue elle-même, il est dans le rapport au monde des sujets : pas de sémantique sans pragmatique. Ceci pose un problème technique difficile, car les méthodes actuelles ne permettent pas d'investigation dans le monde vécu des sujets, ce que Flament (1989) appelle "l'univers praxéo-discursif", car nous savons mal décrire les perceptions "internes" au sujet.

Heureusement, les associations des objets du monde vécu se retrouvent dans le monde des idées, et celui-ci s'exprime par la langue. Dès 1660, Spinoza décrit clairement la chaîne qui, fixant dans la mémoire les associations sensorielles, *propage* dans le monde idéal une image du monde vécu, faisant que la structure du monde des idées est analogue à celle du monde des choses.

L'ordre et l'enchaînement des idées est le même que l'ordre et l'enchaînement des choses (Ethique, livre II, théorème VII)

car :

Comme les pensées et les idées des choses s'ordonnent et s'enchaînent dans l'Âme, de même exactement les affections du Corps et les images des choses s'ordonnent et s'enchaînent dans le corps. (Ethique, livre V, théorème I).

car :

Si le corps humain s'est trouvé une fois affecté simultanément par deux ou plusieurs corps, dès que, dans la suite, l'âme s'imaginera l'un quelconque d'entre eux, elle se souviendra aussitôt des autres. (Ethique, livre II, théorème XVIII)

Parce que l'âme enregistre, sous forme d'associations d'idées, les associations entre choses rencontrées par le corps sentant dans le monde des choses, le réseau des idées sera analogue au réseau des choses perçues.

La scholie du théorème XVIII, livre II, nous indique la voie à suivre pour repérer les connexions : c'est celle de *l'association libre*:

"Par là nous comprenons immédiatement ce qu'est la mémoire. Ce n'est pas en effet autre chose qu'un certain enchaînement d'idées qui enveloppent en elles la nature de choses existantes en dehors du corps humain, enchaînement qui se produit dans l'âme, suivant l'ordre et l'enchaînement des modifications du corps humain. (...)

Ainsi par exemple, de la pensée du mot *pomus* un Romain tombera immédiatement dans l'idée d'un fruit qui n'a aucune ressemblance avec ce son articulé, ni avec lui rien de commun, si ce n'est que le corps de cet homme a souvent été affecté par ces deux choses à la fois ; c'est-à-dire que ce même homme a souvent entendu le mot *pomus* en même

temps qu'il voyait le fruit. Ainsi chacun tombera d'une pensée dans une autre suivant que, dans le corps de chacun, l'habitude aura ordonné les images des choses. Car un soldat, par exemple, voit dans le sable les traces d'un cheval, aussitôt, de la pensée du cheval il tombera dans celle du cavalier, de là dans la pensée de la guerre, etc. Le paysan, au contraire, tombera de la pensée du cheval, dans celle de la charrue, du champ, etc. Ainsi chacun, suivant qu'il aura contracté l'habitude d'associer et d'enchaîner les images de telle ou telle façon, tombera d'une pensée donnée dans une autre. (Spinoza, Ethique, livre II, scholie du théorème XVIII).

Par de tels protocoles associatifs, qui nous amènent à découvrir, dans le territoire de l'esprit, les chemins tracés par la mémoire, et dont les travaux de Freud ou Jung, notamment, ont montré la fécondité, on peut chercher à recueillir des corpus de vues locales des objets qui nous intéressent. On s'adresse alors non pas à l'esprit individuel, mais à l'Esprit de la culture, au sens où en parle Bateson (1979). Cet esprit du corps social peut s'exprimer à travers les locuteurs individuels ou collectifs. De ces vues locales exprimées par le corps social, nous chercherons, par des méthodes distributionnelles, à faire émerger par commutation les traits qui sont en connexion fréquemment attestée.

Les protocoles considéreront des vues locales exprimées en langue naturelle, ce qui est une limite importante par rapport au corpus praxéo-discursif, dont nous n'obtiendrons que des projections déformées. Consolons-nous en rappelant que la langue, instrument de communication privilégié, est précisément conçue pour exprimer les corpus praxéo-discursif du monde vécu. L'objet sera donc dans ce cadre d'abord un mot ou une expression.

C'est dans la langue que je suis allé chercher les objets du monde, en considérant que la langue est une mémoire sociale, et que dans son réseau sémantique elle a sédimenté les visions du monde produites par la culture.

Voilà donc comment je suis arrivé dans un terrain exploré depuis longtemps, où l'on cherche à représenter le sens d'un mot par une relation d'association entre d'autres objets de la langue.

Il est maintenant temps de dire quelques mots de la théorie d'interprétation qui a été implicitement utilisée pour obtenir les "noyaux de sens", ou éléments de la représentation.

### 3. L'art de l'interprétation

Pour déterminer de tels noyaux de sens, dont une combinaison forme la représentation, on repère par des méthodes statistiques ce qu'il y a de commun dans différentes "vues" de la chose, énoncées par une ou plusieurs sources de connaissances imprégnées de la culture dans laquelle existe l'objet (une population, un dictionnaire...). L'espoir de faire émerger les éléments de base du sens (de la représentation de l'objet) par une telle méthode vient de deux idées. D'une part celle, énoncée plus haut, qu'il existe effectivement *quelque chose de commun* dans les différentes vues de l'objet, quelque chose de stable qui fonde l'existence de l'objet en tant qu'il est relativement indépendant des vues qui peuvent en être obtenues. D'autre part que c'est par comparaison entre ces vues que l'on fera émerger le même par contraste avec le différent ; c'est-à-dire, en termes distributionnels, ce qui se répète dans un contexte variable.

Pratiquement, on a récupéré, sous forme de texte, un grand nombre de vues locales énoncées à propos de l'objet. Ces textes sont rassemblés dans un unique corpus, qui est ensuite scindé en énoncés, qui sont considérés comme des associations de traits. Ce sont ici les mots, les locutions, ou leurs racines, qui instancient localement les éléments que nous cherchons.

Les énoncés sont classés, par analogie et contraste, sur la base de leur contenu lexical. On obtient des classes qui contiennent des énoncés. Les énoncés classés sont similaires entre eux dans une même classe, et aussi différents que possible des énoncés des autres classes. Chaque classe possède des traits typiques lexicaux, ceux qui reviennent fréquemment dans les énoncés. Si ces traits renvoient tous à un même noyau de sens, celui-ci peut être considéré comme un des éléments constitutifs du sens du mot.

On a alors trouvé, par abduction, les éléments que nous cherchions, à travers l'interprétation des classes. Pourquoi par abduction ? Examinons cette phase obscure et rarement décrite qu'est l'opération d'interprétation, où l'opérateur introduit subrepticement sa connaissance du monde.

La méthode de classification descendante utilisée ici construit d'un même mouvement les traits caractéristiques et les classes, car la probabilité que le trait "a" (en l'occurrence, une forme graphique, le lexème) puisse être un trait pertinent provient précisément de ce que "a" est observé sur un certain nombre d'énoncés, mais pas sur tous. En d'autres termes, émergent comme traits classificatoires les traits *discriminants*, ceux qui sont effectivement susceptibles d'être productifs de classifications intéressantes. Ceux-ci permettent de

caractériser un nombre suffisant d'expressions à classer, ils sont donc efficaces. Les traits trop répandus (non discriminants) ou les traits trop rares (hapax, traits aberrants) ne sont pas intéressants. C'est là une résolution empirique, idoine, du problème du choix des traits pertinents. Les traits non discriminants sont rendus inopérants par leur répartition homogène dans les énoncés, ils constituent le "fond". Certains traits qui pourraient être discriminants mais sont présents en trop petit nombre n'ont pas une fréquence assez élevée pour influencer l'analyse : ils sont également rendus inopérants par la classification descendante. Mais nous n'avons pour le moment que procédé à un classement, pas encore à une interprétation.

L'opération d'interprétation des classes réside dans la décision de considérer les traits typiques d'une classe comme représentant tous une "idée" unique, à laquelle et par laquelle cette classe est identifiée. Autrement dit, d'abord l'analyste remarque que la classe est caractérisée par un ensemble ( $a_j$ ) de traits typiques. Il *suppose* que la classe correspond bien à une idée unique, c'est-à-dire que ces ( $a_j$ ) sont des avatars d'un unique paradigme, à déterminer, qui serait un noyau de sens, l'idée. En examinant les traits ( $a_j$ ) il cherche à inférer la nature de ce paradigme générateur. Par exemple, "*viande*" lui semble renvoyer à NOURRITURE. Il s'agit là d'une induction.

L'analyste pose alors que la classe réfère au paradigme "A" (ex : NOURRITURE"). Ce faisant, l'analyste décide que "A" est effectivement un noyau de sens acceptable comme signifié d'une classe paradigmatique. L'analyste vérifie alors que tous les traits ( $a_j$ ) sont effectivement acceptables comme avatars lexicaux de A, aux artefacts de méthode près. Si c'est le cas, l'analyste reconnaît que la classe signifie bien "A", et il l'interprète comme telle. Ceci se fait progressivement, en essayant éventuellement plusieurs "A" candidats.

Par exemple, l'induction que "*viande*" et "*poisson*" renverraient à une classe ANIMAUX (signifié potentiel pour l'interprétation de la classe) se voit invalidée par la présence de "*légume*" et "*pain*". Réciproquement, il semblerait peu intéressant d'interpréter la classe en question par le paradigme OBJETS MATERIELS, qui est trop large. L'analyste retient finalement un signifié tel que son paradigme soit quelque chose comme le plus petit commun paradigme (PPCP) rendant compte des traits observés.

Il s'agit donc d'une opération d'abduction, comme nous allons le voir en passant l'opération "au ralenti".

Lorsque l'analyste compare plusieurs expressions, il constate que :

(1) - plusieurs expressions ( $X_i$ ) possèdent le trait " $a_i$ "

Il en induit :

(2) - il existe *probablement* une "bonne classe"  $C(A)$ , dont " $a_i$ " est un trait caractéristique

soit : il est probable que (Il existe  $C(A)$  tel que, si  $X_i$  contient  $a_i$ , alors  $X_i$  appartient à  $C(A)$ )

*L'analyste examine alors la plausibilité de  $C(A)$  en tant que classe, c'est-à-dire qu'il cherche dans sa connaissance du monde le  $A$  en question. S'il trouve un tel  $A$  acceptable et pertinent (PPCP), alors :*

Il abduit de (1) et (2) :

(3) - alors  $X_i$  fait partie de la classe  $C(A)$ .

*Dans cette opération d'abduction, ce qui caractérise une "bonne" classe, c'est qu'elle renvoie à "quelque chose" identifié par l'analyste, dans notre cas un signifié  $A$ , le PPCP, dont les ( $a_i$ ) seraient des expressions, ou signes.*

Dans la pratique, le logiciel, sourd au sens, constitue les classes sur la base de tests statistiques (partition optimale et stable). Ces classes peuvent être "bonnes" (interprétables : on reconnaît " $A$ ") ou pas "bonnes" (non interprétables : on ne reconnaît rien). L'analyste "vérifie" donc que ces classes correspondent effectivement à "quelque chose" d'interprétable, c'est-à-dire un objet du monde ou un concept connu. Mais cette opération de vérification est subreptice. Et la manière dont sont présentés les "résultats de l'analyse" la passe sous silence, ou plutôt, fait croire qu'elle est postérieure à la constitution des classes.

Car, une fois l'analyse terminée, l'analyste présente toujours les résultats de telle manière que, effectivement, les ( $a_i$ ) sont "naturellement" interprétables comme "signes" de " $A$ ". Donc, il ne présente que des "bonnes" classes dans le contexte culturel de son public. Ceci occulte la façon dont le travail réel s'est fait. Puisque, en fait, souvent l'analyste a obtenu une, ou plusieurs classifications préalables, dans lesquelles certaines classes "ne faisaient pas sens" : certains des ( $a_i$ ) n'étaient pas admissibles comme représentants de " $A$ ". C'est-à-dire, il n'a pas trouvé de bon " $A$ " candidat à être le signifié dont tous les " $a_i$ " seraient des représentants lexicaux. Il a alors, utilisant les facilités techniques de son dispositif de classification (ici, les paramètres du logiciel) réalisé une autre classification, modifiant par essais successifs la partition pour obtenir le résultat final, qui, lui, est acceptable. Il m'arrive souvent, personnellement, de réaliser plusieurs dizaines d'essais pour obtenir une "bonne" analyse. Un analyste



prudent se doit d'ailleurs de réaliser systématiquement plusieurs essais en variant les paramètres, pour vérifier que les résultats obtenus sont solides (stables, c'est-à-dire rebelles à ses manipulations). C'est au cours de ces analyses successives que les contours de l'objet se précisent pour lui. Et, naturellement, il choisit de présenter "le meilleur", c'est-à-dire celui qui correspond le mieux à sa conception du monde.

La boîte noire qu'est l'analyste utilise sa connaissance du monde pour reconnaître l'objet "A" à travers ses traits manifestes. L'interprétation suppose donc, implicitement, que l'analyste connaisse déjà plus ou moins les objets qu'il cherche, et soit capable de les identifier à partir des combinaisons que son matériau lui présente comme pertinentes.

La seule différence entre une analyse appuyée sur du matériel empirique et la réflexion en chambre consiste donc en ce que la technique (ici, le logiciel ALCESTE) sélectionne certaines combinaisons particulières de traits candidates à être les "bons" noyaux de sens ("signifiés"), suivant un certain principe légitime (ici : la redondance dans les différents énoncés). Ce point crucial est rarement explicité par les praticiens. Or c'est la connaissance préalable de l'objet, et celle des limites de la technique utilisée, qui fait la différence entre les bonnes et les mauvaises analyses.

Résumons les opérations, dans ce que j'appelle pompeusement une théorie de l'interprétation :

1) Rassembler un corpus d'énoncés "à propos de" l'objet qui nous intéresse.

2) Classer ensemble les énoncés qui ont des traits linguistiques analogues

3) Interpréter les classes en termes sémantiques. Soit :

a) Repérer les combinaisons de traits lexicaux typiques de chaque classe ;

b) En induire, dans l'espace sémantique, des noyaux de sens, constitués de combinaisons de traits sémantiques correspondants ;

c) Formuler des lois d'expression de cette combinaison sémantique dans la langue. (Pratiquement, on considérera toutes les expressions linguistiquement "bien formées" du noyau sémantique sont des productions acceptables.)

4) Vérifier que les expressions effectivement classées dans chaque classe sont interprétables comme expression de ces noyaux sémantiques sous-jacents, (règles de langue, relations sémantiques sur le monde issues non seulement des

associations attestées du corpus, mais de la connaissance du sens commun sur le monde).

Pour répondre à une question qui m'a été posée lors du séminaire de l'INALF : s'agit-il d'une théorie ("construction intellectuelle méthodique et organisée, de caractère hypothétique -au moins en certaines de ses parties- et synthétique") ou d'une méthode ("ensemble de moyens raisonnés employés pour parvenir à un but") ? Il s'agit, tel qu'exposé plus haut, d'une méthode, quand on applique à des corpus de textes la procédure que j'ai suivie. Mais je penche pour l'appeler théorie, puisqu'il s'agit ici davantage d'une description, en partie hypothétique, des actions dont je prétends qu'elles sont réalisées par les interprétateurs, chaque fois qu'ils essaient de modéliser un phénomène à partir d'une grande quantité de descriptions locales, par exemple des énoncés. Peut-être est-ce de la méthodologie ?

Je pense donc que cette théorie correspond à une démarche scientifique plus générale, appliquée dans les sciences expérimentales, et que l'on pourrait décrire de la façon suivante, en appelant "syplexe" (Lahlou, 1990) les combinaisons -ordonnées ou non-, de traits observés ou décrites :

1. On obtient, par une série d'expériences (observations...), des vues subjectives et partielles de l'objet présumé, exprimées sous la forme de syplexes de signes conventionnels (descriptions en langue, mesures...).
2. On compare ces expressions pour en tirer des régularités, c'est-à-dire des syplexes analogues (c'est-à-dire : ayant des traits communs). Lors de cette opération, on ne retient que des régularités considérées comme "pertinentes", et cela par rapport à une connaissance du monde dont on dispose préalablement<sup>9</sup>.
3. On construit des règles (ou lois) de production de ces syplexes (règles d'assemblages des signes conventionnels).
4. On vérifie que les productions de ces règles correspondent effectivement à des observables, par des expériences de validation.

Mon approche paraîtra sans doute naïve, et je ne sais pas si cette "théorie" est bonne en l'état, mais je crois qu'il y a là quelque chose à creuser. Il y a, en tous cas, à l'heure actuelle un manque cruel dans ce domaine (Lebart, 1993).

## 4. Information

A la lumière de ces considérations théoriques, voici maintenant, obtenue par les mêmes méthodes, l'analyse du mot "information" (235 définitions, 115 000 mots, 625 000 caractères, 7256 uce). Contrairement à la précédente, cette analyse porte sur un concept extrêmement flou, et dont l'interprétation est difficile à vérifier intuitivement. C'est là que la méthode risque de montrer son utilité. J'ai volontairement fourni une longue liste de mots typiques, afin que le lecteur puisse s'adonner aux joies de l'interprétation, et constater combien cette opération est discutable.

L'analyse fournit 7 classes qui correspondent à des (présomptions) de noyaux sémantiques constitutifs du concept d'*information*. J'ai appelé ces classes : CONNAISSANCE, CODAGE, MEDIA, ACQUERIR, AVISER, INSTITUER, APPRENDRE (voir Lahlou, 1995a, pour une analyse plus détaillée).

Trois classes fournissent le coeur technique du concept, et correspondent assez bien aux aspects pris en compte par la théorie de l'information de Shannon (1948). Il s'agit des classes CONNAISSANCE, CODAGE et MEDIA. Ces trois classes décrivent la forme de l'information : celle-ci est un contenu idéal, une représentation, la CONNAISSANCE (pensée, sentiment, etc.), présenté sur un support (courrier, publication, messenger...), le MEDIA, en utilisant un certain système de CODAGE (code, informatique, signal...).

### 4.1 MEDIA.

Voici une liste des traits typiques :

*journa+l, lettre+, courrier+, renseignement+, presse+, reclame+, annonce+, feuille+, depeche+, escrit+, affiche+, prospectus, adress+er, imprime+, envo+yer, pli+, publicite+, expres+, missive, pli+er, bulletin+, livre+, papier+, correspondance+, depli+er, page+, spectacle+, note+, cahier+, service+, televise+, reportage+, communiqu+er, billet+, carnet+, reception+, vente+, periodique+, publication+, agent+, publicitaire, revue+, message<, agence+, radio+, edition+, expedition+, faire-part, actualite+, distribution+, magazine, nouvel+, titre+, destinataire+, voiture+, chronique+, recueil+, commerce+, recommand+er, achet+er, article+, somme+, depos+er, imprim+er, piece+, texte+, communicat<, poste+, diplomati+que, fiche+, traite+, mise+, television+, fourn+ir, pa+yer, reponse+, carte+, preface+, marque+, mur+, fete+, illustr+er, annonc+er, copie+, voyage+, signataire+, avis, dossier+,*

*impression+*, *maison+*, *gros+*, *film+*, *remettre.*, *bre+f*, *exemplaire+*, *large+*, *accuse+*, *campagne+*, *colonne+*, *marchandise+*, *tissu+*, *etranger+*, *emission+*, *commercia+l*, *verba+l*, *explication+*, *lire+*, *roman+*, *publi+er*, *demand+er*, *ambassad+eur*, *reference+*,

La notion de support/véhicule apparaît sous de nombreux avatars classiques du média : les supports papier (*journal*, *lettre*, *courrier*+...), oraux (*renseignement*, *annonce*, *avis*...), mais aussi la *radio*, la *télévision*, le *film*, et même des supports auxquels on penserait moins en première intention : le *mur*, l'*ambassadeur*, l'*agence*, le *spectacle*...

Signe peut-être de la vétusté du dictionnaire utilisé, qui est pourtant utilisé pour cette analyse dans son édition 1994 sur CD-ROM, les média informatisés n'apparaissent pas ici. On verra plus loin que c'est sans doute parce que toute " l'informatique " apparaît regroupée avec la fonction de codage-décodage dans la classe suivante : CODAGE. Celle-ci rassemble " dans le même sac " tout ce qui relève de la théorie de Shannon : le code, le signal, l'émetteur, le récepteur, le message ... Ceci n'est pas si bête puisqu'il s'agit finalement d'un *système de codification* dans lequel le codage et ses instruments sont solidaires en ce qu'ils constituent les maillons d'une chaîne de codage-décodage.

## 4.2 CODAGE

Voici maintenant les traits typiques de la classe CODAGE. Je recommande une lecture rapide, impressionniste, de ces traits. Les résultats ne doivent pas être considérés comme un modèle achevé, mais comme des indices, des indications pour construire un tel modèle. Les résultats ici présentés sont plus intéressants par ce qu'ils suggèrent que par ce qu'ils sont. Comme je l'ai dit plus haut, il serait malhonnête de faire croire qu'entre les mains d'un analyste un morceau de listing, fut-il issu d'ALCESTE, est beaucoup plus qu'un *prétexte* à l'interprétation :

*signa+l*, *information+*, *systeme+*, *recepteur*, *cana+l*, *signe+*, *code+*, *theorie+*, *ordinateur+*, *emetteur*, *transmission+*, *support+*, *langage+*, *techn*, *message<*, *appareil+*, *indice+*, *element+*, *traitement+*, *permettre.*, *didact*, *graphi+que*, *dispositif*, *analyse+*, *chiffre+*, *communicat<*, *electrique+*, *semiotique*, *informatique*, *fonction+*, *dessin+*, *machine+*, *decodage*, *nombre+*, *unite+*, *symbole+*, *cybernetique*, *detecteur*, *inform(atique)*, *cod+*, *reproduction+*, *calculat<*, *distance+*, *produit+*, *artificiel+*, *automatique+*, *linguistique*, *utilis+er*, *tuyau+*, *mobile+*, *regulation*, *encod+*, *operation+*, *acoustique<*,

*onde+, surface+, effectu+er, logi+que, phrase+, utilisation+, destin+er, base+, tube+, decel+er, servant+, mecanisme+, quantite+, reproduire., forme+, ensemble+, discours, caractere+, eau+, effet+, trait+, locuteur, moyen+, transmettre., sc(ientifique), organique+, sonore+, constitu+er, frequ+ent, visuel+, magneti<, procede+, et, rapport+, passage+, media, sons, naturel+, entre, volontaire+, combinaison+, processus, provoqu+er, disque+, electronique+, ou, materiel+, oppos+er, partie+, programme+, plan+, aide+, correspondant+, energie+, programm+, entree+, entre-deux, conversion+, gaz, micro+, syntaxe+, taux, mouvoir., genetique, bruit+, mecanique+, math+, relation+, organi+que, suite+, notamment, terme+, notion+, technique+, selon, exterieur+, interrupteur, masse+, production+, spot, changement+, instrument+, representant+, desinform+, prelud+er, trac+er, traduire., television+, espece+, neuf+, represent+er, centre+, fonctionnement+, representation+, complexe+, mesur+er, ligne+, langue+, evolution+, navire+, sortie+, circulaire+, emission+, exposition+, produire., conventionnel+, archiv+, directiv+, enonc+er, en-vue, application+, phenomene+, lumineu+x, fer+, assur+er...*

J'ai volontairement laissé ici tous les termes. Mon analyse en terme de "CODAGE" est évidemment discrétionnaire. Je néglige par exemple les traits "taux", ou "navire" qui me semblent peu pertinents.

### 4.3 CONNAISSANCE

La troisième classe, je l'ai appelée CONNAISSANCE, bien que ce terme ne recouvre qu'imparfaitement son contenu. En voici les traits les plus typiques :

*voir., savoir+, pens+er, savoir., sentiment+, je, jeune+, apprendre., connaitre., chose+, connaissance+, nouveau+, don+, se, dieu+, c'-est, esprit+, ignorance+, qu+, on, rien, etre, il, ce-qu+, pas, ai, facilite+, vivre., memoire+, avoir, homme+, coeur+, suivi+, sentir., peu, aim+er, est, venir., nous, illustre+, bien, experience+, emplo+yer, ma, faire., relativement, souvenir+, vouloir., jug+er, beau+, effort+, plaire., jamais, beaucoup, sa, oubli+er, trouv+er, tout, nom+, gens, vous, plaisir+, rencontr+er, sens, recent+, vie+, etudi+er, avec, toujours, possi+ble, culture+, ecrivain+, intellectuel+, vue+, science+, inconnu+, amour+, fontaine+, objet+, propre+, raison+, quand, quoi, mort+, infiniti+f, comedie+, complement+, habitude+, imagination+, consci+ent, intellig+ent, perdre., sans, arriv+er, croire., notre, elle, femme+, age+, qqch, encore, cela, affecti+f, passion+, prophetie+, ame+, manque+, comment, lui,*

*mal, metier+, ignor+er, capable+, connu+, fille+, joie+, moi, plus, appliqu+er, peuple+, attribut+, foi, voltaire, guere, enfant+, dit+, gard+er, grace+, aucun+, seul+, qui, mon, me, acquis+, oisivete, rappel+er, expression+, dessous, profond+, VRAI,...*

La classe CONNAISSANCE est moins facile à interpréter que les deux précédentes, elle contient un noyau de sens plus flou. Cela se comprend dans la mesure où elle correspond sans doute à l'idée de *contenu* en général, qui transcende le média et le codage, idée finalement assez abstraite qu'il est difficile d'exprimer sans faire référence à la forme. Ce contenu réfère à des expériences subjectives de la réalité, des tableaux du monde : *savoir+(substantif), pensée, savoir (verbe), sentiment, chose, connaissance, mémoire, expérience, souvenir, ...* Le trait *voir* figure dans cette classe essentiellement au titre de son sens référentiel (de *renvoi à*), plus que de son sens perceptif.

La liste contient des traits qui insistent sur les aspects mentaux du contenu de l'information (*esprit, intellectuel, conscient, intelligent...*), ainsi que de nombreux traits abstraits et même métaphysiques, (*dieu, rien, être, vivre, science, sens, raison âme, foi, VRAI,...*), ou moraux (*bien, juger, beau, mal...*) ainsi que des connotations émotionnelles (*sentiment, amour, mort, affectif, passion, joie...*)

Cette ligne de fuite vers un horizon métaphysique, a priori un peu surprenante, n'est pas une nouveauté dans les analyses issues du dictionnaire, et nous l'avons déjà trouvée dans d'autres analyses (cf. Supra "VIVRE"). Elle caractérise les aspects finaux (au sens de buts) de toutes les actions humaines socialisées dans une vie de relation, lesquelles débouchent en dernière analyse, systématiquement, sur des justifications d'ordre moral, et sur les relations entre le sujet étudié (ici : information) et le reste de l'activité des individus. Il semble que l'on trouve dans cette classe des liens, des pointeurs, avec le reste du réseau du monde vécu, le lien du lieu particulier ("information") du réseau sémantique où nous sommes avec le contexte général. Cet aspect dépasse à première vue un peu le cadre de cette analyse. Il n'en reste pas moins qu'il est présent à titre de connotation implicite dans les finalités de l'information.

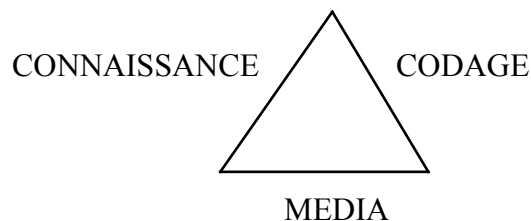
Notons que, justement, il se trouve que les débats sur les systèmes d'information ont toujours implicitement des enjeux moraux, comme en témoignent, proches de nous, la création de comités d'éthique, ou l'existence de la Commission Nationale Informatique et Libertés, et de façon plus lointaine les mythes de la boîte de Pandore ou de la pomme de la Connaissance dans le jardin d'Eden. Les

connotations de cette classe nous rappellent que les enjeux des systèmes d'information plongent dans des connotations de l'inconscient collectif qui peuvent avoir une forte réactivité morale et émotionnelle, pouvant expliquer certaines réactions qui semblent logiquement disproportionnées à l'impact réel des outils. Sans doute parce que l'information, c'est du pouvoir.

Notons que l'idée de nouveauté, qui était cruciale dans la définition de l'information par Bateson (1967), apparaît (*nouveau*), mais sans être centrale.

Ces trois premières classes (MEDIA, CODAGE, CONNAISSANCE) me paraissent devoir être considérées comme un noyau formel de l'information, qui "enveloppe" ou "contient" l'information ; bref, le message.

Fig.7 : Trois facettes du message



A ces trois classes s'ajoutent quatre autres : ACQUERIR, AVISER, INSTITUER, ENSEIGNER.

#### 4.4 ENSEIGNER

Sans ambiguïté, elle rassemble les aspects institutionnels de la transmission du savoir de maître à élève, que cette opération soit vue du maître (*enseignement*) ou de l'élève (*apprentissage*). Voici ses traits typiques :

*ecole+, enseignement+, classe+, scientifique+, primaire+, etablissement+, socia+l, etude+, secondaire+, cycle+, societe+, professeur+, groupe+, categorie+, superieur+, divis+er, formation+, travail, scolaire+, organisation+, gazette+, pedagogie+, class+er, politique+, medecine+, eleve+, diplome+, latin+, moderne+, professionnel+, maitre+, examen+, discipline+, aise+, orienta+l, langue+, classi+que, paris, quotidien+, college+, economi+que, an+, philosophique+, science+, norma+l, sporti+f, candidat+, saint+, oeuvre+, france, religieu+x, personnel+, technique+, grand+, cours, physique+, litteraire+, theorique+, etudiant+, ancien+, manuel+, specia+l, histoire+, methode+, ecolier+, siecle+, dirig+er, rousseau, dans, activite+, revue+, preparatoire+, composition+, prepar+er, argot, histori+que, genre+, travaux, savant+, hierarchi+que, organise+, salle+, institution+, utilis+er, type+,*

*domaine+, plante+, centrale+, enfant+, concours, apprentissage+, regiona+l, maternel+, centre+, education+,...*

On remarquera que cette classe ne privilégie pas un point de vue particulier (enseignant ou enseigné). Elle est symétrique, ce qui me paraît suggérer que l'enseignement/apprentissage est un processus coopératif, comme en témoigne d'ailleurs en français l'ambivalence du terme apprendre<sup>10</sup>.

#### 4.5 ACQUERIR et AVISER

J'insistais sur cet aspect symétrique car les deux classes suivantes, ACQUERIR et AVISER représentent deux aspects différents d'une transmission d'information qui me semble *unilatérale*, tous les deux en vue subjective depuis celui qui est à l'initiative de l'opération. Notons d'abord, pour modérer l'interprétation qui va suivre, que ces deux classes s'agrègent au noeud suivant de l'arbre de classification, et qu'elles sont donc assez proches.

Mais, dans la première classe, ACQUERIR, il s'agit de prolonger dans l'espace ou le temps les capacités sensorielles pour percevoir ailleurs. Il s'agit de cueillette, et la propagation de la connaissance est centripète du point de vue de l'acteur. L'information est alors l'acte d'acquérir une vision d'un état de choses du monde situé à distance de l'observateur. Tandis que dans la seconde, la propagation est centrifuge du point de vue de l'acteur.

Le noyau de sens qui sous-tend la classe ACQUERIR pourrait être caricaturé par la télésurveillance (caméras vidéos qui envoient à l'opérateur situé dans sa salle de contrôle des tableaux du monde lointain). Ce noyau, qui n'apparaît pas très clairement dans les traits typiques, me semble correspondre à une fonction d'acquisition cognitive, avec une forte connotation sensorielle. De manière imagée, les traits évoquent de façon très crue et concrète (*marcher, pied, oeil*) le processus matériel d'*aller voir (regarder...)*, de se déplacer (*marcher*) pour *éclairer, entendre (voix, bruit...)*. Les voici :

*march+er, voix, marche+, eclairc+ir, bruit+, pied+, eclair+er, front+, haut+, vent, avanc+er, oeil, eclaire+, yeux, soleil+, chien+, parl+er, lev+er, basse+, visage+, per, metaphore+, face+, entendre., mouvement+, cheva+l, chant+er, clair+, mont+er, loin, pas, boniment+, gagn+er, terre+, plein+, champ+, chant+, coup+, regard+, bout+, anima+, aller., air+, oreille+, allure+, claire+, rendre., men+er, tres, cote+, jou+er, en-tete, pa+y+er, vive+, oiseau+, tourn+er, juste, voir., bien, manqu+er, lumiere+, moins, note+, court+, frai+, figure+, penible+, bouche+, echelle+, ville+, tete+...*

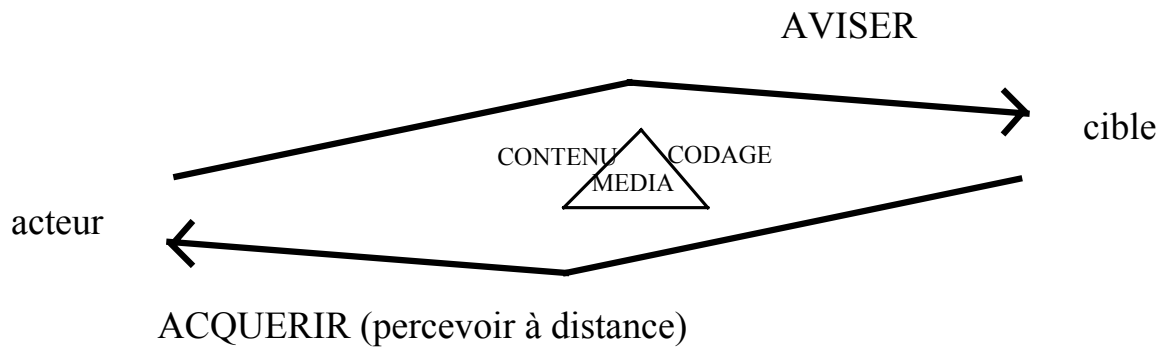


A l'inverse, la classe AVISER est plutôt une prolongation de l'action, et non de la perception. L'acteur *avise, prévient, avertit*. Il utilise l'information pour modifier l'état d'un autre opérateur distant, il signale. Ce n'est plus une caméra vidéo, mais un porte-voix qui symboliserait le contenu de ce noyau de sens.

Cette classe insiste sur l'aspect "différence qui crée une différence" (Bateson, 1967), le déclenchement de quelque chose chez la cible. Tournée vers le destinataire, la propagation d'information a d'abord ici un aspect conatif. Il s'agit d'une transmission centrifuge, du sujet vers *quelqu'un* ou *quelque chose*. Cette transmission est un stimulus que l'émetteur *donne (prévenir, informer, avertir, aviser, document (tation), recommandation, conseil, suffrage, menace, exhortation, estimation...)* ou une requête (*enquérir, interroger, consulter, demander, question, appel...*), dont le sujet est la source et non pas le destinataire, même si cette transmission peut prendre place dans un processus d'échange. Voici les traits typiques.

*avis, qqn, prevenir., inform+er, opinion+, avert+ir, avis+er, renseign+er, qqch, avertissement+, enquérir., interrog+er, informe+, avise+, document+er, consult+er, consultation+, recommandation+, donn+er, precepte+, instruire., conduite+, conseil+, unanim+e, suivre., favorable+, averti+, medecin+, suffrage+, neglig+er, presage+, inspiration+, courant+, notifi+er, conseiller+, obtenir., sage+, demand+er, partage+, menac+er, exhortation, bon+, faveur+, enquet+er, annonc+er, sur, meme, instruit+, cas, ecout+er, attention+, cherch+er, me, prevenu+, lui, document+, intention+, appu+yer, estimation, ses, demande+, question+, tenir., prendre., appel+, soi, mauvais+, symptome+, aller., mieux, qu+, vote+, academie+, partag+er, garde+, attendre., il, etre, vous, circonstance+, exemple+, examin+er, inspir+er, bonne+, votre, explication+, sond+er, son, porte+, accident+, spécialiste+, dessein+, expose+, ciel+, emettre., sujet+, exprim+er, telle, quelque-chose, nuit+, indication...*

Je crois voir ici intervenir deux modes différents de la transmission d'information : de la cible vers le sujet (acquisition) et du sujet vers la cible (fourniture d'avis).



Ces deux classes s'agrègent au noeud supérieur : elles sont deux formes *d'extension ubiquitaire* du sujet, qui étend son rayon de perception (ACQUISITION) ou d'influence (AVISER) à distance.

#### 4.6 INSTITUER

Dernière classe, celle que j'ai appelée INSTITUER, et qui correspond aux traits typiques suivants :

*dr, enquete+, procedure, juge+, conseil+, ordre+, instruction+, tribuna+l, instance, assemblée+, acte+, loi+, justice+, publique+, decision+, civil+, civile+, capitant, public+, administrat<, expertis+er, criminel+, membre+, ministre+, vote+, declaration+, officier+, police+, proces, affaire+, militaire+, autorite+, armee+, judiciaire+, arret+, etat+, droit+, code+, president+, proced+er, interrogatoire, prescription, devant, reunion+, obligation+, investigation, juridiction, travaux, nationa+l, cour+, juridique+, prive+, public, temoin+, premier+, verifi+er, officiel+, lega+l, ordonnance+, roi+, constitution<, contrat+, renvoi+, audience, defense+, cause+, charge+, execut+er, ancienn, election+, reglement+, trait+er, instructeur, seconde+, faite+, élu+, milit, genera+l, projet+, superieur+, preparatoire+, chambre+, proclamation, institution+, laquelle, etabl+ir, ministere+, paix, faits, universel+, delit+, mission+, condition+, principa+l, deliber+er, reun+ir, commission+, par, demande+, art+, obligatoire+, huissier+, sign+er, lors-d+, execution+, hierarchi+que, religieu+x, action+, chef+, partie+, electora+l, attention+, direction+, second+, mariage+, regl+er, port+er, crime+, honneur+, enregistrement, degre+, matiere+, lequel, consigne+, expert+, point+, estimation, sondage, resultat+, suffrage+, securite+, communaute+, jugement+, bureau+, disposition+, gouvernement<, question+, particulier+, corps, division+, guerre+, registre+, util+e, notification, dossier+, appreciation+, commandement+, conge+, engagement+, liste+, statistique...*

L'idée qui prévaut ici est celle d'une explicitation officielle de quelque chose, devant *témoins*, au sens *juridique* du *constat*. Les traits *conseil, tribunal, assemblée, jugement, vote* etc. amènent à penser que ce qui est ici crucial, c'est l'aspect **validation collective** de la procédure. Cette validation est sanctionnée, *établie*, par une *assemblée* légitime, une *administration* ou des *experts* délégués à cette fin par l'organisation (*docteur, juge...*). L'information est ici la mise sur la place publique d'un état de choses, afin de parvenir à un consensus sur "ce qui est", ce qu'on peut appeler une *INSTITUTION*. On dépasse ici le schéma standard, sans doute dérivé du schéma A-B-X de Newcomb dans son papier séminal de 1953, qui sous-tendait jusqu'ici les théories de la communication, et dans lequel une source communique un message à un (ou des) destinataire(s). Ici, c'est l'organisation qui parle à l'organisation.

Et, réflexion faite, aucune organisation ne saurait fonctionner sans procédures de ce type. Il est clair que les circulaires, les *décisions*, les *arrêtés, actes, règlements, prescriptions, ordonnances, notification, contrats, constats*, produits par les *commissions, assemblées* et autres *instances* sont des informations de ce type. Il s'agit ici d'une institutionnalisation d'un état de choses. On notera avec amusement qu'à côté des procédures juridiques, qui représentent la forme traditionnelle de l'institutionnalisation, apparaissent également la forme plus moderne des *estimations, sondages, suffrages, statistiques*, qui constituent désormais les modernes fonds de *dossiers*.

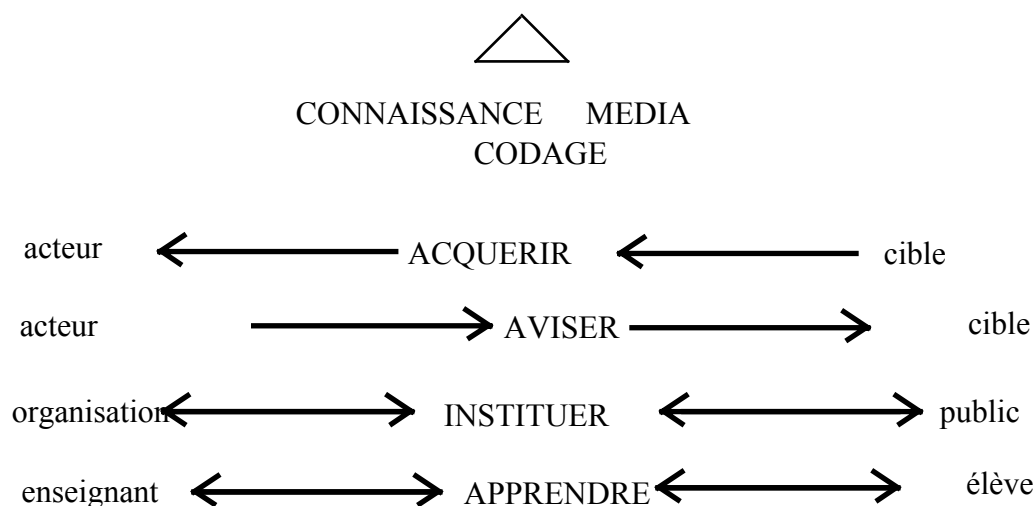
Cette dernière classe INSTITUER est particulièrement intéressante pour une étude dans le cadre de l'entreprise :

- dans la mesure où elle distingue une information collective de l'information dont l'échange concerne simplement une destination et une source individuelles ;
- parce qu'elle révèle un mécanisme social important et fondateur des organisations : l'institution, ou validation collective.

#### **4.7 Vue générale**

On obtient donc finalement le schéma suivant :

Fig. 8 : Les avatars de l'information



Ce schéma se lit de la façon suivante : l'information prend la forme d'une CONNAISSANCE CODÉE dans un MÉDIA. Ce tableau du monde se propage d'un objet (en général : un acteur) à un autre. On peut alors distinguer quatre grands types de propagation de l'information :

- ACQUERIR : un acteur acquiert, à distance, une perception de l'objet (qui peut être un autre acteur) ; l'objet est relativement passif, il y a collecte d'information.
- AVISER : un acteur stimule à distance un objet en lui fournissant de l'information ; l'objet est perçu comme un système réactif dont l'acteur cherche à déclencher une réaction en réponse à sa stimulation (réaction qui peut être simplement : " être au courant ").
- INSTITUER : un système organisé produit un constat, un tableau du monde à titre de repère de coordination pour l'ensemble des acteurs concernés (ceux qui coopèrent dans l'organisation) ; ce faisant il acte un consensus, capitalisant ainsi un état de la connaissance sous une forme réutilisable ultérieurement par l'organisation.
- APPRENDRE : un couple d'objets communique pour qu'une partie des connaissances de l'un (enseignant) soit transférée à l'autre (élève).

Ces cas constituent-ils une classification pertinente des situations de communication ? Sont-ils exhaustifs ? Probablement pas, ou du moins il serait imprudent de le tenir pour acquis à ce stade. Mais celle-ci correspond en tous cas à une certaine vérité culturelle et représente ce qu'on peut tirer du savoir

accumulé dans la structure interne de la langue française sur ce sujet avec les méthodes dont nous disposons actuellement.

La schématisation avec les flèches est à revoir. Elle vise surtout à montrer que la propagation est unilatérale dans les deux classes ACQUERIR et AVISER, mais que dans INSTITUER et APPRENDRE, il se produit une interaction plus complexe entre source et destinataire.

L'examen de la manière dont se raccordent les classes au niveau supérieur amène à distinguer, d'un côté le support (MEDIA + MESSAGE), les procédures de mise en commun, ou, comme dirait Bateson, de reproduction du même (APPRENDRE & INSTITUER), la transmission de signal, centripète ou centrifuge (ACQUERIR, AVISER), et enfin le contenu de l'information, c'est-à-dire CONNAISSANCE, avec tout le flou que recouvre cette classe (savoir, affects, normes, culture...).

Fig. 9 : L'arborescence des classes



Considérons donc cette modélisation conceptuelle comme une base de travail, qui devra être enrichie et validée par l'expérience empirique. Il est probable que l'interprétation que je fais aujourd'hui de ces classes changera un peu quand j'en saurai plus, et un autre analyste tirerait sans doute du même matériau quelque chose d'un peu (mais sans doute pas tellement) différent : l'interprétation, subjective, est en effet toujours dépendante des connaissances (implicites dans l'analyse) que l'analyste a de la question.

Néanmoins, il est possible d'obtenir un certain accord des analystes. Dans une expérience, au séminaire de formation à ALCESTE (Carcassonne, 1995), j'ai fourni à 16 participants la liste des traits typiques des deux classes que "média" et "acquérir" décrites plus haut (sans autre préparation que de leur dire comment avait été constitué le corpus, et évidemment sans leur donner mon interprétation) en leur demandant, individuellement, en 20 minutes, de nommer ces classes à partir de l'examen des traits. C'est un exercice extrême, dans la

mesure où les participants n'avaient à leur disposition qu'un temps restreint et seulement une petite partie des résultats habituellement disponibles.

Pour la première classe, dont les traits sont très homogènes, les réponses sont très concordantes : « presse écrite », « media », « informations comme messages et leurs supports », « journalisme », « moyens », « presse », « média, supports d'écriture », « supports de l'information sans aspects de contenu », « supports de communication », « média/catégories de messages », « énonciation au sens linguistique », « canal utilisé pour transmettre l'info », « support et transfert de l'information dans ou par les agences de presse », « support d'information », « supports visuels de communication », « supports de communication sur le plan écrit ».

Pour l'autre classe, choisie justement en raison de sa difficulté d'interprétation, les réponses sont moins concordantes, mais ne sont pas non plus aussi divergentes que je l'attendais (j'ai noté d'une étoile les interprétations qui convergent avec la mienne) : \*« s'informer (action de s'informer) », « influencer, agir », « l'information au naturel, dit de façon non technique », \*« point de vue », « relations de pouvoir, influence. en opposition aux objets », « approche sensorielle. relation de confiance (ou manque de confiance) », \*« sensoriel. destination, provenance. », « outils d'échange de l'info, communication au sens strict avec une notion dynamique », \*« appréhension du réel », \*« perception : rapport du sujet à l'information (dont physique et symbolique: jugement) », \*« proprioceptif. entre soi et le monde extérieur. et évaluation », « finalité de l'information », « doute sur la véracité et la qualité de l'information », « classe d'action », \*« flux et passage. transfert de quelque chose », \*« perception ».

On le voit, l'interprétation est un processus relativement robuste, mais qui a des limites.

## **5. Conclusion :**

La méthode décrite me semble intéressante si elle est maniée avec prudence, et dans une perspective exploratoire. La technologie est mûre. Du travail reste à faire sur la théorie de l'interprétation, mais, concrètement l'ouverture d'une famille de solutions empiriques à la question du sens semble pouvoir être soutenue.

Les sources savantes, comme le dictionnaire, sont plus riches que les sources profanes, et plus faciles d'accès. Elles sont également moins porteuses d'implicites que la conversation ordinaire, ce qui facilite la découverte de traits "évidents" dont l'explicitation est difficile avec les méthodes traditionnelles.

Je pense que cette technique, ou d'autres analogues, pourraient être appliquées avec profit pour la construction, notamment, de réseaux sémantiques ayant de réels fondements distributionnels.

Dans cette perspective, pour finir, trois remarques.

La méthode que j'ai présentée reste encore très dépendante de l'interprétation par l'analyste, qui nomme lui-même les classes. C'est lourd pour construire des réseaux sémantiques. Mais l'opération d'abduction repose finalement sur une base de connaissances (l'analyste), qui pourrait peut-être être remplacée par une autre base, moins complète mais plus économique (ex : un dictionnaire, un gros corpus...). Repérer le paradigme "A" à partir des traits ( $a_j$ ) caractérisant la classe pourrait alors se faire par la minimisation d'une distance lexicale, par exemple entre les vecteurs des ( $a_j$ ) pondérés par leurs fréquences d'occurrence, dans la classe à interpréter, et dans les mini-corpus constitués par chaque entrée d'un dictionnaire des "A" potentiels...

J'ai utilisé le dictionnaire comme source d'énoncés. Mais on pourrait en prendre de nombreuses autres, par exemple (en prenant comme corpus l'ensemble des phrases qui contiennent le mot étudié) FRANTEXT...

Enfin, il faut être conscient que la nature du corpus influe considérablement sur les résultats. Même si tous les corpus parlent, plus ou moins, de notre Monde, ils le font suivant des perspectives différentes, qui en donnent une image particulière. La nature du corpus doit être adaptée à la stratégie de recherche : on ne peut pas *tout* trouver dans *n'importe quel* corpus ! Je n'ai pas insisté sur ce point, car c'est une évidence ; mais que l'on oublie souvent au moment de l'analyse, car les choix de corpus sont souvent autant dictés par la commodité, ou la disponibilité, que par leur pertinence pour la recherche entreprise.

## 6. Bibliographie

- ABRIC, Jean-Claude (1994). Les représentations sociales : aspects théoriques. In : Abrid, J-C. (éd.), *Pratiques sociales et représentations*. Paris : P.U.F., 1994. pp.11-35. ISBN 2 13045895 5.
- ANZIEU, Didier (1974). Le moi-peau. *Nouvelle Revue de Psychanalyse*, n° 9, printemps 1974. "Le dehors et le dedans". pp. 195-208.
- BATESON, Gregory (1967). Style, grâce, information dans l'art primitif. In. Gregory Bateson, *Vers une écologie de l'esprit*. Tome I. pp. 140-164. Paris : Seuil, 1977. ISBN 2. 02. 004700 4. pp. 209-224.
- BATESON, Gregory (1979). *La nature et la pensée*. Paris : Seuil. 1984. ISBN 2. 02.006794.3.
- BEAUDOUIN, V, LAHLOU, S (1993). L'analyse lexicale, outil d'exploration des représentations. *Réflexions illustrées par une quinzaine d'analyses de corpus d'origines très diverses*. Cahiers de recherche du Crédoc, n°48, septembre 1993. et 48bis.
- BEAUDOUIN, Valérie, HÉBEL Pascale (1994). *Avancées en analyse lexicale*. Paris : Crédoc, Cahier de recherche, n°61, 1994.
- BEAUDOUIN, Valérie, LAHLOU, Saadi, YVON, François (1993). Analyse d'une question ouverte : incidence du mode de questionnement. Actes des secondes journées internationales d'analyse statistique des données textuelles. Montpellier, 21-22 octobre 1993. Paris, ENST. pp. 133-145.
- CHIVA, Matty (1979). Comment la personne se forme en mangeant. *Communications*, 31, Automne, 1979. pp. 107-118.
- CHIVA, Matty (1985). *Le doux et l'amer : sensation gustative, émotion et communication chez le jeune enfant*. Paris : P.U.F, 1985. 256 p.
- CODOL, Jean-Paul (1969). Note terminologique sur l'emploi de quelques expressions concernant les activités et processus cognitifs en psychologie sociale. *Bulletin de Psychologie*, 1969, 23. pp. 63-71.
- CORDIER, Françoise (1991). Les représentations privilégiées dans tous leurs états. *Psychologie française : Le traitement cognitif du texte*. n° 36-2, 1991. pp. 119-128.
- DURKHEIM, Emile (1912). *Les formes élémentaires de la vie religieuse*. Paris : Editions Générales Françaises, Le livre de poche, 1991. ISBN 2 253 05849 1.
- FISCHLER, Claude (1990). *L'omnivore*. Paris : Odile Jacob, 1990. ISBN 2 7381 0101 1.



FLAMENT, Claude (1993). Structure et dynamique des représentations sociales. In : Jodelet, Denise (éd.), *Les représentations sociales*. Paris : P.U.F. , 3ème édition 1993. pp. 204-219. ISBN 2 13 045615 4.

GRICE, Paul. H. (1975). Logique et conversation. *Communications*, 30, 1979, pp. 57-72. Traduit de : *Logic and Conversation, Syntax and Semantics*, vol. III. In : P. Cole et J.L. Morgan, (eds), *Speech Acts*. Academic Press, Inc., 1975. pp. 41-58.

HALBWACHS, Maurice (1925). *Les cadres sociaux de la mémoire*. Paris : P.U.F, 1925. Réed. Paris - La Haye, Mouton.

INSEE (1991). Les Dépenses de consommation dans huit états membres de la Communauté Européenne en 1988. Statistiques Rapides. Population et conditions sociales, n°7, novembre 1991.

LAHLOU, Saadi (1990). *Eléments de formalisation pour une théorie de l'évolution des systèmes : la système-compatibilité*. Paris : CREDOC, Cahier de recherche, n° 4, octobre 1990.

LAHLOU, Saadi (1994). L'analyse lexicale. *Variances*, n°3, oct. 1994. pp. 13-24.

LAHLOU, Saadi (1995a). L'utilisation de l'information dans l'entreprise : quelques réflexions théoriques et une analyse lexicale. EDF-Direction des Etudes et Recherches. Note HN 5194055, janvier 1995

LAHLOU, Saadi (1995b). *Penser Manger. Les représentations sociales de l'alimentation*. Thèse, Ecole des Hautes Etudes en Sciences Sociales. 1995

LAPLANCHE, Jean, PONTALIS, J. -B. (1990). *Vocabulaire de la psychanalyse*. Paris : P.U.F. 10ème édition, 1990. ISBN 2 13 043288 3.

LEBART, Ludovic. Allocution de clôture. Actes des secondes journées internationales d'analyse statistique des données textuelles. Montpellier, 21-22 octobre 1993. Paris, ENST.

LE ROBERT ÉLECTRONIQUE (1991), outil d'aide à la rédaction sur la base du Grand Robert de la Langue Française. Paris : Dictionnaire le Robert, 1991.

MESCHONNIC, Henri (1991). *Des mots et des mondes. Dictionnaires, encyclopédies, grammaires, nomenclatures*. Paris : Hatier, 1991. ISBN 2 218 03726 2.

MOSCOVICI, Serge (1961). *La psychanalyse son image et son public*. Paris : P.U.F. , 1976.

MOUNIN, Georges (1963). *Les problèmes théoriques de la traduction*. Paris : Gallimard, 1963.

NAHOUM-GRAPPE, Véronique (1991). Histoire et anthropologie de l'excès : quelques pistes  
In : Pierre Chambat (éd.), *Modes de consommation, mesure et démesure*. Paris : Editions  
Descartes, coll. Université d'été, 1991. pp. 145-161. ISBN 2 909728 00 5.

NEWCOMB, Theodore A. (1953). Ebauche d'une théorie de la communication. Trad. fr. In :  
Faucheux, Claude & Moscovici, Serge (eds.), *Psychologie sociale théorique et expérimentale*.  
Paris : EHESS, Mouton, coll. Les textes sociologiques, 8. 1971. pp. 21-38.

RASTIER, François (1994). Microsémantique, lexique et contexte. Pp. 109-147. IN Eveline  
Martin (éd.) Traitements informatisés de corpus textuels. Collection "Etudes de sémantique  
lexicale". INaLF. Paris : Didier Erudition. 1994. ISBN 2-86460-224-5.

REINERT, Max (1987a). Classification descendante hiérarchique et analyse lexicale par  
contexte : application au corpus des poésies d'Arthur Rimbaud, *Bulletin de Méthodologie  
Sociologique*, n° 13.

REINERT, Max (1987b). Un logiciel d'analyse des données textuelles : ALCESTE.  
*Communication aux 5èmes Journées Internationales "Analyse de données et Informatique"*.  
INRIA.

REINERT, Max (1990). ALCESTE, une méthode d'analyse des données textuelles.  
Application au texte "Aurélia" de Gérard de Nerval. *Bulletin de Méthodologie Sociologique*,  
26. pp. 25-54.

ROUQUETTE, Michel-Louis (1994). Une classe de modèles pour l'analyse des relations entre  
cognèmes. In : Guimelli, C. (éd.), *Structures et transformations des représentations sociales*.  
Neuchâtel : Delachaux et Niestlé, 1994. pp. 153-170. ISBN 2 603 00945 1.

SAUSSURE, Ferdinand de. (éd. posthume, 1915). *Cours de linguistique générale*. Paris :  
Payot, 1972, 1985. ISBN 2-228-50070-4.

SHANNON, Claude E. , WEAVER, Warren (1948). *The Mathematical Theory of  
Communication*. Illini Books edition, University of Illinois Press, 1963. ISBN 0-252-72548-4.

SPERBER, Dan (1989). L'étude anthropologique des représentations : problèmes et  
perspectives. In : Jodelet, Denise (éd.), *Les représentations sociales*. Paris : P.U.F. , 1989.  
pp. 115-130. ISBN 2 13 042102 4.

SPINOZA, Baruch (1677). *Ethique*. Trad. Raoul Lantzenberg. Paris : Flammarion, 1947.

WHORF, Benjamin Lee (1927). On the Connection of Ideas. In : John B. Carroll (ed.),  
*Language, Thought and Reality*. Cambridge (Massachusetts) : The MIT Press, 1956. pp. 35-  
39. ISBN n° 0 262 73006 5.

WITTGENSTEIN, Ludwig (1958). *Le cahier bleu*. Paris : Gallimard, 1965.

## Notes

---

<sup>1</sup> "(...) le langage n'est pas plus fait de mots que les encyclopédies n'énumèrent des personnes ou des objets. Ces listes ne tiennent que parce qu'on projette de la langue sur des choses, et du discours sur la langue." (Meschonnic, op. cit. p. 9).

<sup>2</sup> Techniquement, on fait une ACP sur le tableau de contingence UCE x lexèmes, et on fait glisser le premier hyperplan factoriel sur le premier axe factoriel pour obtenir une bipartition optimale en termes d'inertie interclasses, qui construit C1 et C2 comme les contenus respectifs des deux demi espaces complémentaires séparés par l'hyperplan. L'algorithme, qui utilise une procédure de classification croisée pour optimiser les classes, est plus sophistiqué que la procédure ici décrite. Voir Reinert (1990).

<sup>3</sup> Le chi2 mesure l'écart à une distribution régulière du vocable sur le corpus. Il est d'autant plus élevé que le % dans la classe est "anormalement" élevé.

<sup>4</sup> L'analyse des seuls analogues de premier niveau de l'entrée "manger" fournit les quatre premières classes, mais pas ces deux là.

<sup>5</sup> Sur ces questions, voir par exemple Flament (1993), Codol (1969), Abric (1994), Rouquette (1994), et Lahlou (1995) pour une présentation détaillée.

<sup>6</sup> "Le signe (la phrase) prend un sens par référence à un système de signes ou langage, auquel il appartient." (Wittgenstein, 1958, 1965, p. 31)

"la partie conceptuelle de la valeur [d'un terme] est constituée uniquement par ses rapports et ses différences avec les autres termes de la langue (...)" (de Saussure, 1915, 1985 p. 163)

<sup>7</sup> Par exemple, les "taxèmes" ou les "champs" (Rastier, 1994, pp. 125-128). Il nous faut cependant abandonner l'espoir de déterminer des éléments "atomiques", en raison même de la nature fractale du sens : tout élément, vu de près, se scindera à son tour en éléments.

<sup>8</sup> "Il existe un véritable postulat de Bloomfield (jamais assez mis en relief au cours des discussions) qui justifie la possibilité de la science linguistique en dépit de la critique bloomfieldienne de la notion de sens, postulat qu'on doit toujours remettre au centre de la doctrine bloomfieldienne après l'avoir critiquée : "comme nous n'avons pas les moyens de définir la plupart des significations, ni de démontrer leur constance, nous devons adopter comme un postulat de toute étude linguistique, ce caractère de spécificité et de stabilité de chaque forme linguistique, exactement comme nous les postulons dans nos rapports quotidiens avec les autres hommes. Nous pouvons formuler ce postulat comme l'hypothèse fondamentale de la linguistique : dans certaines communautés (communautés de langue) il y a des énoncés linguistiques qui sont les mêmes quant à la forme et quant au sens" (Bloomfield, Language, p. 144, cité par Mounin, 1963, p. 30)

---

<sup>9</sup> Autrement dit, l'analyste choisit les traits "pertinents" de façon discrétionnaire, suivant un critère subjectif d'idonéité à ses motivations de recherche.

<sup>10</sup> Qui est une aubaine pour mon interprétation ici, mais ne semble pas être un universel linguistique.