

DEUXIÈME ÉCOLE D'ÉTÉ  
DE DIDACTIQUE DES MATHÉMATIQUES  
OLIVET 5-17 JUILLET 1982

SUR LES CORPUS EXPÉRIMENTAUX

Yves Chevallard

## AU LECTEUR

Le thème des "corpus expérimentaux" (et les questions connexes d'épistémologie et de méthodologie) est l'un des principaux thèmes sur lesquels le collectif d'organisation de la 11<sup>ème</sup> école d'été de didactique des mathématiques a voulu attirer l'attention des intervenants et des participants. Le présent document se veut une contribution au travail que nous mènerons à ce propos. De même que le document "Sur l'ingénierie didactique", les lignes qui suivent sont extraites de Notes écrites à l'intention du collectif d'organisation; le style, les contenus, et jusqu'à la lisibilité, en portent la marque. J'espère que le lecteur voudra bien ignorer les multiples imperfections dont, les circonstances aidant, je suis le seul responsable.

Y. C.

juin 1982

## SOMMAIRE

- Toute science est expérimentale (p.1)  
Ce qu'est la "matière" de nos corpus (p.3)  
Premier corpus, première théorie (p.14)  
"Introspection" et "anamnèse"... (p.20)  
Le mouvement théorique - expérimental (p.24)  
Le mot même de corpus, son origine (p.27)  
Corpus et empirisme (p.29)  
La notion de schème générateur et l'exemple  
du schème laplacien (p.32)  
L'enclanchement du processus théorique - expérimental (p.41)
- Annexes (p.42)

## SUR LES CORPUS EXPÉRIMENTAUX

[ TOUTE SCIENCE EST EXPÉRIMENTALE ]

3.3. J'ai poussé à la roue pour que la dimension "ingénierie" soit prise en compte préférentiellement à Orléans ; j'ai ajouté que je ne suis pas moins attaché à la dimension "méthodologie", et que je serais heureux si elle faisait effectivement l'objet d'un questionnement systématique dans le cadre de l'École d'Été. Je le serai d'autant plus que c'est à ce niveau, me semble-t-il, que la fermeture de l'espace de parole, que j'évoquais dans ma Note 2, peut avoir les effets les plus dommageables, et qu'elle apparaît dans le même temps le plus assurée. Il faut donc travailler à ouvrir un espace de parole où en effet puisse se faire entendre, et être entendu, plus d'un discours ; et qui donne sa chance à une imagination méthodologique féconde pour notre discipline. Il m'apparaît qu'aujourd'hui - et je voudrais qu'aient raison ceux qui me répondront qu'en cela je me trompe - un unique discours se tient, et quelquefois rouffère ; un discours sans nuance épistémologique, à la fois

discours dominant, discours normalisateur, discours luminant ; celui d'un certain "experimentalisme" d'importation, soutenant ses prétentions d'une référence, critique en ce qui concerne sa problématique, à critique pour ce qui est des procédures, à quelques secteurs de champs voisins (psychologie expérimentale, etc.). Discours fortement empiriste - j'essaierai de le montrer. Mais posons d'abord un postulat épistémologique : toute science est expérimentale\*. Le problème est donc des formes propres d'expérimentation qui confèrent à chacune son visage particulier. Or, aujourd'hui, la didactique des mathématiques est un champ scientifique méthodologiquement amorphe (alors que le débat scientifique a relativement bien fait progresser la structuration en ce qui concerne sa problématique\*\*) ; la dimension méthodologique constitue son "ventre mou" - quel que soit le discours méthodologique tenu, nous revenons bientôt à l'étiage de l'experimentalisme. Voilà pourquoi je crois indispensable que le fer soit maintenant porté dans la plaie - parce que je vois que nous avons tous à y gagner. Il est devenu aujourd'hui nécessaire d'examiner sans détour la question des voies et moyens de notre rencontre avec le réel : ne pas le faire serait confirmer tacitement notre assujettissement au seul paradigme méthodologique actuellement légitime (et donc, par une contagion qui va du droit au fait, du seul paradigme pratiquement disponible).

3.4. La question sous laquelle on peut ranger toute l'interrogation méthodologique - ou du moins la question qui m'apparaît

\* Même les mathématiques ; mais c'est une autre affaire.

\*\* Bien entendu, il reste beaucoup à travailler sur ce point...

la plus pressante entre nous, à la suite des conversations que j'ai eues récemment - est celle-ci : étant donné une assertion sur notre objet d'étude (par exemple : "en 4<sup>e</sup>, il arrive régulièrement que des élèves, discutant entre eux à propos d'une figure de géométrie, disent "rond" pour "cerle" ), qu'est-ce qui fait que je tiendrai ou non cette assertion pour vraie ? Observons tout de suite qu'il est déjà plus particulier, et si j'ose dire plus engagé (épistémologiquement), de formuler ainsi la question : " De quel corpus vais-je construire la confirmation de cette assertion ou, en sens contraire, instruire son procès ? " C'est à peu près cette formulation (me semble-t-il) qui a été retenue implicitement dans nos discussions préparatoires jusqu'à présent - et la mention explicite, à propos du titre I, du souci de " présentation du travail de construction des corpus nécessaires..." paraît en prendre acte. Ce découpage a l'immense avantage de porter l'attention au lieu du trauma, sans tourner autour du pot ; et je crois que c'est par là qu'il convient d'attaquer.

### [ CE QU'EST LA "MATIÈRE" DE NOS CORPUS ]

3.5. Question préliminaire : sur quoi fait-on assertion, et de quoi fait-on corpus quand on fait de la didactique des mathématiques ? Je propose de trier un peu à partir d'un petit schéma théorique (eh oui !) que j'ai déjà pris la liberté

de présenter.

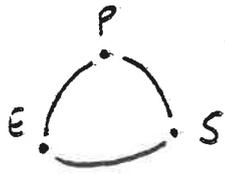


Fig. 1.

L'objet d'étude stricto sensu du didacticien est le système didactique, soit la triade Professeur - Elève - Savoir - les trois termes "Professeur", "Elève", "Savoir" désignant des objets du système (des sous-systèmes) - plus les interactions entre ces objets.

Ce que nous cherchons à savoir, quand nous faisons de la ddm, c'est le jeu qui se mène entre ces trois sous-systèmes - avec l'idée que ce savoir nous permette d'exercer ensuite une action, non totalement opaque, sur le système didactique (ingénierie didactique). Par exemple l'assertion que j'ai énoncée plus haut affirme quelque chose à propos de ce système interactif de trois "places" : des élèves communiquent entre eux à propos d'un certain élément de savoir (une figure géométrique) - voilà déjà deux des trois fonctions annoncées. Et le "Professeur" (la fonction-professeur), dira-t-on ? Eh bien, il se manifeste ici en ceci que le professeur est supposé absent, ou du moins que les propos échangés entre les élèves ne s'adressent pas à lui (par exemple parce que les élèves travaillent entre eux en dehors de la classe). Notons en passant que l'"absence" du professeur n'est pas ici l'effet (que l'on pourrait alors considérer comme une limitation, ou au contraire comme une condition intentionnellement recherchée) d'une certaine procédure expérimentale ; elle se rapporte à des phases du fonctionnement didactique "normal" (soit en classe : travail - licite - à plusieurs autour d'un problème, ou communication - toléré ou illicite - dans une recherche en principe individuelle ; soit hors classe : travail en commun à la maison) ; et le questionnement expérimental de l'assertion qui la met en jeu suppose

par conséquent la résolution du problème de son observation in vivo ou de sa reproduction approchée — ce qui est un autre problème, de procédure expérimentale précisément.

3.6. Cela étant, il se trouve qu'on ne parvient pas toujours à comprendre ce qui se passe, dans un système didactique, en ne considérant que ce système seulement. Il faut aller voir ce qui se passe à l'extérieur du système, bien souvent, pour éclairer le fonctionnement didactique stricto sensu. La première zone dans laquelle on pénètre en quittant le système didactique (dans le schéma que j'avance :

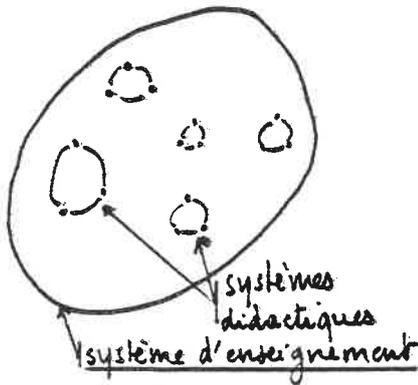


Fig. 2.

je poserai plus loin le problème de sa "vérité"), c'est ce que j'appellerai le système d'enseignement. Outre les systèmes didactiques, le système d'enseignement contient, par exemple, un ensemble de dispositifs de régulation des flux d'élèves, qui concourent à la formation des systèmes didactiques. Ces dispositifs peuvent

être officiels (orientation scolaire) ou officieux (intervention discrète mais systématique sur la composition des classes de la part de l'administration de l'établissement); se référer à la matière considérée (les mathématiques) ou à une autre (classes de latinistes, etc.), ou adopter encore d'autres critères plus ou moins occultes. La relation entre la survenue de certains faits au sein du système didactique et certaines caractéristiques du système d'enseignement doit être considérée, du point de vue de la didactique des mathématiques (il n'en irait pas forcément ainsi du point de vue de la sociologie de l'éducation par exemple),

par le biais de la retraduction didactique ( en des termes se rapportant au fonctionnement du système didactique uniquement ) des caractéristiques du système d'enseignement que l'on aura éventuellement mises en cause.\* Supposons ainsi que l'on ait pu établir ( d'une manière que je suppose satisfaisante - écartant par là le problème-même qu'il s'agit d'examiner ) le fait didactique suivant :

" Par rapport aux années immédiatement antérieures, l'année 1981/82 a vu, dans les classes de seconde, une diminution, de la part donnée aux mathématiques "non algorithmiques", au profit des mathématiques consistant surtout en la mise en œuvre d'algorithmes ".

On pourra alors être tenté d'"expliquer" ce fait didactique en le reliant au fait - qui concerne le système d'enseignement, et plus précisément les mécanismes de régulation des flux d'élèves à l'intérieur de celui-ci - que

" En 1981/82 on a pour la première fois abandonné la différenciation ( dans le cadre de filières courantes jusqu'au baccalauréat ) des classes de seconde ".

On ne tiendra une ( éventuelle ) explication didactique qu'en désignant un fait didactique qui procéderait immédiatement de la cause ainsi énoncée ( qu'on peut bien appeler une "cause", au sens où elle produit l'effet attesté ). Par exemple,

\* Cette affirmation posée dogmatiquement n'est qu'une autre manière de souligner, en paraphrasant Durkheim, qu'un fait didactique doit être expliqué par un fait didactique - et non par un fait d'une autre espèce.

" En 1981/82, il y a eu diminution sensible de l'homogénéité des classes, en ce qui concerne les mathématiques, par rapport aux années immédiatement antérieures ".

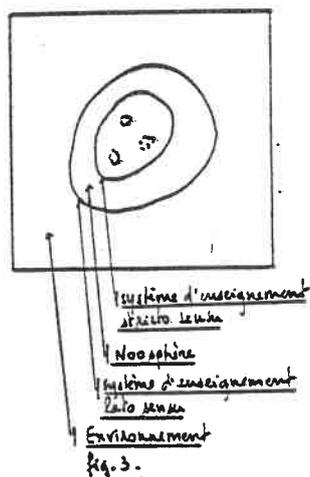
(Il resterait alors à montrer en quoi ce fait didactique permet d'expliquer - théoriquement et expérimentalement - le fait didactique primitivement examiné ).

3.7. Le caractère systémique du schéma présenté permettra alors de se livrer à la prédiction suivante :

" En 1982/83 on peut s'attendre - en dehors d'une éventuelle décision officielle dans ce sens - à la mise en place, au sein du système d'enseignement, de mécanismes compensatoires officiels visant à rétablir au moins partiellement l'homogénéité antérieure "

(faisant ainsi l'hypothèse d'une retroaction négative, allant des systèmes didactiques vers le système d'enseignement). Etc.

3.8. Il faut parfois - et même souvent - aller chercher au-delà du système d'enseignement pour comprendre telle particularité



du fonctionnement didactique. Entre le système d'enseignement dont j'ai parlé, et son environnement (son écosystème), il y a une zone intermédiaire, une couronne que j'ai proposé d'appeler la noosphère (la sphère où l'on pense...), et qui contient tout ce qui "pense" sur le système d'enseignement : aussi bien les "militants" de ce système (au sein de l'APM par

exemple) que les Inspecteurs Généraux, etc. La noosphère doit être considérée comme une couche-tampon entre système et environnement, sorte de sas qui est à la fois lieu d'expression privilégié des conflits entre système et écosystème, lieu spécifique où se mènent les négociations et s'élaborent les transactions permettant la résolution locale de ces conflits. Par rapport à ce schéma d'ensemble, le système d'enseignement que j'ai introduit en premier lieu peut être appelé plus précisément "système d'enseignement stricto sensu", et l'ensemble qu'il forme avec la noosphère "système d'enseignement lato sensu".

3.9. D'une manière générale, on peut suivre en remontée, depuis le système didactique jusqu'à l'environnement, les enchaînements causaux dont nous\* observons les effets au sein du système didactique. Cela ne signifie nullement que cette remontée soit toujours possible. Par exemple un remaniement au sein du système d'enseignement, tel l'indifférenciation des classes de seconde, peut très bien résulter d'une pression exercée par l'environnement (mettons: les parents d'élèves), reprise et reformulée par la noosphère; mais on peut aussi imaginer, en l'espèce, que ce remaniement trouve sa motivation dans un phénomène pour l'essentiel interne au système d'enseignement (par exemple parce qu'on attend de lui qu'il amène la réduction de certaines tensions, etc.\*). Mais quelquefois il faut remonter plus haut. C'est ainsi que le fait didactique dont nous sommes partis (la part plus grande prise par l'"algorithmique") peut être mis en relation avec un fait (bien observable) affectant la noosphère: l'existence d'un courant - s'exprimant par des textes, des slogans ("faïe des mathématiques", etc.), des "rumeurs", etc. - mettant en avant l'"importance" de l'activité de l'élève, et traduisant cela par le thème, dûment nommé, et négociable avec l'administra-

\* Pour éclairer d'un seul exemple cette indication: lors d'une

tion, des "activités", thème qui peut s'inscrire dans la nomenclature des actes pédagogiques. Ce courant semble être l'expression d'un conflit, non pas interne au système d'enseignement, mais au niveau même de la noosphère, où il retraduirait par exemple certaines tensions entre système et environnement\*. Dans cette ligne, on mettra éventuellement en rapport le fait didactique de la majoration des mathématiques à base d'algorithmes avec le fait didactique suivant (qui résulte immédiatement d'une décision de la noosphère):

"En 1981/82, la part d'activité accordée aux élèves, sous la forme d'activités, a sensiblement augmenté par rapport aux années immédiatement antérieures".

On notera que, dans la réalité, les deux facteurs didactiques invoqués ici (diminution de l'homogénéité, accroissement des phases d'activité des élèves) ont pu conjugué leurs effets (un malheur n'arrive jamais seul) ; et que, dans leur principe, les

---

interview réalisé en juin 1981 (donc avant la mise en œuvre de la décision d'indifférenciation) auprès d'une conseillère d'orientation qui se déclarait favorable à la réforme des classes de seconde, j'ai eu en guise de justification de cette attitude la réponse que voici : depuis des années les conseillers d'orientation doivent affronter, à la fin de la classe de troisième, le "fantasme de la seconde C", source d'attentes irréalistes et de frustrations qu'il est à peu près impossible de "gérer" de manière supportable - pour les uns et les autres. Avec l'indifférenciation le problème disparaît, chacun peut imaginer faire une seconde C, et tenter sa chance...

\* Là encore, un mot d'éclaircissement : l'échec de la réforme des "maths

décisions qui sont à l'origine possible de ces faits didactiques (l'indifférenciation des classes; l'activité des élèves) sont apparues comme de bonne volonté (de bonne démocratie pour la première; de bonne épistémologie - ou de bonne psychologie - pour la seconde). Mais on notera aussi que tout cela fait l'économie d'un savoir approprié concernant ce jeu qui se joue au sein d'un système didactique - une "économie" chèrement payée quelquefois...

3.10. Pour ne pas allonger, j'arrête là les exemples donnés à titre d'illustration. Et je reviens à la question dont j'étais parti: sur quoi portent nos assertions, et de quoi nos corpus sont-ils faits? Je ne crois pas qu'il soit trop contraignant, ni trop engagé, de répondre: nos assertions et nos corpus se rapportent aux "objets" du schéma qui vient d'être tracé (bien entendu, ce schéma n'est ici qu'à peine esquissé: par exemple il faudrait y préciser la "place" des manuels, etc.). Premier effet de cette réponse: effet de "totalisation". En incluant dans notre champ de question

modernes" amènent "les gens" à se pencher sur le problème de cet échec: Les mathématiciens (dans l'environnement) et, dans la noosphère, ceux qui prennent en charge (pour éventuellement les retraduire de manière appropriée) les points de vue des mathématiciens, examinant ce qui se passe dans cet enseignement moderne, voient qu'on y parle de mathématiques plus qu'on n'en fait, et concluent que les mathématiques, ou plutôt "faire des mathématiques", ça n'est pas ça. Et qui il y aurait dans cette perversion, pour le moins, un facteur de l'échec. Le thème des activités est une réponse adaptée à ce type de pressions (on sait que par ailleurs ces pressions convergent avec des influences qui se nourrissent à d'autres râteliers - au piagetisme "scolarisé" notamment).

nement bien plus que le système didactique stricto sensu (les trois places, du Professeur, de l'Élève, du Savoir), elle prétend arracher - en droit, sinon dans les faits, ce qui est un combat de plus longue haleine - au régime de l'opinion (de la "doxa"), donc du prescriptif, du blâme ou de la raillerie, tout un ensemble de "faits" (concernant l'activité de la noosphère, les opinions et les actions des mathématiciens, des parents d'élèves, etc.) que nous ne recevons ordinairement (je veux dire dans la perspective d'un centrage exclusif sur le système didactique) que dans le registre de l'anecdote, du fait empirique bien connu, trop connu même (de nous), pour que nous songions à l'aborder par un détour théorique, qui assigne à cet ordre de faits un "intérêt" théorique, qui "fasse parler" l'empirie, en lui ouvrant un domaine de pertinence en égard à ce que nous prétendons étudier. Si ce qui se joue dans le système didactique a partie liée avec ce qui se joue dans ses environs, alors il convient de soumettre tout le champ que le schéma désigne comme pertinent à un même ordre d'exigence épistémologique - ce qui ne signifie pas, toute fois, je vais y revenir, que toutes les "régions" à explorer soient justiciables des mêmes méthodes. On peut déjà se convaincre de l'importance décisive de cette extension du champ de visée de la d.d.m. en considérant seulement le système didactique lui-même. Il est encore largement vrai, je crois, que l'on caricature à peine les positions des uns et des autres en disant que l'enseignant ne voit que, dans la triade Professeur - Elève - Savoir, que le troisième terme, le Savoir, qu'il remaniera, éventuellement, à saboter; et que le psychologue est voué à ne voir que l'Élève - et encore: le "Sujet" - , prenant

le savoir pour un donné qui ne ferait pas problème, et le professeur pour un personnage à qui il suffirait de communiquer les résultats de l'analyse du sujet (à explorer) aux prises avec un savoir non questionnable. L'énonciation de la pertinence de la prise en charge de trois termes - et non d'un seul, quel qu'il soit - est à mon avis à mettre, historiquement, à l'actif de la didactique.

3.11. Deuxième effet de la réponse proposée : effet de différenciation.

En traçant des lignes de démarcation entre des régions différentes du champ à explorer, elle marque une première structuration de l'ensemble des éléments qui peuvent venir figurer dans nos corpus par exemple : un ensemble de copies d'élèves et un ensemble de textes publiés par le bureau de l'APM, ça n'est pas seulement empiriquement différent ; leurs positions respectives dans le schéma présenté les différencient a priori comme outputs de sous-systèmes distincts

(entre lesquels existent des interrelations que l'on pourra ensuite s'attacher à élucider), donc comme théoriquement distincts, et entrant de manière a priori distincte dans la théorisation en train de se faire. Quel bénéfice quant au problème de la validation de nos assertions, donc quant à la constitution de nos corpus, peut-on alors retirer de cette différenciation ? Essentiellement, en premier lieu, celui de nous mettre en garde contre l'attitude de dogmatisme méthodologique qui consisterait à imposer à toutes nos assertions, sur quelque région qu'elles portent, d'être comptables de leur vérité devant les mêmes procédures méthodologiques. Pour éclairer d'un exemple, étranger à notre domaine, cette position épistémologique : si je fais de la physique élémentaire et si je prétends avancer pour vraie l'assertion suivante

" Si je tiens dans ma main une pierre, et si je lâche alors cette pierre, elle tombe "

ne me demandez pas de quel "corpus d'observations" je puis nourrir ma certitude ! Je n'irai certainement pas magnétoscooper mille lâchers de pierres, ni même dix, pour complaire à une exigence de pur psittacisme méthodologique - et de pur ridicule. Dans l'histoire de la physique (et non dans quelque histoire rêvée d'une physique fantasmée), nul ne s'est jamais posé cette question. On s'est demandé comment (selon quelles lois) les pierres tombaient (Galilée); et pourquoi elles tombaient (Newton). On ne s'est pas demandé si elles tombaient... Le fait empirique de la chute des pierres n'était pas le problème; le problème était celui de rendre compte théoriquement d'un fait patent - et la réponse amenait à chaque fois un savoir qui n'était pas dans notre connaissance (notre certitude) empirique brute. Par exemple, Galilée nous a appris que les pierres tombent également vite (en première approximation: mais c'est elle-là qui compte) quel que soient leurs poids - phénomène observable, mais non observé, qu'on ne nous avait pas fait observer jusqu'à lui\*, donc qui n'était pas dans notre connaissance de la "chute des graves". Newton ne nous a pas dit que les pierres tombent parce que la Terre les attire (ce qui aurait été nous faire le coup de l'explication par la "vertu dormitive" de l'opium), mais que les masses s'attirent, selon une certaine loi; et donc que la pierre attire la Terre (et tout autant!), chose qui n'est certainement pas dans notre connaissance "spontanée" de la chute des pierres...

\* Si peu en vérité que la physique d'Aristote tenait pour assuré qu'une pierre d'un poids double tombe deux fois plus vite.

[ PREMIER CORPUS, PREMIÈRE THÉORIE ]

3.12. Ces dernières remarques ont pour ambition de montrer d'abord qu'il n'est pas de l'ordre de l'évidence qu'il faille soumettre des choses dissemblables à des traitements identiques : c'est un premier point, qui accomplit l'intention dialectique (de totalisation et de différenciation) du schéma avancé plus haut. En second lieu, elles font apparaître que le problème de l'objectivation de nos connaissances (à propos de notre objet d'étude) ne peut se ramener à une affaire de corpus expérimentaux seulement : la théorie pointe le bout de son nez - nécessairement. Toute objectivation suppose une formulation, dont le contrôle expérimental (en terme de vérification ou de falsification, peu importe ici) puisse faire sa pâture ; soit donc, sinon une théorie, du moins une amorce de théorisation. Mais le contrôle expérimental lui-même suppose plus que cela : en général il ne peut rien dire d'une formulation isolée, d'un isolat théorique, parce que nous ne savons pas alors quelles formes concrètes pourrait prendre un tel "contrôle", quels "faits" et quels "corpus" de faits pourraient autoriser ce contrôle. Permettez-moi ici le recours aux auteurs (simple accès de paresse) : seule une théorie scientifique peut apporter "la résistance organisée d'un corps systématique de concepts et de relations défini autant par la cohérence de ce qu'il exclut que par la cohérence de ce qu'il établit ; seule, elle peut construire le système de faits entre lesquels elle installe une relation systématique ; seule, enfin, elle peut donner à l'expéri-

mentation son plein pouvoir de démenti en lui présentant un corps d'hypothèses si systématique qu'il est tout entier exposé en chacune d'elles"\* . Pour aborder valablement la question des corpus expérimentaux, il faut donc bien se pénétrer de cela : pas de contrôle expérimental qui ne soit contrôle d'un système, assez vaste et organisé, de propositions : "une expérience, notait Durheim, ne peut jamais condamner une hypothèse isolée mais seulement tout un ensemble théorique"\*\*. Inutile de se déchaîner, en termes d'enquêtes, d'observations, etc., pour les beaux yeux d'une formulation non insérée en une théorie significative du domaine auquel elle s'attache : on ne se fait pas martyr de la science si facilement !

3.13. Je note en passant une question que ne manqueront pas d'amener les indications qui précèdent : si la théorie a le pas (le primat) sur le contrôle expérimental, d'où nous vient la "première théorie" ? Est-elle condamnée à s'auto-proclamer sans n'avoir rien (aucun "corpus justificatif", précisément) sur quoi s'appuyer ? - questions qui semblent fort préoccuper certains d'entre nous, et auxquelles je vais revenir en ce qui me concerne. Mais je retourne au corpus, et à sa constitution ; je résume ce qui a été dit jusqu'ici par cette formule : la constitution d'un corpus pertinent est une affaire qui n'est ni aussi simple ni toujours aussi dramatique qu'on le croit quelquefois. Elle n'est pas

\* P. Bourdieu, J.-Cl. Chamboredon, J.-Cl. Passeron, Le métier de sociologue, Mouton, Paris, La Haye, 1973, pp. 88-89. Voilà l'annexe 1, de saine lecture !

\*\* Cité par Bourdieu et al., op. cit., pp. 89-90.

simple, ni évidente, parce qu'il ne s'agit pas seulement de rassembler des faits, mais de déterminer des catégories de faits qui soient, au regard de la théorie, pertinents, et en lesquels quelque chose soit risqué. Or de tels faits ne se rencontrent pas nécessairement à tous les carrefours - la traque est souvent un peu incertaine. Je prends un exemple personnel. J'ai développé (notamment dans Pour la didactique) une théorie du système didactique selon laquelle ce qui distingue l'enseignant de l'élève dans leur rapport au savoir au sein de la relation didactique, ce n'est pas que l'enseignant sait plus que l'élève, mais qu'il sait avant, qu'il peut anticiper (qu'il a la maîtrise du temps du savoir). Comment constituer le corpus de cette théorie ? De différentes manières sans doute (j'en ai indiqué plusieurs dans la référence citée). Mais voici une catégorie d'observations d'une autre nature : une institutrice bavardant avec une mère d'élève lui indique que, dans la classe où se trouve son enfant, deux élèves sont particulièrement remarquables de leurs camarades, qui les appellent "le petit maître" et "la petite maîtresse"; pourquoi ? eh bien, parce qu'ils savent toujours (sic) quelles questions elle va poser\*... Cette seule observation, qu'on chercherait vainement à reproduire de manière systématique, est l'attestation fulgurante d'une certaine "vérité" que la théorie permet d'énoncer à propos du système didactique. Une telle observation, certes, coûte cher, et ne peut guère être produite en série ; mais elle nous introduit au cœur du fonctionnement didactique, qui se dévoile

\* Observation rapportée par M.-A. Joshua (la mère d'élève). :

en sa vérité à l'insu même de ceux qui en sont les agents; elle nous fait participer à l'intimité, à la respiration profonde, de la relation didactique. Il n'est pas douteux pourtant qu'on ne pourra recueillir même un petit nombre d'observations de cette catégorie...

3.14. Dans le même temps, on peut soutenir aussi que la constitution des corpus ne requiert pas également (selon les "régions" auxquelles on s'adresse) la même gravité, la même composition, le même sérieux: pour forcer un peu, je dirai qu'il y a des corpus qu'on peut traiter avec une certaine légèreté, qu'il y a des corpus qu'on peut se limiter à évoquer pour les invoquer, sans nécessairement les faire comparaître, parce qu'une juridiction à la saint Thomas, qui veut voir, et toucher, pour croire, ne s'assurerait pas de plus grande certitude... C'est ici qu'il faut se souvenir des pierres qui tombent, et se moquent de nos corpus. C'est ici qu'il faudrait mobiliser un certain "feeling" épistémologique, celui de notre inégale proximité à notre objet d'étude: il y a des "parties" de cet objet sur lesquelles nous\* avons des connaissances, et pour lesquelles une compatibilité "expérimentale" stricte n'accroîtrait dans un premier temps ni le volume de nos connaissances ni leur qualité (elle satisferait peut-être chez certains d'entre nous un goût particulier pour la "propreté" méthodologique); et puis il y a des zones qui, à l'opposé, sont pour nous terres inconnues, scellées de méconnaissance; et tout d'abord, cette terra incognita par excellence, l'élève...

\* nous, nous qui ne sommes pas des Martiens, nous qui avons été élèves, nous qui avons fait des études de mathématiques, nous qui sommes (souvent) des enseignants (de mathématiques), etc.

Entre ces deux extrêmes, il y a tout un dénivelé le long duquel les conditions de l'objectivation évoluent, et dont l'exploration exige de la souplesse - et non le dogmatisme de l'intégrité - dans l'emploi de la "méthode". Je prends un exemple, vieux truc didactique qui servira bien encore quelques siècles. Soit un didacticien qui est (comme la plupart) en même temps enseignant. Il est un peu l'ethnologue indigène; cela pose des problèmes particuliers, certes, mais pas les mêmes que ceux qui se poseraient à celui qui n'aurait, de sa vie, vu une classe fonctionner (ce qui pourrait être le lot de qui voudrait se vouer à la didactique historique, étudier les systèmes didactiques au Moyen Age par exemple). Le problème majeur, bien sûr, est celui de l'illusion de la transparence, et de la nécessité (et de la difficulté) de rompre avec elle: difficulté d'assurer les conditions de la rupture épistémologique avec la didactique spontanée des enseignants. Mais lorsqu'on s'est assuré (collectivement) une prise sur ces conditions, qu'on sait pouvoir donner congé à une illusion qui pourtant toujours menace, il ne faut pas pour autant jeter le bébé avec l'eau du bain! C'est là qu'il faut, avec de la circonspection, un certain sens du ridicule: n'allons pas "prouver" que les pierres tombent. Notre enseignant-didacticien fera bien de se méfier de ce qu'il voit savoir touchant l'élève: il est particulièrement bien placé pour ne pas voir certaines choses (ce qui appartient, notamment, au "monde caché" de l'élève, ses procédures, son langage en marge du langage officiel, etc.), et il

devra alors, comme didacticien, redoubler d'exigence sur les conditions de l'objectivation d'un savoir à construire (presque) de toutes pièces ; en revanche il pourra déjà, avec un peu de métier (de métier de didacticien), repérer les mouvements (de protestation, de mutisme, d'incertitude, etc.) qui manifestent par exemple une tension au niveau du contrat didactique ; mieux encore, s'il s'agit pour lui de travailler sur cet objet particulier que sont les manuels, ce n'est pas, dans un premier temps du moins, d'une fréquentation accrue de ceux-ci - de ceux-là même qu'il a depuis des années mis à contribution comme enseignant - qu'il pourra attendre un surcroît de savoir - ce n'est pas avec "plus de la même chose" qu'il obtiendra autre chose que ce qu'il a déjà ; et son "premier corpus", il l'aura dans la tête, mieux sans doute que le didacticien non enseignant\*, faisant pesamment son métier, ne pourrait y prétendre.

3. 15. Le premier corpus sur lequel travaille le didacticien est donc un corpus "imaginaire" (au sens où André Malraux, prix Goncourt 1933, parlait de musée imaginaire), qu'il aura rassemblé - ou plutôt, thésaurisé - à son usage. C'est à partir de ce corpus qu'il devra engager une rupture épistémologique, en recadrant ce corpus par le moyen d'une certaine visée théorique. Ce recadrage a pour ambition un recyclage épistémologique, dûment contrôlé, des connaissances dont le didacticien est toujours déjà le dépositaire. Je crois très naïve l'idée qu'il serait possible de

---

\* non enseignant au niveau concerné.

faire l'économie de ce retravail, de faire le vide, de se purger des "connaissances", bonnes ou moins bonnes, par lesquelles notre regard spontané serait habitué, et téléquidé; encore plus irréaliste apparaît l'idée qu'il serait possible de "faire le tri", de conserver le bon, d'écarteler le mauvais. Il ne s'agit pas d'oublier, mais bien plutôt de se ressouvenir (fine allusion platonicienne!) et de travailler ce ressouvenir.

### [ "INTROSPECTION" ET "ANAMNÈSE".. ]

3.16. Il faut en ce point marquer fortement une distinction et prévenir les ravages d'un vieux débat qui pourrait - sait-on jamais? - nous faire écran à de plus difficiles questions: veux parler de l'introspection - car on aura pu, dans les lignes qui précèdent, voir s'en profiter la menace. Ne nous laissons pas égarer. Revenons froidement aux mots, pour les peser à leur juste poids. Qu'est-ce que l'introspection? J'emprunte à quelqu'un qui ne plaisante pas avec la liturgie\*: c'est une méthode, nous dit-il en se référant à ce pauvre Wundt (qui eut le tort d'écrire une Volkerpsychologie en 10 volumes), c'est une méthode "qui consiste à demander à l'individu lui-même un rapport sur ce qu'il pense, ce qu'il ressent, en un mot sur ses états subjectifs". On voit tout de suite que l'introspection n'a pas forcément à voir avec le ressouvenir (le résultat de l'introspection peut être livré immédiatement); mais surtout, on voit que le caractère spécifique de l'introspection est cons-

\* Maurice Reuchlin, dans son Histoire de la psychologie (PUF, Paris, 1978), p. 18.

titulé par l'objet auquel elle nous fait accéder : elle porte sur des "états de la conscience", elle prétend nous entretenir de la "vie" intrapsychique. Cela n'est pas en soi un mal, mais bute sur une exigence épistémologique fondamentale : comment objectiver ce que je suis le seul à connaître, etc. ?\* . Or si le souvenir peut bien en effet porter sur des états de conscience, et tombe alors sous le coup de la critique, (qui n'est toutefois pas absolue) que l'on vient de rappeler, en général il n'en est rien : si je me demande si j'ai déjà vu tel phénomène se produire dans une classe, par exemple si j'ai déjà vu des élèves écrire (au tableau, mettons) que  $0 \times 4$  égale 4, je ne fais nullement de l'introspection! (tout de même!). En fait il y a un mot pour désigner ce que je fais alors, c'est un mot bien banal chez les médecins et les psychiatres, c'est "anamnèse" (du grec ana, remontée, et mnémè, souvenir : voir l'annexe 2). On m'accordera que le fait de ramener toute anamnèse à de l'"introspection," c'est pour le moins - comment dirai-je ? - manquer de rigueur...

3.17. Bien entendu, un mot ne fait pas de miracle. Le problème du contrôle de l'"anamnèse" reste entier. Simplement, dans la majorité des cas (en particulier lorsque l'objet de l'anamnèse est fait de situations objectives, et non d'événements intrapsychiques), le problème du contrôle des "résultats" du souvenir peut être résolu assez facilement. Nous arrivons là à une question qui me paraît essentielle, parce que précisément la réponse qu'on peut lui apporter n'est pas une réponse méthodologique générale, en cela qu'elle met à profit le caractère

\* Voir Reuchlin, op. cit., pp. 22-23.

particulier de l'objet d'étude du didacticien (sans être pour autant toutefois, entièrement spécifique de la d.d.m., tel un bibelot unique et singulier). En didactique en effet, nous avons souvent affaire à des situations objectives observables et répétées (non pas forcément répétables : mais qui se répètent, comme les éclipses de lune...). Prenons le cas alors de situations observables par l'enseignant en tant que tel, telle celle que j'ai notée un peu plus haut (il y a des situations observables, qui ne sont pas observables par l'enseignant-en-train-d'enseigner : tout ce qui se rapporte au "monde caché" de l'élève notamment). Si je demande à un enseignant s'il a déjà observé tel phénomène didactique, et s'il me répond positivement, j'obtiens déjà une information ; il s'agit d'anamnèse, de l'"état zéro" du souvenir. Après cela je peux essayer de passer à une phase plus avancée de la constitution de mon corpus (pour le moment, mon "corpus" est entièrement dans la tête de l'enseignant auquel je me suis adressé - ou des enseignants, s'il y en a plusieurs) ; mais souvent le phénomène auquel je m'intéresse n'est pas, bien qu'on puisse l'observer de manière répétée, reproductible (ou du moins reproductible de manière suffisamment approchée), et je veux l'observer "in vivo". On court donc le risque, à ne pas introduire entre simple anamnèse et expérimentation un niveau intermédiaire fait d'une methodologie appropriée, de laisser de côté une grande partie, voire la plus grande partie des faits didactiques intéressants - on court le risque de chercher, comme le raconte un apologue méthodologique bien connu, la clé sous le rebord, non parce

\* C'est le cas si, en météorologie, je veux étudier la formation des orages par exemple (il n'est pas inutile d'aller chercher l'astronomie, la géologie, etc.).

qu'on l'aurait perdue là, mais parce que c'est le seul endroit éclairé ! Par rapport à la catégorie de phénomènes que j'ai indiquée, la stratégie suivante est possible pourtant : demander à l'enseignant (à soi-même le cas échéant !) de noter aussi précisément que possible le phénomène attendu dès lors qu'il se reproduira (note réalisée sous forme d'une "fiche d'observation" que l'on peut légèrement standardiser, etc.) ; rendre compte assez rapidement de l'observation faite afin qu'éventuellement des questions puissent être posées qui rencontrent une mémoire encore fraîche de l'événement ; et (le cas échéant) enrichir la fiche d'observation des remarques et additifs apportés par ce débat-compte rendu. Quel est l'avantage de cette procédure ?

A mon avis, il est triple. En premier lieu, on s'assure qu'il n'y a pas eu de malentendu sur le type de phénomène mentionné au cours de la première évocation : et cela par le fait que l'enseignant-observateur rapporte du matériel confirmatif. Ensuite, on détient une attestation effective du phénomène ; enfin, on crée les conditions pour qu'une telle observation soit répétée, et banalisée (on crée les conditions de variation - et d'extension - du corpus).

3.18. L'anamnèse est donc pour le didacticien, bien souvent (non pas toujours : je vais y revenir), le point de départ, l'ancrage de la constitution du corpus. Ensuite, il faut continuer par d'autres voies. J'ai simplement voulu souligner qu'il n'y avait pas que les voies royales, les autoroutes méthodologiques, où il suffirait de s'acquitter régulièrement du péage pour progresser sans embûches. Je crois qu'il est bon d'être attentifs au thème des "méthodologies"

intermédiaires", appropriées à une classe de faits didactiques éventuellement rebelles aux méthodes standard. Je vois qu'une partie essentielle de l'"identité scientifique" de la didactique, de sa scientificité et de son identité propre dans l'univers des sciences, va se jouer là ; et que c'est là que nous aurons beaucoup à gagner, et donc, ce qui a bien no plaisir, beaucoup à perdre aussi.

### [ LE MOUVEMENT THÉORIQUE-EXPÉRIMENTAL ]

3.19. Je prends un exemple pour illustrer une partie des thèses précédentes, et faire le lien avec la suite de mon propos. Il s'agit d'un travail exemplaire par sa banalité méthodologique\*, et très instructif (pour nous) parce que légèrement décalé par rapport à nous, par rapport à la didactique des mathématiques. Je veux parler du travail d'André Cauty intitulé Reflexions sur le langage mathématique, présenté pour l'obtention du DEA de linguistique en juin 1981. Je me référerai ici à la première partie du mémoire ("Généralités sur le langage mathématique") pour l'essentiel. L'auteur se présente normalement comme chercheur se livrant à une "recherche linguistique de terrain". Dans cette catégorie il aurait pu par exemple choisir de travailler (j'imagine) sur le vocabulaire des jeux d'enfants (simple exemple). Comme il n'est plus un enfant, et que de toute façon le vocabulaire en question évolue, il aurait dû alors

\* On aura compris que le mot de banalité n'est ici nullement péjoratif : au contraire !

apprendre à connaître son "terrain", d'abord par l'observation nue (voie "participante": les enfants, les neveux, etc., sont là pour ça), puis en se constituant un corpus par enregistrement; etc. Mais voilà: le terrain qu'il a choisi c'est le "langage mathématique", tout au moins (il s'en explique) celui qui se rencontre dans les manuels. Va-t-il agir de même que s'il étudiait le vocabulaire des jeux d'enfants? Nullement! Il est enseignant de mathématiques, il a fait des études de mathématiques! Cette particularité est ici pleinement pertinente: il en fait mention. C'est un bénéfice, une sorte de situation, et non une tare (comme nous serions presque tentés de le croire, aujourd'hui, en didactique des mathématiques). Il connaît ce terrain mieux qu'aucun autre terrain\*. Son corpus imaginaire, à cet égard, est immense... Le vrai problème pour lui est de faire fonctionner une certaine théorie (ou une certaine orientation théorique) sur cet objet qu'il connaît - empiriquement - bien mieux que Lévi-Strauss ne connaissait les Bororo... C'est le travail théorique (linguistique) sur l'objet, ou plutôt le travail théorique de l'objet qui est ici l'axe directeur. Bien entendu, pour borner son champ, il désigne explicitement, dans ses références, un sous-corpus fini (comme opposé à "infini") comprenant un très petit nombre de titres (11 en tout) par rapport à son corpus "dans la tête". Pourtant, c'est bien me semble-t-il par rapport à son corpus imaginaire qu'il va définir les points d'ancrage de

\* A ceci près (à ma connaissance) qu'André Cauby a beaucoup travaillé sur les langues amérindiennes - un autre terrain.

son travail théorique; c'est grâce à un "connaître" non explicité qu'il pourra écrire: " Sans entrer dans le détail d'une détermination très fine\*, nous pensons pouvoir\*\* retenir trois grandes genres mathématiques (la définition, l'énoncé de théorème et la démonstration de théorème) ainsi que deux genres liés au didactisme (le commentaire, et l'exercice)". Le corpus matériel qu'il a pu effectivement manipuler pour son travail (les manuels, en petit nombre, qu'il a dû effectivement feuilleter, parcourir, etc.) n'a nullement pour objet de donner la "base la plus large" à la théorisation, mais seulement de soutenir la théorisation, de donner un appui au travail théorique; ses références ne sont pas là pour "fonder" sa théorisation, mais pour permettre (pour guider, contrôler) le mouvement de théorisation, la production d'un certain savoir. En voulez-vous un exemple? Au cours de ce mouvement, loin de chercher à élargir son corpus, il est au contraire amené à le restreindre. Et il s'en explique: " L'inventaire que nous proposons, écrit-il, ne prétend pas à l'exhaustivité. Pour ne pas favoriser systématiquement les verbes étroitement liés à une branche mathématique particulière (compactifier, intégrer, estimer, plonger...) nous avons décidé de limiter le corpus à des énoncés de baccalauréat (série Terminales C) portant sur la question géométrique du barycentre. Les résultats obtenus ont ensuite été vérifiés sur des exercices de manuels".

3.20. Nous pouvons partir de ce passage pour avancer. Tout d'abord, il convient de noter que la méthodologie réelle mise en œuvre par le chercheur doit constamment se défendre contre l'interpellation

\* Cette formule d'excuse - en regard aux prétendues bonnes manières méthodologiques - est l'expression (impensé?) du coup d'état par lequel s'inaugure nécessairement la théorisation - la première prise théorique est un coup de force.

\*\* Même remarque.

d'une méthodologie fantasmée, à la fois méthodologie officielle, qui hante le chercheur comme norme légitime, et méthodologie d'opérette, inopérante et dérisoire - surmoi du chercheur. Car pourquoi s'exposer de la "non exhaustivité" de son inventaire? Qui exige cette exhaustivité? Et que serait-elle? La précision apparaît comme une réponse à une question qui n'a pas été posée, et qui n'est pas formulable - une fiction méthodologique, mais une fiction nuisible, qui établirait facilement le chercheur dans une mentalité d'assrégé. Je reviendrai plus loin sur le fantasme du "corpus complet" - fantasme d'empiriste s'il en est! Mais reprenons la question initialement posée, celle de la validation de nos assertions, et du rôle des corpus expérimentaux. D'une part on a vu que les corpus expérimentaux ne doivent pas, en regard au processus de validation, être considérés comme pierres de touche en eux-mêmes (cf. les points 3.5 à 3.12.); ils n'ont de sens, de pertinence d'intérêt qu'au sein d'un ensemble théorique-expérimental, plus justement au sein d'un mouvement théorique-expérimental (qui s'enclanche solidairement avec la "première prise théorique" et le "premier corpus"). Et cela a des implications essentielles pour la constitution des corpus, question qu'on ne peut prétendre traiter in abstracto, mais qu'il faut au contraire aborder en fonction des caractères propres de notre objet d'étude (cf. les points 3.13. à 3.19.). D'autre part, et c'est cela que nous apporte de neuf l'exemple examiné en 3.19., il est particulièrement réducteur de ne voir dans les corpus que le moyen de la garantie expérimentale. Cette réduction - par rapport à la réalité du travail scientifique - est le propre d'une obsession des "fondements"

de la connaissance scientifique. Elle ne permet pas de saisir autrement qu'à titre de paradoxes ou d'approximations - d'imperfections dues à la faiblesse de nos cerveaux humains - certaines pratiques pourtant essentielles et constantes dans le travail scientifique, telle la "restriction" des corpus. Les corpus garantissent la théorisation en ceci d'abord qu'ils la rendent possible, qu'ils en sont l'objet même : d'un même mouvement ils soutiennent la production du savoir et le processus de validation du savoir - sans épuiser le débat.

### [ LE MOT MÊME DE CORPUS, SON ORIGINE ]

3.21. A la lumière des remarques faites jusqu'ici, il faut tout de même en venir à l'examen de la notion de corpus - et du mot. Cet examen est nécessaire parce que, là encore, nous courons le risque de la fausse innocence. Car le mot de corpus, qui nous vient de la linguistique, n'y est pas, dans ses usages premiers et spécifiques, synonyme d'"ensemble de données" tout court ! Il se trouve historiquement lié à une méthode particulière, une méthode qui ne va pas de soi, dont on peut au moins dire qu'elle est très exigeante et, comme le notera Chomsky, très (trop ?) ambitieuse - à prétendre promouvoir un certain angélisme méthodologique\*. Je reproduis tout d'abord une définition proposée de "corpus" : " Ensemble homogène et significatif de données linguistiques observées et à partir desquelles pourra s'élaborer la théorie " (voir l'annexe 3). Quant à

\* qui, à s'affirmer comme règle, outrepassant le simple fait de recherches spécifiques, devient une véritable perversion épistémologique.

la méthode à laquelle la notion se trouve liée, c'est la méthode dite distributionnelle d'analyse structurale en syntaxe (dont le père est Z. S. Harris). Dans cette approche, le "corpus" n'est pas là pour seulement "soutenir" la théorisation; la théorie doit procéder toute du corpus lui-même, d'où on doit la faire surgir par une méthode (le distributionnalisme) qui permettra de "faire parler" le corpus: tout le corpus, mais rien que le corpus! (cf. l'annexe 4, sur l'analyse distributionnelle). Dans les débats que nous aurons à Orléans à propos des corpus expérimentaux, nous ne devons pas oublier cet emploi originare du mot, qui l'associe à une méthode dont la pertinence ne peut être établie qu'en rapport avec un certain type d'objets d'étude, et ne saurait de toute façon se réclamer d'une exigence épistémologique a priori. J'insiste sur ce point parce qu'il me semble que l'"ascèse" du "corpus - et - rien - que - le corpus" peut, dans la perspective positiviste-empiriste si spontanément répandue, se trouver facilement investie d'une réputation d'hyper-scientificité fascinante et d'une aura de pureté méthodologique, à la fois âge d'or et terre promise de tout travail scientifique. Pour que les choses soient bien concrètement précisées, je mets en annexe 5 à cette note une présentation critique (due à G. Mounin) d'une analyse distributionnelle en sémantique: plus que l'intérêt et les limites de la méthode mise en oeuvre, on y saisira le caractère de moyen propre à une méthode particulière d'analyse qui, dans ce cadre, est attribué au corpus.

3.22. La critique fondamentale a été portée par Chomsky. Un résumé en est donné dans l'annexe 4, auquel, cher lecteur, vous pouvez vous reporter. La parenthèse distributionnaliste explique assez bien, je crois, l'illusion d'exacte superposition, à laquelle nous serions tentés d'ajouter foi, des deux fonctions des corpus que nous avons recensés : dans la méthode distributionnelle, le corpus est du même coup ce qui permet de produire le savoir et ce qui garantit le savoir produit. L'analyse présentée plus haut substituée à cette ambition du "d'un seul coup, d'un seul" la réalité d'un processus théorique-expérimental de production/validation.

### [ CORPUS ET EMPIRISME ]

3.23. Il est à mon avis une situation dans laquelle la confusion que je viens d'indiquer (à quoi il faudrait substituer une dialectique de rôles) est particulièrement sujette à se manifester, tout en manifestant les racines empiristes auxquelles elle se rattache. C'est la situation où le corpus se rapporte à une "population parente" par nature finie. Un exemple-type, pour nous, est celui des manuels : si l'on se réfère à la population "ensemble des manuels de 4<sup>e</sup> publiés correspondant au programme de 1971", la population en question est par nature finie. On peut donc ambitionner d'en constituer le corpus complet - chose qui, en l'espèce, ne doit pas être très difficile. Ce type de situation apparaît comme

la matrice d'un fantasme que je nommerai le fantasme du corpus complet. La poursuite (empiriquement possible dans certains cas particuliers) de corpus terminés entraîne bientôt à une quête indéfinie de corpus interminables. Monsieur X a examiné 20 copies (ou a interviewé 20 sujets), c'est un exercice "exploratoire"; trois ans plus tard, il en est à 2000. Expérimentateur et martyr, on ne pourra lui refuser sa place au panthéon de la didactique. Au delà des mystifications que l'on peut craindre (et dont notre communauté, à l'étape actuelle de son histoire, souffrirait cruellement), il faut analyser l'ornière épistémologique qui, par là, se creuse. Tout d'abord le fantasme, ici, participe à l'évidence d'un mécanisme de défense, mis en place contre l'effroi suscité par l'incertain, et par l'incertitude dans l'énonciation de ce qu'on voudrait pouvoir, au bout du compte (du compte de manuels, de copies, de sujets...), présenter enfin comme la "vérité" (voir ma Note 2). Or, le corpus complet se présente, dans une perspective empiriste, comme l'arme défensive absolue : à la fois origine de tout savoir, et pièce de touche de tout énoncé proposé ! Je ne reviendrai pas sur le second point (le caractère d'expérience "cruciale" attribué à la mise en rapport avec "le" corpus, complet ou non d'ailleurs), puisque j'ai essayé de souligner plus haut l'insuffisance de la conception à laquelle il se réfère. Pour aborder le premier point, maintenant, je poserai ceci : dénoncer le mythe du corpus-origine-du-savoir, c'est en même temps réaffirmer le primat de la théorie et affirmer qu'un corpus n'est jamais complet, au regard de la construction du savoir, même s'il est

empiriquement complet parce que tiré sans reste d'une population empirique par nature finie ! Proposition provocatrice ? Nullement. Proposition des plus banales en ce qui concerne l'activité scientifique. Mais je vais lui donner une forme plus provocante encore : même si vous avez réuni tous les manuels de 4<sup>e</sup> correspondant au programme de 1971, vous n'avez pas un corpus complet, parce qu'il vous manque les manuels possibles, ceux qui n'ont pas été publiés, ceux même qui n'ont pas été écrits — mais qui auraient pu être écrits, et publiés, et diffusés. Bien entendu la question à poser est ici : au regard de quoi "auraient-ils pu" ... ? C'est indiquer la béance dont l'existence se trouve déniée par les conceptions positivistes de l'activité scientifique. Le corpus ne fait pas savoir ; et en un sens excède tout savoir — dans sa matérialité dont aucune théorisation ne peut rendre compte sans reste. L'approche scientifique veut les choses autrement : derrière les éléments matériels effectivement produits, attestés, recueillis, "connus", elle vise à déterminer les principes de production et de structuration de ces éléments, soit le schème générateur (ou, comme P. Bourdieu n'hésite pas à le dire, la "grammaire générative"\*) qui explique ce corpus en même temps qu'il explicite les éléments possibles de tout corpus de son espèce.

\* Bourdieu dit en fait "grammaire génératrice" : voir sa postface au livre d'E. Panofsky, Architecture gothique et pensée scolastique (Minauit, Paris, 1967), p. 152.

[ LA NOTION DE SCHÈME GÉNÉRATEUR ET  
L'EXEMPLE DU SCHÈME LAPLACIEN ]

3. 24. J'ai essayé de donner de cela un exemple - celui de la classification de Mendeleev - dans Pour la didactique. Mais je voudrais dire quelques mots d'une situation bien plus proche de nous - celle des corpus "statistiques".\* Un corpus n'est pas de droit justiciable d'un traitement statistique; ce n'est pas sans que certaines conditions - touchant à son mode de constitution - soient satisfaites qu'on peut légitimement le considérer du point de vue de la statistique. Sans vouloir de trop demander, on demande au moins que le corpus présente une certaine homogénéité, que ses éléments soient réunis selon certains critères, certes non entièrement déterminants, mais délimitant une certaine communauté opératoire. C'est là, me semble-t-il, le degré zéro de l'exigence statistique. Toutefois, la statistique classique (i.e. antérieure à l'analyse de données) est beaucoup plus exigeante quant aux conditions de son application aux corpus expérimentaux (ou d'observation). Je crois qu'il faut insister sur ce point, parce que nous sommes devenus - le temps et l'oubli aidant - très laxistes à cet égard. Or, il est remarquable que l'exigence posée par la statistique - telle qu'elle s'élabore sous la plume de Laplace, et s'étend aux sciences du social à partir de Quetelet - est une exigence formulée en termes de schème

\* Dans ce qui suit, il s'agit (sauf mention expresse du contraire) des corpus rapportés à des populations parentes supposées répondre à la loi normale; il est clair qu'alors un tel corpus ne saurait être complet (on n'enregistre pas d'individus de 5 m de taille); il est non moins clair que la théorie dans laquelle est pris le corpus justifie cette incomplétude.

générateur - précisément ! Pour qu'un corpus puisse être envisagé comme corpus statistique on doit pouvoir postuler à son origine un certain type de mode d'engendrement. C'est une

question que l'on peut aborder - pour déconstruire quelques uns de nos "réflexes" les mieux installés mais non les mieux fondés - par une interrogation sur la taille (l'effectif) de nos corpus : à partir de quelle(s) valeur(s) de  $n$  voudrait-on bien considérer le lot d'éléments recueillis comme justiciable de la perspective et de l'analyse statistiques ? A partir de 1000 ? De 100 ? De 20 ? De 10 ? De 5 ? De 2 ? Je suppose que tout le monde (ou presque) sera d'accord pour répondre : de 1000, certainement ; de 2, vous plaisantez ! Entre les deux, cela dépend. On place la cuisse, nul ne le sait exactement ; mais chacun d'en craindre le défaut invalidant. Nous savons seulement que la barre numérique, autrefois située assez haut (quelques centaines, au moins plusieurs dizaines : les "grands nombres"), a pu ensuite (à partir de Student et de Fisher) être abaissée (quelques unités, la dizaine). Nous savons aussi, à vrai dire, que l'occasion, en la matière, fait souvent le laron : nous savons les jugements de circonstance - et que 20 paires de jumeaux homozygotes valent amplement 2000 copies d'élèves de sixième...

3.25. Dans la vie quotidienne, nous constituons fréquemment des corpus de taille  $n=1$ , qui sont pratiquement et théoriquement fort satisfaisants. J'ai devant moi, sur mon bureau, deux chemises : l'une contient des copies provenant d'une classe de sixième, l'autre, des copies venant d'une classe de cinquième. Ce sont ces

dernières que je veux consulter. Je me saisis d'une des chemises, je l'ouvre et considère la première copie venue: Magali Dupont, 6<sup>e</sup> 3, etc. C'est donc l'autre qu'il me faut! Faisons un peu dans la custrerie: je viens de tester une hypothèse ("le lot de copies dont je me suis saisi, c'est celui de la classe de cinquième") et de la rejeter; et j'ai pour cela constitué un échantillon de taille 1. Bien entendu, le succès de cette manœuvre est lié à la nature particulière de la loi régissant le caractère examiné (il y a constance!), cette loi étant l'expression observable du schéma générateur qui définit la population à laquelle je me réfère (les copies proviennent d'une même classe). Cette situation d'apparence singulière est pourtant une version spécifiée du cas général: le nombre  $n$  d'éléments qu'il convient de regarder afin de parvenir à quelque conclusion que ce soit dépend de la loi du caractère examiné, qui résulte elle-même du schéma générateur définissant la population où l'on puise les observations. Simplement, le seul cas où l'on soit assuré qu'un échantillon d'effectif 1 suffit est le cas où il y a constance du caractère étudié. Cela est assurément fort banal, mais ce n'est pas là que je voulais arriver: c'est de là que je voudrais partir. S'il n'y a pas constance (présence ou absence, pour tous les éléments de la population), un échantillon d'effectif 1 n'est pas très informatif, certes (encore que, bien souvent en pratique, il permette d'atteindre un ordre de grandeur). Mais sur quoi fondons-nous cette "idée" qu'un échantillon d'effectif plus grand (2, 5, 10, 100, ...) nous apporterait de plus précieux renseignements? Sur une évidence de la Raison? Alors cette Raison n'existe pas depuis très longtemps - je rappelle quelques faits.

3.26. Le développement de l'astronomie au XVI<sup>e</sup> puis au XVII<sup>e</sup> siècles ("l'âge d'or" de l'astronomie) a amené à la considération du problème suivant : lorsqu'on dispose, à propos d'une certaine grandeur (un angle, etc.), de plusieurs mesures différentes - comme c'est toujours le cas pour les mesures astronomiques ou géodésiques - quelle mesure doit-on, en définitive, adopter ? En fait, assez tôt, l'usage se répand de prendre pour mesure la moyenne arithmétique des observations effectuées : c'est déjà ainsi que, dans son Palais d'Uranie, procède Tycho Brahé, à la fin du XVI<sup>e</sup> siècle. Pourtant cette façon toute pratique de répondre à une difficulté pratique soulève des oppositions : une opinion contraire se fait entendre, qui voudrait qu'on ne considérât que le résultat d'une seule observation - mais d'une observation soignée... La polémique où s'engagèrent alors quelques-uns des meilleurs mathématiciens du temps (Simpson, Lagrange, Daniel Bernoulli) montre au moins que l'argument n'était pas sans portée ; et qu'il n'allait pas entièrement de soi que n observations valent mieux qu'une ! Mais ce qui est encore plus intéressant, c'est d'observer comment les partisans du "n mesures valent mieux qu'une seule, même soignée" vont parvenus, en 50 ans de travail, à fonder leur précepte méthodologique : ils n'y parviennent qu'en produisant un schéma générateur (causal) de la population des mesures d'une grandeur donnée. Je résume. Simpson, dans un ouvrage publié en 1757, entreprend "An Attempt to shew the Advantage arising

by Taking the Mean of a Number of Observations, in Practical Astronomy". Il suppose - sans pouvoir avancer de mécanisme causal qui expliquerait le fait - que les valeurs possibles de l'erreur (i.e. l'écart entre l'observation et la valeur vraie) sont

$$-v, \dots, -3, -2, -1, 0, 1, 2, 3, \dots, v.$$

La seule chose qui lui paraisse évidente, c'est que des erreurs de même valeur absolue, et de signes opposés, ont la même probabilité de se produire; quant aux valeurs des probabilités, il hésite. Sa première hypothèse - selon un réflexe d'époque, qui est encore, de nos jours, celui du profane - est de considérer toutes les valeurs comme équiprobables. Puis, se ravisant, il propose comme plus vraisemblable une distribution "en triangle isocèle" (les probabilités étant supposées proportionnelles à

$1, \dots, v-2, v-1, v, v+1, v+2, \dots, 1$ ). A partir de là, il lui est possible de calculer les probabilités  $P\left(\frac{\sum e_i}{n} = \frac{m}{n}\right)$  et  $P\left(\frac{\sum e_i}{n} \leq \frac{m}{n}\right)$ , et de soutenir son précepte pratique d'une description mathématique précise. Le problème devient un problème majeur pour les mathématiciens probabilistes (Daniel Bernoulli lui consacre par exemple 8 mémoires; sans parler des travaux de Lagrange, etc.). Mais l'éclucidation totale ne sera atteinte qu'avec les travaux de Laplace, consignés dans sa Théorie des probabilités de 1812: Laplace fournit un schéma générateur vraisemblable et bien ajusté à l'expérience.

Je cite les auteurs: "Si l'on réfléchit sur l'origine et la génération des erreurs de mesure, écrit l'un d'eux\*, on peut consi-

\* En l'occurrence G. Darrois, in "Histoire générale des sciences", 3, La science contemporaine, 1, Le XIX<sup>e</sup> siècle, Paris, PUF, 1969, p.81.

dérer, et ce fut là une grande idée de Laplace, qu'une mesure peut se rattacher aux phénomènes collectifs aléatoires. En effet, la grandeur que nous considérons, mesure topographique ou géodésique, rayonnement total d'un astre, etc., est à chacune des mesures, soumise à un processus très complexe de déformation, de sorte qu'en réalité chaque lecture travaille sur un être appartenant à une population d'images possibles. Sans doute ces images sont-elles, disons, assez voisines, mais dans un processus de fabrication industrielle, elles sont nécessairement affectées de différences légères, ce que Laplace appelle les erreurs élémentaires. Et c'est ici qu'apparaît le génie de Laplace, et, malgré tout ce qu'on sait de cet autre immense génie de Gauss, une nette supériorité sur ce point. D'abord une conception claire d'un mécanisme concret possible. Et, de plus, Laplace a osé penser que cette superposition d'erreurs élémentaires peut, sous des conditions raisonnables, aboutir à une loi universelle valable pour le résultat de l'addition d'un grand nombre d'aléatoires".

3. 27. Le schéma générateur laplacien est donc le suivant : l'erreur qui affecte la mesure réalisée est la résultante d'un grand nombre d'erreurs très petites et indépendantes entre elles. C'est un schéma qui paraît adapté à un certain nombre de phénomènes : messages divers, tir à la cible, etc. Bien entendu, l'usage du formalisme statistique classique ("paramétrique") ne se justifie que par rapport à des populations dont on puisse supposer, sans trop d'in vraisemblance, qu'elles

procèdent d'un schème générateur de type Laplace. Quoiqu'on puisse croire aujourd'hui, les meilleurs parmi les utilisateurs de la statistique, au XIX<sup>e</sup> siècle, mettent explicitement ce schème générateur à l'origine de leurs corpus expérimentaux. Il en est ainsi, par exemple, de celui à qui nous devons tant de mauvaises habitudes dans le domaine des sciences de l'homme : Adolphe Quetelet. Celui qui n'est pas encore le créateur de la théorie de l'"homme moyen", qui n'est encore qu'un jeune et brillant géomètre belge, vient à Paris en 1823, et y tombe amoureux de la doctrine de Laplace. Rentré à Bruxelles, il a trouvé sa vocation - l'étude des êtres humains, fondée sur la statistique, et sur les statistiques ; ce qu'il nommera la "physique sociale". Le projet de Quetelet a au moins le mérite d'une parfaite cohérence\* : s'il prétend appliquer à l'humain la théorie des erreurs, il n'entreprend de le faire qu'appuyé sur cette hypothèse essentielle que la théorie en question est dûment applicable au fait humain, que l'homme - tel individu particulier - est (n'est que) la résultante de la superposition d'un grand nombre de facteurs, de portée limitée pris un à un, agissant indépendamment les uns des autres. Il est intéressant de voir comment M. Reuchlin, dans un livre déjà cité, rend compte de l'affaire : "...le mathématicien belge A. Quetelet (1796-1874) avait découvert (1835) que des ensembles de données rassemblées sur des groupes importants d'hommes (en particulier les tailles des soldats de la Grande Armée napoléonienne) se répartissaient selon une loi que les mathématiciens avaient déjà

\* Cohérent, même s'il apparaît plus tard comme non pertinent en regard au réel dont il prétend rendre compte théoriquement.

étudiée (loi de Laplace-Gauss) et que l'on avait retrouvée dans l'étude des erreurs de mesure : beaucoup de soldats avaient une taille moyenne et l'on en rencontrait de moins en moins lorsqu'on s'éloignait de la moyenne vers les tailles très grandes ou très petites. Dans chacun des deux sens, la déviance du nombre d'hommes n'était pas égale d'une taille à l'autre : elle était d'abord rapide, puis lente, obéissant précisément à la loi mathématique étudiée par Laplace et Gauss. On obtient des résultats de même forme lorsqu'on mesure plusieurs fois le même objet : on fait souvent des erreurs accidentelles de mesure en plus ou en moins, et les erreurs faibles sont beaucoup plus fréquentes que les erreurs importantes. Le rapprochement de ces deux ordres d'observations conduisaient à des hypothèses métaphysiques contre les séductions desquelles Quetelet se défendit mal : la Nature tenterait, à chaque naissance, de réaliser le même "Homme moyen" et ce serait l'effet de "causes accidentelles" qui donnerait naissance à la diversité des individus vivants".\* Comme on le voit, Reuchlin fait reproche à Quetelet, précisément, de ce pour quoi il conviendrait de le louer : la cohérence très ferme de sa méthodologie - chapitre sur lequel nous sommes en réalité devenus bien moins exigeants que les pères fondateurs, et où une méthodologie à la va-vite, la même pour tout et pour tous, a eu tôt fait de "régler" le débat scientifique en l'enterrant\*\*.

\* Op. cit., pp. 45-46.

\*\* Je crois que l'un des facteurs de dégénérescence de ce débat est le suivant pour prendre part au débat, en l'espèce, une réelle culture mathématique et scientifique était nécessaire ; on sait quel degré de démathématisation a

3.28. Il me semble que la leçon à tirer est ici la suivante : il n'y a eu, pour les promoteurs de la statistique classique, de corpus "statistiques" que justifiés par l'hypothèse d'un schéma générateur spécifique. Par une extension induite de la pratique de la "statustisation" cette exigence, en laquelle ~~se maintenaient~~ ~~pourtant~~ les diverses pratiques statistiques, a été perdue de vue. Que l'exigence elle-même ait changé de contenu est une chose (je pense au développement de la statistique non paramétrique, notamment); qu'on ait pu croire qu'il était possible d'inférer à partir de là qu'on pourrait jeter par-dessus <sup>board</sup> cette exigence elle-même est une tout autre chose. En particulier, il faut noter ceci : les récents "succès" (scientifiques et mercantiles) de l'analyse de données ont suscité un discours euphorique (tenu par au moins un certain nombre de ses promoteurs) selon lequel nous serions aujourd'hui enfin dispensés de toute hypothèse sous-jacente sur les mécanismes de production de la variation soumise à l'analyse - cela constituant un progrès libérateur éminent par rapport aux exigences archaïques de la statistique laflouienne ! Bien entendu, il n'en est rien : l'analyse de données, bien qu'apportant sans doute un énorme surcroît de puissance par rapport à la statistique descriptive à l'ancienne (les moyennes de Tycho Brahe), n'en demeure pas moins de la statistique descriptive; et nous en sommes vis-à-vis de son utilisation au stade où en étaient au XVIII<sup>e</sup> siècle astronomes et géomètres : nous n'avons pas de moyen de poser théoriquement le problème de sa pertinence et de son applicabilité aux corpus que nous voudrions lui soumettre. Ce qui ne signifie pas qu'il faille l'écarter, mais ce

---

atteint l'outil statistique entre les mains des utilisateurs "ordinaires" - des psychologues tels qu'ils sont formés par nos maîtrises de psychologie par exemple.

qui ne saurait justifier, à rebours, un discours triomphaliste. L'un des mérites de l'analyse de données a été de familiariser les communautés scientifiques <sup>avec l'idée</sup> qu'il fallait décidément renoncer à l'hypothèse Laplaceenne pour une quantité de phénomènes pour lesquels cette hypothèse était malgré tout maintenue parce qu'elle constituait jusqu'alors notre seul recours...

### [ L'ENCLANCHEMENT DU PROCESSUS THÉORIQUE-EXPÉRIMENTAL ]

3.29. J'avais annoncé (au point 3.13.) une remarque à propos de la première théorisation. L'idéologie régressive usuelle, qui prétend fonder une connaissance sur une connaissance antérieure, trouve là son impasse : entre "premier corpus" et "première théorisation", il y a la même relation qu'entre la poule et l'œuf. Il y a bien primat de la théorie, mais non antériorité de la théorie sur le corpus, non plus que du corpus sur la théorie. Ce qui "fonde" la théorie, ce n'est pas son passé : c'est son avenir. Première théorisation et premier corpus entrent ensemble dans le discours scientifique, pour permettre l'enclanchement du processus théorique - expérimental. On doit se contenter de noter leur surgissement, et d'avancer à partir de là.

Annexes

## Le primat de la théorie

### 3.2. *Système de propositions et vérification systématique*

Si les opérations de la pratique valent ce que vaut la théorie qui les fonde, c'est que la théorie doit sa position dans la hiérarchie des opérations au fait qu'elle actualise le primat épistémologique de la raison sur l'expérience. Il n'est donc pas étonnant qu'elle constitue la condition fondamentale de la rupture, de la construction et de l'expérimentation, et cela par la vertu de la systématisme qui la définit : seule une théorie scientifique peut opposer aux sollicitations de la sociologie spontanée et aux fausses systématisations de l'idéologie, la résistance organisée d'un corps systématique de concepts et de relations défini autant par la cohérence de ce qu'il exclut que par la cohérence de ce qu'il établit<sup>1</sup>; seule, elle peut construire le système de faits entre lesquels elle instaure une relation systématique [L. Hjemslev, *texte n° 37*]; seule, enfin, elle peut donner à l'expérimentation son plein pouvoir de démenti en lui présentant un corps d'hypothèses si systématique qu'il est tout entier exposé en chacune d'elles.

On voudrait pouvoir dire de la sociologie ce que Bachelard disait de la physique expérimentale : « Le temps des hypothèses décousues et mobiles est passé, comme est passé le temps des expériences isolées et curieuses. Désormais l'hypothèse est synthèse. »<sup>1</sup> En fait, la vérification pointilliste qui soumet à des expérimentations partielles une série discontinue d'hypothèses parcellaires ne peut jamais recevoir de l'expérience que des démentis sans grande conséquence. Que l'on pense par exemple aux commodités que se donne l'analyse des résultats d'une enquête lorsqu'elle prend le tableau statistique pour unité d'interprétation : en omettant de poser la question de l'articulation des propositions dégagées de chaque tableau ou de ces chapelets de tableaux traînant chacun après soi le commentaire sur mesure qui le double, on s'évite d'exposer tout un corps systématique de propositions au démenti que pourrait lui opposer chacun des tableaux. Rien n'est mieux fait pour préserver la bonne conscience positiviste que la démarche qui consiste à aller d'une observation à une autre, sans autre idée que l'idée qu'une idée pourra surgir, puisque l'épreuve du démenti global qu'encourrait par exemple un modèle théorique est constamment repoussée et que les faits pris un à un n'ont rien à opposer à l'interrogation discontinue et inchoative de ces états crépusculaires de la conscience épistémologique où s'engendre « le même-pas-faux ». La rigueur apparente des techniques de preuve n'a pas en ce cas d'autre fonction que de dissimuler une dérobade : comme le jeune Horace, le chercheur s'assure une victoire aisée sur les faits en les fuyant pour pouvoir les affronter un à un.

Au contraire, lorsque l'hypothèse engage une théorie systématique du réel, l'expérimentation, qu'il faut alors appeler expérimentation théorique, peut exercer systématiquement son plein pouvoir de démenti. Comme l'observait Duhem, « une expérience ne peut jamais condamner une hypothèse isolée mais

1. Le pouvoir des prénotions, qu'elles soient populaires ou savantes, tenant au caractère systématique de l'intelligibilité qu'elles procurent, il est vain d'espérer les réfuter coup par coup. Historiquement, c'est toujours une théorie systématique qui a pu avoir raison d'illusions elles-mêmes systématiques, comme le font voir à propos des sciences physiques T. S. Kuhn (« The Function of Dogma in Scientific Research », in A. C. Crombie [ed.], *Scientific Change, op. cit.*, p. 347 sq.) et N. R. Hanson (*Patterns of Discovery*, Cambridge University Press, Cambridge, 1965).

2. G. Bachelard, *Le nouvel esprit scientifique, op. cit.*, p. 6.

seulement tout un ensemble théorique »<sup>1</sup>. Par opposition à une série discontinue d'hypothèses *ad hoc*, un système d'hypothèses tient sa valeur épistémologique de la cohérence qui fait sa pleine vulnérabilité : d'une part un seul fait peut le remettre en question tout entier et, d'autre part, construit au prix d'une rupture avec les apparences phénoménales, il ne peut recevoir la confirmation immédiate et facile que fourniraient les faits pris à leur valeur faciale ou les documents pris à la lettre. En effet, choisissant de s'exposer à tout perdre afin de tout gagner, le savant met à tout moment à l'épreuve des faits qu'il interroge tout ce qu'il engage dans son interrogation des faits. S'il est vrai qu'en leur forme la plus achevée, les propositions scientifiques sont conquises contre les apparences phénoménales et qu'elles présupposent l'acte théorique qui a pour fonction, selon l'expression de Kant, « d'épeler les phénomènes pour pouvoir les lire comme expériences », il s'ensuit qu'elles ne peuvent plus trouver leur preuve que dans la cohérence entière du système complet des faits créés par — et non pour — les hypothèses théoriques qu'il s'agit de valider. Pareille méthode de preuve, où la cohérence du système construit de faits intelligibles est à elle-même sa propre preuve, en même temps qu'elle est le principe de la vertu probatoire des preuves partielles que le positiviste manipule en ordre dispersé, suppose évidemment la décision systématique d'interroger les faits sur les relations qui les constituent en système. Ainsi, lorsque Erwin Panofsky présente comme un « élément de preuve » le *inter se disputando* de l'*Album* de Villard de Honnecourt, il n'ignore pas que cette inscription ne répond pas à une question de fait — par exemple celle de l'influence directe des scolastiques sur les architectes —, comme le voudrait l'historiographie positiviste pour qui l'interrogation est un simple questionnaire auquel le réel répondrait question après question, par oui ou par non; en réalité, ce petit fait tire sa force probatoire de ses relations avec d'autres faits qui, eux-mêmes insignifiants tant qu'ils sont considérés indépendamment des relations que le système d'hypothèses instaure entre eux, ne prennent toute leur valeur que comme termes organisés d'une série : « Qu'il s'agisse de phénomènes historiques ou naturels, l'observation particulière ne présente le caractère d'un 'fait' que lorsqu'elle peut être reliée à d'autres observations analogues, de telle sorte que l'ensemble de la série 'prenne sens'; le 'sens' peut donc être légitimement utilisé, à titre de contrôle, pour interpréter une nouvelle observation particulière à l'intérieur du même ordre de phénomènes. Si toutefois cette nouvelle observation particulière refuse, indiscutablement, de se laisser interpréter conformément au sens de la série, et s'il est prouvé qu'il n'y a pas d'erreur possible, le 'sens' de la série devra être reformulé de manière à inclure la nouvelle observation »<sup>2</sup> [E. Wind, *texte n° 38*]. C'est le même mouvement circulaire qu'opère le sociologue soucieux de ne pas imposer au donné ses propres présupposés lorsque, dans le dépouillement d'une enquête, il déchiffre à partir de l'ensemble des réponses au questionnaire le sens de chacune des questions par lesquelles il a suscité et construit ces réponses, tout en reformulant le sens de l'ensemble en fonction de ce qu'il apprend de chacune d'entre elles. Duhem n'employait pas un autre langage pour décrire la logique du progrès de la science physique, « tableau symbolique auquel de continuelles retouches donnent de plus en plus d'étendue et d'unité [...], tandis que chaque détail de l'ensemble, découpé et isolé du tout, perd toute signification et ne représente plus rien » et où un regard naïf ne saura voir qu'« un monstrueux fatras de pétitions de principe et de cercles vicieux »<sup>3</sup>.

1. P. Duhem, *La physique théorique*, op. cit., p. 278.

1. E. Panofsky, « Iconography and Iconology », *Meaning in the Visual Arts*, Doubleday, New York, 1955, p. 35.

2. P. Duhem, *La théorie physique*, op. cit., p. 311.

La preuve par la cohérence du système des preuves condamne au *cercle méthodique* où il serait trop facile de dénoncer un cercle vicieux : réinterprétant cette logique de la preuve par référence à une définition analytique de la vérification, le positivisme ne peut voir dans cette construction systématique des faits que le résultat d'une manipulation des données inspirée par l'esprit de système. C'est la même cécité qui porte certains à apercevoir dans l'analyse structurale d'un mythe la projection des catégories de pensée du chercheur ou même le protocole d'un test projectif, et un effet du parti pris dans le parti méthodique d'interpréter chacune des relations statistiques établies par une analyse multi-variée en fonction du système de relations entre les relations dont chaque relation tient sa signification. La force probatoire d'une relation empiriquement constatée n'est pas seulement fonction de la force de la liaison statistique : la probabilité composée de l'hypothèse mise à l'épreuve est fonction du système complet des propositions déjà établies (qu'il s'agisse de relations statistiques ou de régularités de tout autre type), c'est-à-dire de ces « concaténations de preuves », selon l'expression de Reichenbach, qui « peuvent être plus fortes que leur chaînon le plus faible et même que leur chaînon le plus fort »<sup>1</sup>, parce que la validité d'un tel système de preuves se mesure non seulement à la simplicité et à la cohérence des principes mis en œuvre, mais encore à l'extension et à la diversité des faits pris en compte, et enfin à la multiplicité des conséquences imprévues auxquelles il conduit [C. Darwin, *texte n° 39*].

1. A. Kaplan, *The Conduct of Inquiry*, op. cit., p. 245.

[ Extrait de : P. Bourdieu, J.-Cl. Chamboredon, J.-Cl. Passeron,  
Le métier de sociologue, Mouton, Paris - La Haye, 1973,  
pp. 88 - 92 ]

## ANNEXE 2

### La notion d'anamnèse

#### ANAMNÈSE

Issu des mots grecs *ana* (remontée) et *mneme* (souvenir), l'anamnèse signifie rappel du souvenir. Pour Platon, elle est la restauration de l'idée contemplée, avant l'incarnation, par l'âme humaine dans le ciel des idées et dont le souvenir serait resté inconscient sans l'opération de la « réminiscence ». Aristote refuse cette conception de la *theoria* et fait de l'anamnèse la faculté, propre à l'homme, de rappeler volontairement un souvenir d'origine empirique et de le localiser dans le temps. La *mneme* passe ainsi de la potentialité à l'état d'entéléchie.

C'est à la langue philosophique que les médecins ont emprunté le terme d'anamnèse pour désigner cette partie de l'enquête diagnostique qui reconstitue le passé de la maladie en ayant recours à la mémoire du malade et, par extension, à sa documentation et aux indications données par l'entourage. Comme une maladie s'inscrit toujours dans une histoire pathologique et physiologique totale l'anamnèse doit explorer aussi bien les antécédents personnels du malade que ses antécédents héréditaires, c'est-à-dire parentaux.

Le renouveau de la médecine hippocratique et le mouvement psychosomatique ont valorisé depuis un demi-siècle cet aspect de l'acte médical, mais c'est surtout avec la psychanalyse qu'il s'est souverainement imposé, puisqu'on a pu dire que la psychanalyse est une anamnèse prolongée. Si, en effet, le médecin ordinaire se contente de quelques renseignements factuels, le psychanalyste, lui, a forgé un instrument qui oblige le patient à ramener à sa conscience claire et vécue des souvenirs inconscients ; et, si l'anamnèse ne suffit pas à la guérison (due plutôt à l'analyse du transfert et aux abréactions), il n'empêche que, tout au long de son déroulement, la cure est un rappel indéfini, patient, méfiant, critique et rusé de souvenirs que le sujet croyait et voulait perdus.

Dans la psychanalyse inspirée de Carl Gustav Jung, l'anamnèse découvre des « archétypes » qui n'appartiennent plus à l'individualité du patient, mais à l'inconscient collectif. En parapsychologie, d'inquietantes et suspectes anamnèses, obtenues par hypnose, remontent très haut dans la vie du sujet et prétendent même extraire du tréfonds de l'âme des souvenirs de vies antérieures.

[ Extrait du Thesaurus de l'Encyclopaedia Universalis ]

**ANAMNESIS** Medical term for the history of an illness as given by the patient at consultation, i.e. for his recall of the relevant past.

[ Extrait de Charles Rycroft, A Critical Dictionary of Psychoanalysis, Penguin Books, 1972, p. 7 ].

La notion de corpus

**CORPUS**

Ensemble homogène et significatif de données linguistiques observées et à partir desquelles pourra s'élaborer la théorie. La notion de corpus est évidemment fondamentale dans la linguistique structurale : désireuse de substituer à la normativité de la grammaire ou aux fondements pseudo-logiques de la signification un enregistrement de l'usage, écrit ou oral, ainsi qu'un relevé des différences en fonction des différents contextes, elle pose en principe l'absolue nécessité de s'appuyer sur des productions de parole. Les caractéristiques d'un corpus significatif sont : l'homogénéité (le groupe qui le produit est socialement défini), la synchronie (on ne peut travailler sur des énoncés trop éloignés chronologiquement les uns des autres), la moindre redondance possible (on supprimera des énoncés représentatifs d'un phénomène identique). Nanti de son matériau, le linguiste distributionnaliste induira, par l'intermédiaire d'une formalisation des contextes et en s'appuyant sur l'hypothèse des deux axes de fonctionnement (paradigmatique et syntagmatique), les lois du code. La notion de corpus prête le flanc à une critique néanmoins sérieuse : notamment, dans le modèle de langage impliqué par le recours à l'échantillon, on ne peut que rester en deçà de la perspective créative par laquelle le sujet est capable de produire aussi bien que de comprendre un nombre infini d'énoncés qui ne figurent pas dans le corpus. Il faut alors reconnaître la fragilité d'une observation inductive comparée aux théories hypothético-déductives capables d'intégrer cette créativité. On est alors conduit à renoncer au corpus pour le remplacer par l'épreuve de grammaticalité auprès d'un locuteur natif.

[ Extrait du Thesaurus de l'Encyclopaedia Universalis ]

## Linguistique générative et distributionnalisme

Élève d'abord de Z. S. Harris, qui a poussé le distributionnalisme [49s.] jusqu'à ses conséquences les plus extrêmes, l'Américain N. Chomsky, après s'être intéressé lui-même à la formalisation (au sens logico-mathématique de ce terme) des notions distributionnalistes de base, a proposé une conception nouvelle, dite générative de la linguistique, conception qui contredit les dogmes distributionnalistes, et les a rapidement remplacés comme fondement de la recherche linguistique américaine.

Du distributionnalisme, Chomsky souhaite retenir le caractère explicite. Le distributionnalisme est explicite en ce sens que les descriptions de langues auxquelles il aboutit, n'utilisent, comme concepts élémentaires (= non définis), aucune notion dont la compréhension implique déjà la connaissance, soit de la langue décrite, soit du langage en général : son concept de base, la notion d'environnement (telle unité, dans tel énoncé, est entourée par telles et telles unités) est compréhensible pour qui, par une hypothèse absurde, n'aurait aucune expérience personnelle de la parole. C'est là, pour Chomsky, la supériorité du distributionnalisme sur les grammaires traditionnelles, et aussi sur la linguistique dite fonctionnaliste [42], qui recourent à des notions comme la dépendance (« tel mot se rapporte à tel autre »), la fonction-sujet (tel mot représente ce dont on parle), etc., dont la compréhension est partie intégrante de la faculté du langage, et qu'on ne saurait donc, sans cercle vicieux, utiliser pour décrire cette faculté.

Mais Chomsky reproche au distributionnalisme de payer son caractère explicite par des abandons impossibles à admettre. D'abord, par une limitation excessive du domaine empirique qu'il prend pour objet. Car une langue est tout autre chose qu'un corpus [50].

a) Alors qu'un corpus est par définition un ensemble *fini* d'énoncés, toute langue rend possible une *infinité* d'énoncés : puisqu'il n'y a pas de limite au nombre de propositions que l'on peut introduire dans une phrase française, on peut, à partir de tout énoncé français, en fabriquer un autre, aussi régulièrement construit (en ajoutant, par exemple, une proposition relative) : le distributionnalisme est condamné par sa méthode à ignorer ce pouvoir d'infini inclus dans toute langue.

b) Bien plus, une langue, ce n'est pas seulement un *ensemble* d'énoncés (fini ou infini), mais tout un *savoir* à propos de ces énoncés. Car on ne dira pas de quelqu'un qu'il connaît une langue, s'il ne sait pas distinguer les énoncés ambigus des énoncés à une seule interprétation, s'il ne sent pas que tels et tels énoncés ont des constructions syntaxiques semblables, tels autres, des constructions très différentes, etc. Or ce savoir des sujets parlants concernant leur propre langue, les distributionnalistes l'excluent délibérément de leur champ descriptif, et se contentent de décrire la façon dont les unités se combinent dans les énoncés (voir, ici-même, la notion de *compétence* selon Chomsky, p. 158 s.).

Même si l'on admettait cette réduction du domaine décrit (on ne peut prétendre tout décrire), il y a un deuxième abandon que Chomsky reproche au distributionnalisme, c'est justement de se contenter de *décrire*, et de renoncer à *expliquer*. En cela, les successeurs de Bloomfield seraient fidèles à une conception empiriste selon laquelle la science a seulement à décrire les phénomènes, en cherchant à mettre un peu d'ordre dans leur désordre apparent : la tâche essentielle du chercheur serait alors la classification, la *taxinomie*. C'est bien là en effet l'objet unique des distributionnalistes, pour qui une grammaire est simplement une classification des segments (phonèmes, morphèmes, mots, groupes de mots) qui apparaissent dans les énoncés du corpus. Et, dans la mesure où le principe de cette classification est de regrouper les éléments à distribution [50] identique (ou voisine), on peut la considérer, selon l'expression de Harris, comme une « description compacte » du corpus : une fois en possession de cette classification, il doit être facile en effet de reconstituer tous les énoncés du corpus. Selon Chomsky, au contraire, toute science, en se développant, est conduite à se fixer un but plus ambitieux que la description et la classification. Il doit en être de même pour la linguistique, qui peut prétendre présenter des hypothèses à valeur *explicative*, concernant la faculté qui est à l'origine de l'activité linguistique. Il ne suffit pas de dire, même de façon compacte, quels sont les énoncés possibles et impossibles, quels sont les énoncés ambigus, syntaxiquement apparentés, etc., mais il faut que toutes ces remarques de détail puissent être reliées à une conception d'ensemble du langage. C'est pour réconcilier le souci d'être *explicite* et celui d'être *explicatif*, que Chomsky a été amené à proposer une nouvelle définition de ce qu'est une grammaire et de ce qu'est une théorie linguistique.

[ Extrait de : Oswald Ducrot et Tzvetan Todorov :  
Dictionnaire encyclopédique des sciences du langage,  
Seuil, Paris, 1972, pp 56-58 ]

La structure d'un vocabulaire  
politique et social

**Un champ conceptuel dans un corpus**

Parmi les recherches effectuées pour découvrir si le lexique est structurable — au sens où le sont les phonèmes et les formes grammaticales — celle de Jean Dubois, dans son ouvrage sur *le Vocabulaire politique et social en France de 1869 à 1872*, mérite un examen approfondi.

A la différence des tentatives antérieures, qui esquissaient une possibilité sur deux exemples en passant, ou choisissaient un champ privilégié (grades militaires, espèces botaniques, etc.), celle de Dubois est remarquable par le domaine exploré (une science humaine peu systématisée, la sociologie politique), par l'étendue du champ (tout le vocabulaire politique et social), ainsi que par la méthode employée. Ici le travail a porté sur une synchronie vraie (1869-1872) et sur un corpus réel : 650 auteurs de livres, discours, articles, lettres. Témoignages littéraires

écrits, donc, et témoignages oraux ; mais aussi brochures, manifestes, tracts, affiches. Pour la presse, plus de 250 périodiques consultés, pour un, dix, ou plus de 300 numéros. Au total, presque 600 bases lexicales dont l'usage est illustré par 5 408 contextes datés, répertoriés et cités. Si beaucoup de termes (*garibaldien*, *képisme*, etc.) sont épisodiques, *masses* est cerné par 60 contextes, *démocrate* par 100, *classe* par 256.

Aucun doute que les matériaux ainsi dépouillés, classés et analysés ne présentent beaucoup d'intérêt pour les historiens, les sociologues, les francistes qui les exploiteront pour plus d'un nouveau travail. Mais ce qu'on examine ici, c'est l'apport de l'ouvrage en linguistique générale : sa méthodologie.

Tous les efforts en vue d'appliquer le structuralisme au lexique ont buté sur ce point : pour aboutir à des procédures objectives, il faut admettre que seul un appui pris sur des marques formelles est la garantie du caractère proprement linguistique d'une analyse. A chaque différence de signification qu'on affirme, notamment, doit correspondre une différence de forme. Et, réciproquement, les différences de forme ont une fonction linguistique seulement si elles entraînent des différences de signification. Ce type d'analyse nous apprendra bien que les formes lexicales *féodal* et *bourgeois* ont des sens distincts ; mais, apparemment, nous ne pourrions pas aller plus loin. Sauf dans l'analyse des « séries étymologiques », du genre *ouvrier*, *ouvrière*, *ouvriérisme*, *anti-ouvrier*, qui structurent formellement mais partiellement tout champ lexical, par des procédures bien connues : affixation et composition. Depuis Saussure, on

ournait la difficulté en recourant à une procédure conceptuelle, et non proprement linguistique, pour déterminer le contenu et délimiter les frontières d'un champ lexical : on supposait donné par une science autre que la linguistique un *champ conceptuel*. (Quand nous établissons des relations, entre les signifiants *réaction, conservatisme, indifférentisme, réformisme, révolution*, c'est parce que la science politique nous a fourni les définitions classificatoires qui apparentent leurs signifiés).

### Des procédures formelles

L'originalité de Dubois sur ce point c'est — tout en acceptant comme point de départ les données d'un champ conceptuel, celui des « rapports économiques, politiques et sociaux » — d'élaborer des procédures proprement linguistiques pour vérifier les relations établies entre les termes du champ lexical correspondant.

Quelles sont ces procédures ? Dubois, certes, utilise les « séries étymologiques ». Soit pour mettre en évidence l'entrée d'un mot dans le champ (*égalité*, mot du lexique général, est sûrement à mettre dans le lexique politique au moment où il produit *égalitaire, égalitarisme, anti-égalitaire*, en utilisant des affixes de haute fréquence dans ce champ). Soit pour confirmer la validité d'autres procédures (on peut parler de couples antonymiques typiques du champ, puisqu'une soixantaine de ces couples sont déjà décelés par les préfixes *anti-*, ou *in-* ou *non-*, par conséquent sur des critères linguistiques formels). La série étymologique est donc ici surtout moyen de contrôle.

Les autres procédures proposées par Dubois pour structurer le champ sémantique à partir de critères linguistiques formels sont la mise en évidence, entre les termes, de réseaux d'oppositions, d'identités, d'associations, de corrélations. Ce qu'il nomme oppositions, ce sont des « couples antonymiques » qui « ne sauraient être confondus avec les contraires [...] des dictionnaires » : en effet, un contexte peut offrir des couples du type *réaction~liberté*. Mais ce sont aussi des « réseaux d'oppositions » : soit des constellations d'oppositions autour d'un même mot (*révolution~aristocratie, révolution~réformes, révolution~ordre*, etc.), soit des séries d'oppositions parallèles (*classes riches~pauvres, d'en haut~d'en bas, supérieures~inférieures, bourgeoises~populaires*, etc.). L'existence d'une relation d'opposition entre termes est établie, soit par une marque formelle (*ordre~désordre, révolutionnaire~contre-révolutionnaire*), soit « par des commutations qui font ressortir les oppositions ou les identités » (mais Dubois ne les utilise pas pour établir des oppositions) ; soit surtout par des analyses de contextes qui font la preuve de l'usage antonymique (« tenons-nous à égale distance de la *réaction* et des théories *révolutionnaires* » ; « s'il est pour la *révolution* ou pour la *réaction* », etc.

Les identités qui structurent le champ sémantique, dit Dubois, ne sauraient, elles non plus, être confondues avec les synonymes des dictionnaires. Des « unités sémantiques parallèles » comme : *avènement des masses* [...], *des prolétaires* [...], *des travailleurs* [...], etc., indiquent de telles identités, que Dubois nomme aussi

« substituts sémantiques » et « variantes combinatoires ». L'existence d'une relation d'identité entre substituts ou variantes est établie soit par des contextes, comme ci-dessus (*civilisation* et *révolution*, et *ordre social*, et *progrès*, etc.), soit par des commutations opérées à partir des unités lexicales complexes ou des syntagmes les plus fréquents, du type : *affranchissement des travailleurs* [...], *du travail* [...], *des masses*, etc. « Ces syntagmes, dit Dubois, peuvent avoir le caractère de groupements phraséologiques stéréotypés : chaque élément du syntagme perd alors une partie des traits pertinents qui le définissent sur le plan du lexique; deux termes qui sont distincts dans d'autres syntagmes, se trouvent être ici permutable. Ils deviennent des *variantes combinatoires*, et le phénomène est comparable à celui de la neutralisation phonologique dans le cas des phonèmes ».

Les associations sont des relations entre termes plus délicates à bien définir. Tantôt Dubois nomme ainsi de véritables identités (*ouvriers* = *travailleurs* = *salariés* = *pauvres* = *déshérités*), qu'on pourrait démontrer telles par les procédures indiquées au paragraphe précédent. Tantôt il s'agit de relations entre séries parallèles : *paix sociale*, *mal social*, *niveau social*, *hiérarchie sociale*, etc. Tantôt, sans doute, de véritables associations : *révolution* s'associe à *liberté*, à *désordre*, à *progrès*, à *socialisme*, à *travailleurs*. Ici la mise en évidence de la relation se fait exclusivement par le contexte : par la fréquence du voisinage des termes dans la même phrase.

La notion de corrélation se révèle elle aussi délicate à bien cerner. Si l'on écarte les emplois au sens général, on trouve que les corrélations désignent essentiellement les relations contractées entre termes dans des contextes, donc surtout les associations (« Le contenu d'une unité lexicale est donc défini dans une phrase par un double système de relations, par des oppositions paradigmatiques, et par des corrélations syntagmatiques »). Le terme serait synonyme d'association. Mais il peut désigner aussi l'ensemble des relations (opposition, identité, association) concernant un ou plusieurs termes. On considérera donc ici que la corrélation ne désigne pas chez Dubois une procédure distincte de structuration du lexique, mais est un terme générique couvrant les trois autres procédures.

### Champ ou système ?

Comment apprécier cet effort méthodologique considérable ? Sans doute est-ce la meilleure illustration, étendue, et systématique, qui ait été donnée de la thèse saussurienne, selon laquelle « la valeur d'un terme est constituée uniquement par des rapports et des différences avec les autres termes de la langue ». Sans doute est-ce aussi une des meilleures illustrations de la thèse lexicologique de Meillet (évoquée par Dubois), selon qui « le sens d'un mot n'est que la moyenne entre les emplois qu'en font les individus et les groupes d'une même société » — thèse que Wittgenstein redécouvre avec son dicton : « La signification d'un mot n'est rien d'autre que l'ensemble de ses usages ». L'utilisation systématique de milliers de contextes fait

de Dubois l'un des rares linguistes qui ait raisonné non sur des significations *a priori*, mais exclusivement sur des usages, et non pas fabriqués pour les besoins de la cause, mais tous attestés, et nombreux. Chez lui toute relation, c'est-à-dire toute structure lexicale, est justifiée (et n'est justifiée que) par un faisceau de contextes synchroniques, sans comparaison plus riche statistiquement que tout ce qu'on avait fait jusqu'ici.

De plus, le travail minutieux de Dubois attire utilement l'attention sur la fréquence élevée de ces couples antonymiques chers à Jolles (*droit-gauche, jour-nuit*) que Trier avait éliminés péremptoirement comme moyen trop rudimentaire de construire un champ sémantique. Il se pourrait qu'une analyse psychologique et linguistique plus fine montre l'importance, au moins statistique, de cette structuration lexicale binaire, qui correspond peut-être à un caractère fondamental du fonctionnement de la pensée : le recours aux unités discrètes, la mesure par *oui* ou *non* ?

Il reste à vérifier si l'objectif théorique essentiel est atteint : démontrer que les méthodes de l'analyse fonctionnelle et structurale sont utilisables telles quelles intégralement pour le lexique, au point de pouvoir affirmer comme Dubois que « l'unité lexicale joue, sur le plan du lexique, le même rôle que le phonème sur le plan phonologique ». Dubois pose d'emblée le problème exactement : pour que les procédures structuralistes soient efficaces, il faut que les champs sémantiques se révèlent être des *systèmes*, au sens rigoureusement linguistique du terme. Il adopte à demi cette terminologie : toutes les fois qu'il s'agit de discussion théorique ou méthodologique il parle de système lexical autant que de champ. Or un système, au sens opératoire étroit du concept, c'est l'inventaire des unités susceptibles d'apparaître à un point déterminé de l'énoncé, le choix de l'une excluant toutes les autres en ce point. Pour les phonèmes, il y a donc en ce sens un nombre défini de systèmes, lié à un nombre défini de points remarquables de l'énoncé, c'est-à-dire de contextes — une douzaine peut-être : consonnes à l'initiale, à la finale, intervocaliques ; voyelles en syllabe ouverte ou fermée, voyelles initiales, finales, pré-accentuelles, etc. Le système des voyelles et celui des consonnes d'une langue, donc, puis le système phonologique entier, sont des « systèmes de systèmes ». Pour les unités lexicales, les choses sont assez différentes. D'une part, le système lexical *stricto sensu* d'une unité comme *bourgeoisie* par exemple, ce serait l'inventaire des unités (sémantiquement très disparates), susceptibles de se substituer à cette unité dans un énoncé sans le détruire comme tel : « *La bourgeoisie, voilà l'ennemi* » (*Le cléricisme, le phylloxéra, l'alcool, le socialisme, etc.*). D'autre part, le nombre des contextes à explorer serait immense, sinon illimité : dans le corpus de Dubois, à supposer qu'il ait relevé tous les contextes du mot *classe* — ce qui n'est sûrement pas le cas — cela ferait 256 contextes du mot, et donc 256 systèmes au sens propre, dont il faudrait voir les similitudes et les différences. Ainsi, quand on détermine un « système lexical » sans recourir à la rigoureuse procédure formelle de type phonologique, on introduit subrepticement quelque chose de plus que dans cette procédure, ce qui permet d'esquiver les deux obstacles ci-dessus. Mais quoi ? En fait, on

décide, entre toutes les unités qui peuvent « former système » avec une autre à partir d'un ou plusieurs énoncés donnés, de ne retenir que celles qui sont sémantiquement apparentées (*la bourgeoisie, le prolétariat, l'aristocratie, le capitalisme...*, voilà l'ennemi ; mais non *l'alcool, le phylloxéra*, etc.). La présentation semble déterminée par des critères formels ; en réalité, on a réintroduit une décision d'ordre conceptuel, implicite, dans le circuit opératoire. Ceci n'invalide pas la construction de tels champs sémantiques, mais attire l'attention sur le fait que, malgré les apparences, on n'a pas réussi à délimiter le champ sur des marques formelles, et d'après des procédures exclusivement et proprement linguistiques.

### Commutation et traits pertinents

La chose apparaît encore plus nettement quand on examine l'usage de la notion de commutation chez Dubois. Si le champ lexical peut être traité sans métaphore comme un système, l'opération spécifique qui le prouvera, c'est la commutation : Dubois a raison. Mais en fait, il emploie le terme 6 fois, dont 3 dans la conclusion, d'une manière générale (« On détermine la pertinence des unités lexicales par des commutations qui font ressortir les oppositions ou les identités » ; voir aussi l'emploi des synonymes un peu lâches, *permutable* et *permutation*). Les deux seuls exemples que donne Dubois sont intéressants, parce qu'ils montrent qu'on peut utiliser la commutation pour mettre en évidence des identités dans des unités lexicales complexes ou dans des syntagmes fréquents du type *affranchissement des travailleurs*. Ici, la commutation fonctionne correctement, mais négativement, sur un point limité : des formes linguistiques différentes (*des travailleurs, des masses, du prolétariat*, etc.) se révèlent identiques quant au sens parce qu'elles n'altèrent pas la signification de l'unité lexicale complexe. Mais la commutation ne fonctionne ni pour les oppositions, ni pour les associations, sauf si par une décision d'ordre conceptuel, non-linguistique, on choisit d'abord les contextes et les unités sur quoi la faire porter exclusivement. Les paires antonymiques *travailleurs~capitalistes, travailleurs~maîtres, travailleur~oisif*, ne sont jamais définies par commutation de type phonologique, mais par contextes.

La notion de trait pertinent chez Dubois mène à la même observation. Dubois l'emploie peu, une douzaine de fois, trop allusivement peut-être, et ne la définit bien que dans sa conclusion, par référence à Ivanov. Presque toujours, le trait sémantiquement pertinent sur lequel il opère, c'est celui qui manifeste une opposition (« Un mot ne saurait être analysé sans celui qui est son contraire, l'antonymie étant un trait pertinent », ce qui nous ramène en général à une série étymologique : dans *anti-révolutionnaire*, par rapport à *révolutionnaire*, il y a bien, formellement apparent, un « trait pertinent ou sème », représenté par *anti-*, sur lequel peut jouer la commutation (*clérical~anti-, autoritaire~anti-*, etc.), exactement comme dans la corrélation phonologique *sourdes~sonores*, par exemple. Mais par quels « sèmes » se justifient formellement les couples *anarchie~*

*ordre, aristocratie-peuple, bourgeois-ouvrier, etc.* ? Ici aussi, on passe d'un domaine où les traits sémantiquement pertinents sont dégagés par des marques formelles, à d'autres domaines où ils le sont par des procédures conceptuelles, non-linguistiques, au cours desquelles la comutation n'est ni utilisable ni utilisée.

### Une analyse contextuelle exemplaire

Ces limitations de la méthode employée par Dubois, si elles contestent l'interprétation théorique trop optimiste qu'il donne dans sa conclusion, n'entament ni sa validité ni ses résultats. D'abord, les observations ci-dessus n'ont été rendues possibles que parce que la recherche de Dubois existe, et a été poussée loin au-delà du point où s'arrêtent trop d'articles sur la structuration sémantique, celui des esquisses. A cet égard, le travail est et restera exemplaire par sa démonstration du rôle des contextes utilisés scientifiquement, c'est-à-dire exhaustivement, dans ce domaine. Il prouve aussi le rôle auxiliaire non négligeable des séries étymologiques, et le rôle, plus important qu'on ne le soupçonnait, des oppositions lexicales, et peut-être des identités. Mais ces procédures ne sont sûrement ni toujours efficaces, ni les seules qui permettent de structurer tous les champs lexicaux.

Au terme de l'étude, on n'échappe pas à l'impression que la méthode employée par Dubois réussit d'autant mieux que le champ choisi, tout à fait par hasard, était plus favorable qu'il n'y paraît d'abord : à cause du fait qu'un nombre important de dénominations devaient s'appliquer, par la nature même des choses, à des relations non-linguistiques binaires, antagonistes ; du fait aussi que ces relations antagonistes (économiques, politiques et sociales) engendraient un vocabulaire « affectif » (pour~contre, bon~mauvais, etc.) particulièrement important, et, par sa nature même, à tendance polaire. On n'échappe pas non plus à l'impression que cette méthode, parce que son instrument de travail est fondamentalement le contexte, réussit mieux à décrire les relations syntagmatiques entre les termes du champ, que toutes les relations paradigmatiques. Sur ce point, elle ne cerne que celles qui sont révélées par des contextes (oppositions, identités, associations). Cette méthode à elle seule donne des résultats remarquables (voir par exemple l'analyse des distributions complémentaires de *paysan, cultivateur, agriculteur, rural*, pp. 84-87)<sup>1</sup>. Mais elle s'interdit d'analyser les traits sémantiquement pertinents qui ne sont pas révélés par le contexte (qui font que *ouvrier* ≠ *travailleur* ≠ *salarié* ≠ *prolétaire* ≠ *rouge*, etc.), c'est-à-dire arrête l'analyse au point exact où elle cesserait d'être purement linguistique. Cette attitude est loin d'être critiquable ; et Dubois démontre bien qu'elle livre beaucoup plus de la structure d'un lexique qu'on n'osait le supposer, par une méthode qui calque au fond celle de l'apprentissage spontané du lexique par le locuteur. Mais on peut se demander si elle livrera la structure complète de tous les champs lexicaux. D'abord,

1. Voir à l'Appendice LX, ci-après, d'autres exemples de ce type d'analyse, dus à Paul Vicaire et Jean Casabona.

parce que, par les contextes, le locuteur apprend l'usage des *valeurs* saussuriennes des termes, et non la totalité de leur signification. Ensuite, parce que celle-ci s'acquiert pleinement par référence situationnelle, ou logique (définitions), autant ou peut-être plus que par référence contextuelle. La méthode de Dubois ne fournirait pas — même avec des dépouillements étendus sur un corpus de plusieurs centaines de milliers de mots — les traits sémantiquement pertinents qui différencient réellement *cheval*, *hongre*, *jument*, *poulain*, par exemple ; certains contextes seuls (définitoires) en donneraient la signification mais par une analyse de type conceptuel. Les contextes, c'est-à-dire les usages, en fin de compte les marques formelles (ici distributionnelles), ne suffisent pas à définir la signification d'un terme, parce que la structuration du lexique n'est peut-être pas entièrement de nature linguistique ; selon l'avertissement déjà ancien de Martinet, « l'expression [linguistique] est un moyen, le contenu [sémantique] une fin » : ce qui signifie que la structuration sémantique doit refléter la structuration que nous nous donnons du monde non-linguistique, laquelle n'est jamais rigoureusement parallèle à la structuration des signifiants que nous offre la langue. La structuration du lexique est donc, partiellement, mais de façon variable avec les champs lexicaux (*chat*, *chatte*, *chaton* ≠ *étalon* [*cheval*], *jument*, *poulain*), de nature conceptuelle, comme le révèlent les insertions implicites d'opérations conceptuelles, non-linguistiques, qui se sont glissées dans les procédures de Dubois.

[ Extrait de : Georges Mounin, Clefs pour la sémantique,  
Seghers, Paris, 1972, pp. 65 - 77 ]