

# Métaphore et analogie dans les théories scientifiques (des objectifs et des structures des théories scientifiques)

Pierre-Yves Raccah, CNRS  
CeReS, Université de Limoges

Table Ronde de Nice 2006.

Texte très provisoire pour  
démarrer les discussions

## RESUME

L'analyse que je propose ici vise d'une part à caractériser les objectifs des théories scientifiques et, d'autre part, à déterminer les moyens structuraux dont ces théories disposent pour satisfaire ces objectifs. En ce qui concerne le premier point, je critique la conception naïve selon laquelle le rôle des théories est d'"expliquer" les phénomènes relatifs à un domaine, et lui substitue une conception selon laquelle les théories *engendrent des descriptions* des phénomènes en vertu de règles formelles. Au cours de cette discussion, j'examine la notion intuitive de "loi naturelle", le concept d'adéquation explicative et je décris ce que l'on entend par "rendre compte d'un phénomène"; les concepts d'hypothèses externes et internes sont précisés et comparés à ceux de mesure et de calcul; un modèle formel des théories empiriques est alors décrit. J'examine ensuite les moyens dont une théorie dispose pour délimiter son objet d'étude; deux stratégies de découpage de l'observable sont décrites et discutées dans leur application aux sciences du langage: l'approche épistémique et l'approche procédurale.

## ABSTRACT

The present work intends, on the one hand, to characterise the aims of scientific theories and, on the other hand, to determine the structural means by which these theories achieve these aims. I criticise the naive approach according to which the rôle of theories is to "explain" the phenomena of some domain and propose an approach according to which theories *generate descriptions* of phenomena by virtue of formal rules. Within this discussion, I examine the intuitive notion of "Natural Law", the concept of explanatory adequacy and describe what is understood by "accounting for a phenomenon"; the concepts of external and internal hypotheses are made precise and compared to those of measure and calculus; a formal model of empirical theories is then described. I then examine the means which a theory can use for the delimitation of its object; two strategies are described and discussed in their application to the sciences of language: the epistemic approach and the procedural approach.

# 1 Description et explication

## 1.0 Explication et signe

Pour que l'on puisse parler d'« expliquer », certaines conditions doivent être réunies, qui contraignent fortement la notion étudiée:

1. Il doit d'abord y avoir quelque chose à *propos de laquelle* une explication peut être fournie (un comportement, un texte, un schéma ou autre représentation perceptible, bref, un *phénomène*<sup>1</sup>).
2. Cette chose à propos de laquelle une explication peut être fournie doit être *perçue* ou imaginée par un destinataire; cette perception peut être plus ou moins complète.
3. Cette chose, à propos de laquelle une explication peut être fournie, est susceptible d'être *comprise* ou de ne pas être comprise; le destinataire pouvant la percevoir sans la comprendre.
4. Enfin, que cette chose soit présentée au destinataire par un agent, ou qu'elle soit donnée à lui par la nature (auquel cas, peut-être faudrait-il considérer que le destinataire se la présente à lui-même et joue, de ce fait, également le rôle de premier agent), l'explication, si elle a lieu d'être, est *émise* par un agent<sup>2</sup>.

### Comprendre et expliquer

Dans une situation d'explication, nous sommes donc en présence d'un phénomène  $\Phi$ , perçu ou imaginé par un agent  $\alpha$ , phénomène pouvant être compris ou non par  $\alpha$  (indépendamment du fait qu'il soit perçu). On dira, alors, qu'un agent  $\beta$  (différent de  $\alpha$ ) *explique*  $\Phi$  à  $\alpha$  si (i) à un temps  $t_1$ ,  $\alpha$  perçoit  $\Phi$  mais ne le *comprend* pas (ou pas entièrement); (ii) à un temps  $t_2 > t_1$ ,  $\beta$  émet *quelque chose* en direction de  $\alpha$ , qui est perçu par ce dernier; et (iii) à cause de l'action de  $\beta$ , à un temps  $t_3 \geq t_2$ ,  $\alpha$  *comprend*  $\Phi$ <sup>3</sup>.

Malgré le caractère apparemment formel du paragraphe précédent, la caractérisation qu'il fournit de l'explication reste très vague: Que signifie *comprendre*  $\Phi$ ? Quelle est la nature de ce qui est émis par  $\beta$  (qu'est-ce donc qu'une *explication*? On voit que nous sommes encore loin d'avoir répondu à la question initiale...). De quelle manière l'action de  $\beta$  *cause* la compréhension de  $\Phi$  chez  $\alpha$ ? Cette formulation nous permet néanmoins de risquer une analogie intéressante.

---

<sup>1</sup> On s'aperçoit déjà que, en fonction de cette entité à propos de laquelle une explication peut être fournie, les notions d'*expliquer* et d'*explication* varient.

<sup>2</sup> Bien que cela ne soit pas nécessaire pour la suite de cette réflexion, il me semble utile d'ajouter ici que cet agent -explicateur- doit probablement être différent du destinataire (ce qui n'est pas le cas, comme on l'a vu, pour l'agent qui présente la chose à propos de laquelle l'explication pourrait être fournie).

<sup>3</sup> Pour être rigoureux, il faudrait ajouter que c'est en mettant en œuvre les intentions de l'action de  $\beta$  que  $\alpha$  comprend  $\Phi$ , et non pas seulement à cause de l'action de  $\beta$ ...

Tout se passe comme si  $\Phi$  constituait un signe d'un système sémiologique<sup>4</sup>, signe que  $\alpha$  est censé percevoir, mais peut ne pas être en mesure d'interpréter; l'action de  $\beta$  vise alors à donner à  $\alpha$  les moyens d'interpréter  $\Phi$ . Si l'on adopte cette analogie, les questions philosophiques qui viennent d'être posées ne sont plus qu'une spécialisation des questions que l'on peut poser à propos de tout système de signe : on sait que, même s'il n'est pas pensable d'y apporter de réponses générales, cela n'empêche pas l'étude des systèmes de signes.

L'utilisation de cette analogie impose cependant des contraintes sur la nature des explications (ce qui est émis par  $\beta$  pour amener  $\alpha$  à comprendre  $\Phi$ ). En effet, s'il s'agit d'une aide à l'interprétation d'un signe, ce que  $\beta$  émet, c'est à dire l'explication proprement dite, doit lui-même relever d'un système sémiologique. Nous affinerons ici l'analogie dans la direction de la linguistique et nous intéresserons aux énoncés explicatifs. Ce choix n'implique pas que l'on doive considérer toute explication comme relevant nécessairement de la linguistique; néanmoins, si, en accord avec l'analogie décrite plus haut, on conçoit l'explication comme relevant d'un système sémiologique, on est amené à penser que le point de vue linguistique n'est pas seulement la spécification d'un contexte explicatif possible, mais une sorte de paradigme duquel il est possible de s'inspirer pour étudier tout type d'explication.

### **Comprendre et comprendre**

La chose à propos de laquelle une explication peut être fournie, et qui doit donc, selon la perspective choisie, être considérée comme un signe, est ce que l'agent explicateur vise à faire *comprendre* au destinataire. D'autre part, l'agent explicateur produit, pour faire *comprendre* cette chose, une explication (dans la perspective choisie, un segment linguistique), qui est elle-même censée être *comprise* par le destinataire, et qui plus est, *comprise comme une explication*. Il y a donc en jeu deux notions de compréhension, qui n'ont pas de raisons de se confondre: celle qui concerne le phénomène à expliquer et celle qui concerne l'explication elle-même (dans notre cas, le segment linguistique explicatif). Dans le premier cas, il s'agit de relier le phénomène à une interprétation, qui n'est autre que celle qui doit être assignée au segment explicatif (cf. plus haut); dans le second cas, il s'agit de construire l'interprétation du segment explicatif. Le schéma suivant illustre la situation:

---

<sup>4</sup> Par "système sémiologique", il faut entendre tout système d'entités susceptibles d'être interprétées dans ce système, et supposant donc des êtres du genre *émetteur*, des entités ressemblant à des *signes* et des êtres jouant un rôle voisin de ce lui de *destinataire*. Les langues constituent des exemples de systèmes sémiologiques particuliers; d'autres exemples, comme la signalisation routière, l'habillement, le design, etc. ont bien été étudiés.

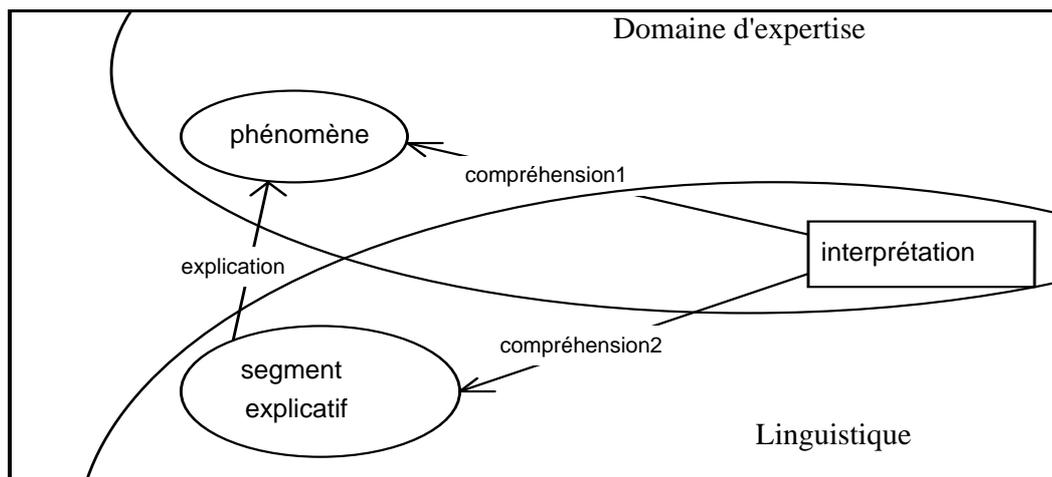


Schéma 4 : explication, compréhension et compréhension

Cette situation explique une confusion, ou du moins une assimilation abusive, que l'on retrouve dans presque tous les travaux sur l'explication : très rares, en effet, sont les auteurs qui opèrent une distinction, dans leurs fondements théoriques<sup>5</sup>, entre les phénomènes, les concepts, et les signes qui y renvoient. Comprendre un phénomène, un concept ou des signes (par exemple une explication...) est, chez ces auteurs, considéré *a priori* comme étant la même activité. Ce manque de précision pourrait être considéré comme de la désinvolture si l'on ne tenait pas compte de cette "circonstance atténuante" que constitue le caractère sémiologique des phénomènes vis-à-vis de leur interprétation dans un système conceptuel explicatif (tel que l'illustre l'angle nord-est du schéma ci-dessus). Il n'en demeure pas moins que, même si les phénomènes sont susceptibles d'être *compris*, ils ne le sont pas de la même façon qu'un énoncé ou qu'un texte exprimé dans une langue (naturelle), ou même dans un langage formel. Ainsi, le mot "comprendre" dans l'expression "comprendre le mot « honnêteté »" n'a pas la même signification que dans l'expression "comprendre le concept d'honnêteté", et ces deux significations sont encore différentes de celle qu'il acquiert dans l'expression "comprendre l'honnêteté d'une action ou d'un comportement". Il résulte de cela qu'expliquer un concept, un phénomène ou un texte sont trois activités différentes.

Un exemple

A ce point, un exemple pourra être utile... Considérons le phénomène  $\Phi$ , qui pourrait être décrit par l'énoncé suivant :

Lorsqu'on chauffe un thermomètre à mercure, la colonne de mercure s'élève

Considérer que  $\Phi$  est susceptible d'être expliqué, c'est à dire considérer qu'il peut être compris ou ne pas être compris, revient à considérer qu'il est un signe de quelque chose, qu'il est possible de lui associer une interprétation dans un domaine d'explications possibles. « Comprendre le phénomène  $\Phi$  » signifie alors lui associer une interprétation dans ce domaine. « Expliquer le phénomène  $\Phi$  » signifie donner au destinataire de l'explication les moyens de construire cette interprétation. En revanche, pour

<sup>5</sup> Pour ne parler, bien sûr, que des auteurs qui prennent la peine de se fonder sur des analyses théoriques et d'en expliciter quelques uns des fondements...

comprendre l'énoncé décrivant le phénomène  $\Phi$ , il n'est, bien sûr, pas nécessaire d'être en mesure de construire une telle interprétation. Enfin, pour comprendre l'explication du phénomène  $\Phi$ , il est nécessaire, d'une part, d'interpréter l'énoncé utilisé pour fournir l'explication, et, d'autre part, d'associer cette interprétation au phénomène lui-même.

### Causalité de dicto et causalité de re

Un examen attentif de ce qui est accessible à notre appareil sensoriel convainc assez rapidement qu'aucune relation causale n'est directement observable : en toute rigueur, nous devons bien admettre que nous ne pouvons, dans les meilleurs des cas, qu'observer des faits, mais qu'il nous est impossible d'observer un lien causal entre deux faits. Toute causalité est donc *attribuée* par les observateurs ; ce qui, bien sûr, n'implique pas que toute attribution causale soit nécessairement fautive, mais seulement qu'elle n'est ni vérifiable ni réfutable sans le recours à des postulats, qui caractérisent le type de rationalité et/ou d'idéologie de ceux qui admettent cette attribution.

Présentation des deux concepts de causalité

Étant donné deux faits F1 et F2, nous parlerons d'attribution causale *de re* entre F1 et F2 de la part d'un sujet  $\beta$ , lorsque F1 est présenté par  $\beta$  comme une cause de F2 ; nous schématiserons cette situation par :

$$[F1 \text{ CAUSE}_r F2]_\beta$$

Ainsi, si  $\alpha$  demande à  $\beta$  pourquoi ce dernier est arrivé en retard et si  $\beta$  répond que c'est parce qu'il a raté le train, nous pouvons décrire la situation de la manière suivante :

F1 = le 'ratage de train' de  $\beta$

F2 = le retard de  $\beta$

Attribution causale :

$$[F1 \text{ CAUSE } F2]_\beta$$

Imaginons maintenant une situation dans laquelle  $\beta$  déclare qu'un individu prénommé Jean n'a pas été gentil et que  $\alpha$  lui demande pourquoi ; et supposons que  $\beta$  réponde que c'est parce que Jean n'a pas voulu lui prêter son livre. Si nous essayons de représenter la situation en suivant la même méthode que précédemment :

F1 = le refus de Jean

F2 = le manque de gentillesse de Jean

Attribution causale :

$$[F1 \text{ ~~CAUSE~~ } F2]_\beta$$

Nous obtenons quelque chose d'incongru : personne n'a prétendu *ni compris* que c'est le refus de Jean qui l'a rendu peu gentil... Le lien causal que  $\beta$  met en scène ne tient pas entre F1 et F2, mais entre F1 et le fait de penser ou de dire F2 : il s'agit d'une causalité *de dicto*, et non pas d'une causalité *de re*, comme dans le cas précédent.

Étant donné deux faits F1 et F2, nous parlerons d'attribution causale *de dicto* entre F1 et F2 de la part d'un sujet  $\beta$ , lorsque F1 est présenté par  $\beta$  comme une cause de la croyance que F2 ; nous schématiserons cette situation par :

$$[F1 \text{ CAUSE}_d \ll F2 \gg]_\beta$$

Munis de ce deuxième outil, nous pouvons décrire la situation de notre dernier exemple de la manière suivante :

F1 = le refus de Jean

F2 = le manque de gentillesse de Jean

Attribution causale :

[F1 CAUSE<sub>d</sub> « F2 »]<sub>β</sub>

Remarquons que ces deux types d'attributions causales, qui répondent toutes deux à la question « Pourquoi ? » correspondent, en général, à des chaînes causales opposées, selon le schéma suivant :

*Si le fait F1 me fait croire que le fait F2 a eu lieu, c'est parce je crois que F2 cause<sub>r</sub> F1.*

Ainsi, si le refus de Jean me fait croire qu'il manque de gentillesse, c'est parce que je crois que le manque de gentillesse cause<sub>r</sub> le refus.

Assimilation abusive...

On a donc la relation générale entre les causalités *de re* et *de dicto* :

**Si** [F1 CAUSE<sub>r</sub> F2]<sub>β</sub> **alors** [F2 CAUSE<sub>d</sub> « F1 »]<sub>β</sub>

Ce que nous avons appelé plus haut « le problème de la confusion *de dicto* – *de re* » est l'hypothèse cognitive suivante :

HC : Les êtres humains, quelle que soit leur culture et quel que soit leur niveau d'éducation ont tendance à prendre les attributions causales *de dicto* pour des attributions causales *de re*.

Cette confusion, qu'il est extrêmement facile de provoquer chez la plupart des êtres humains (et même chez des chercheurs en sciences de la cognition...), a pu être repérée dans une grande quantité de textes de divulgation scientifique et même, hélas, dans des textes académiques... Certaines propriétés apparemment irrationnelles des garants des argumentations, que l'on verra à propos de la réversibilité des *topoi argumentatifs*, ont une similitude intéressante avec cette inversion de la chaîne causale (on se gardera bien, maintenant, de dire qu'elles sont causées par – ou qu'elles causent – cette confusion...).

### 1.1 A propos des lois naturelles

On dit souvent que la Nature obéit à des lois; que les phénomènes naturels s'expliquent par des règles; ou encore que telle théorie explique tel ensemble de phénomènes. Tant que tout cela reste général et métaphorique, ces propos sont inoffensifs; en revanche, si quelqu'un se mettait à croire que la Nature obéit réellement à des lois (en ce sens qu'elle serait *obéissante*), on n'hésiterait pas à le considérer comme mystique ou fou (cette attitude est relativement récente: penser à l'horreur du vide dont on taxait la Nature il y a moins de deux siècles ...).

Pour ma part, je considère comme mystique non seulement l'attitude qui consiste à croire que la nature est obéissante, mais encore celle qui consiste à croire que les théories constituent des *explications* aux phénomènes naturels, dans le sens habituel du mot "explication". En effet, aussi bien la deuxième que la première de ces attitudes présuppose une métaphysique toute particulière, selon laquelle la Nature serait susceptible de *comportements* qui, dans le premier cas, seraient *dictés* par des lois (de nature divine ?) et, dans le deuxième cas, seraient tels que l'on pourrait en fournir des *raisons* (et peut être même des *causes*). Peut-on sérieusement affirmer que la loi de l'attraction universelle est une cause (ou même une raison) du mouvement de la Terre autour du Soleil, au même titre que l'on pourrait dire que la jalousie d'Otello a été une

des causes, une des raisons de la mort de Desdemone ? Il me semble que la seule façon de ne pas personnifier la Nature (avec tout ce que cela implique de mysticisme) est de considérer que les "lois naturelles" cessent d'être des lois lorsqu'on les examine du point de vue de la Nature et cessent d'être naturelles dès qu'on les considère comme des lois: s'il est nécessaire d'admettre que les *régularités* de la nature pré-existent à l'homme et sont indépendantes de leur 'découverte' par celui-ci, soutenir que ces régularités existent à cause d'une loi implique cette personnification de la nature qu'il me semble nécessaire de rejeter. Quelle loi contraindrait la Nature à obéir à telle ou telle loi? Et si l'on voulait en énoncer une, quelle autre loi contraindrait la Nature à se conformer à cette dernière loi<sup>6</sup> ? Le terme "loi naturelle" est bien commode et nous permet, par une métaphore simple, de parler d'un processus assez complexe; mais il ne faut pas prendre l'ombre pour la proie: ce que les expressions de la langue mettent en jeu n'a d'existence que dans la représentation que l'emploi de ces expressions suggère. Le fait que l'on puisse *parler* de lois naturelles n'est pas une caution de leur *existence*. Si l'on considère une théorie scientifique comme un ensemble de règles engendrant une *description* de certains aspects d'une classe de phénomènes et non pas comme *expliquant* ces phénomènes, on dispose d'une conception plus raisonnable des 'lois naturelles', conception que l'on pourrait qualifier de *de dicto*: ce que l'on appelle 'loi naturelle' est une règle qui explique la description d'un phénomène, dans le cadre de telle ou telle théorie, en ce sens qu'elle permet de *générer* cette description. L'examen des théories de la physique illustre bien la nécessité d'une telle prudence. Ainsi, si l'on considérait la loi régissant le mouvement linéaire uniformément accéléré, exprimée par la relation:

$$e = \frac{1}{2} gt^2 + v_0t + e_0$$

comme une *explication* de la trajectoire d'un objet lancé avec la vitesse initiale  $v_0$  au moment  $t=0$  et au point  $e_0$  de l'espace, on devrait aussi expliquer pourquoi cette relation régit effectivement ce type de mouvement; de fil en aiguille, on finirait par proposer comme 'explication' la loi de l'attraction universelle:

$$F = k \frac{M M'}{d^2}$$

à laquelle il faudrait encore trouver une explication, et ainsi de suite. Heureusement, les physiciens ne procèdent pas ainsi. Ces équations n'ont pas pour but d'expliquer les 'mystères' de la nature, mais d'expliquer comment la théorie décrit et prédit les phénomènes dont elle est censée 'rendre compte'. On voit encore une fois réapparaître le fameux adage d'Auguste Comte, selon lequel les questions scientifiques se formulent en "comment ?" et non en "pourquoi ?".

Lorsqu'on examine une théorie syntaxique de ce point de vue, le statut des règles syntaxiques apparaît aussi clairement que celui des lois de la physique: une règle syntaxique a pour fonction, non pas d'expliquer pourquoi la phrase A est grammaticale alors que la phrase B ne l'est pas, mais d'engendrer des descriptions linguistiques de la forme

La phrase A est grammaticale  
La phrase B n'est pas grammaticale.

---

<sup>6</sup> Bas van Fraassen a souvent insisté sur ce point. Voir sa discussion détaillée dans van Fraassen 1985.

Comme en physique, on peut s'élever d'un niveau théorique et formuler des lois qui engendrent des règles: les 'universaux' linguistiques sont aux règles syntaxiques ce que la loi de l'attraction universelle est à l'équation du mouvement rectiligne uniformément accéléré. Mais, comme en physique, les lois d'un niveau théorique supérieur n'ont pas de valeur explicative intrinsèque en ce qui concerne le monde réel.

### *De L'Adéquation Explicative*

L'impression d'"explicativité" produite par une théorie scientifique est extérieure à la théorie. Elle est due à une combinaison de plusieurs facteurs subjectifs, parmi lesquels:

- notre propension à croire que la nature obéit à des règles;
- notre tendance à ne pas distinguer facilement un phénomène de sa description.

La combinaison de ces facteurs est vraisemblablement opérée par le raisonnement abductif suivant:

Puisque la nature obéit à des règles et puisque la règle R engendre une description satisfaisante du phénomène naturel P, ce doit être à cette règle R que la nature obéit pour produire ce phénomène P.

(remarquer le glissement *de dicto* ---> *de re* de la première à la deuxième occurrence de R).

Cette remise en question du caractère intrinsèquement explicatif des théories scientifiques ne s'oppose qu'en apparence à la distinction chomskyenne entre adéquation descriptive et adéquation explicative: le fait que le caractère explicatif d'une théorie soit externe à cette théorie et que l'impression d'"explicativité" repose, en partie, sur des éléments subjectifs ne dispense pas de valoriser l'adéquation explicative par rapport à l'adéquation exclusivement descriptive<sup>7</sup>. Il ne semble pas exagéré de demander d'une théorie qu'elle suggère, même métaphoriquement, des réponses à nos questions métaphysiques, même si, d'un point de vue scientifique, ces questions sont mal posées. Les recherches sur l'adéquation explicative deviennent alors un domaine dans lequel la philosophie de la science, la philosophie de la connaissance et les sciences cognitives, et même l'éthique, doivent coopérer pour élaborer des réponses à la question: "Qu'est-ce qui compte comme explication pour l'Homme?". La nécessité de ce type de recherches se fait évidence lorsqu'on constate que, même dans les travaux de philosophie de la science, ces questions sont, presque toujours, soigneusement évitées, y compris par les auteurs qui valorisent explicitement l'adéquation explicative par rapport à l'adéquation descriptive. Ainsi, cette remise en question n'est pas seulement compatible avec les aspirations de la grammaire transformationnelle: elle prépare les bases d'une justification non exclusivement rhétorique de ces aspirations.

## 1.2 Rendre compte d'un phénomène

L'activité scientifique consiste en gros à construire et modifier des théories destinées à "*rendre compte*" de l'ensemble des *phénomènes* d'un *domaine*. Des hypothèses externes aux théories découpent le réel et constituent ainsi ces domaines:

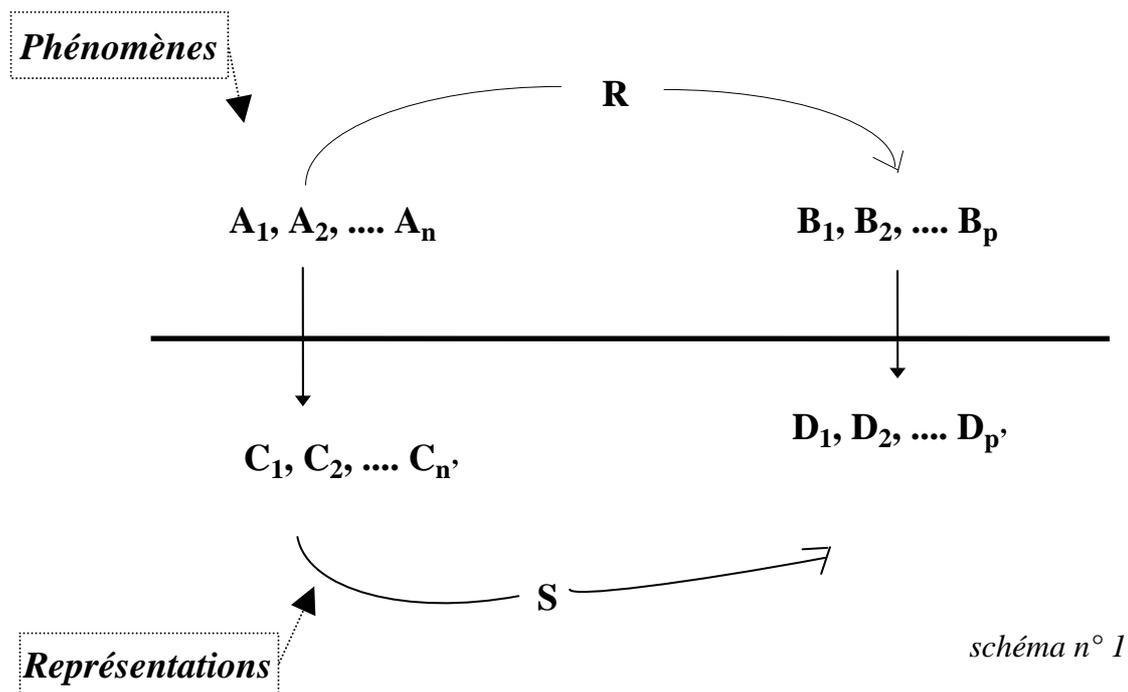
---

<sup>7</sup> La subjectivité de la notion explicative peut-être, d'ailleurs, la subjectivité des débats à son sujet...

ces mêmes hypothèses (externes) déterminent ce qui sera considéré comme un phénomène (pertinent) pour une théorie donnée. Une théorie *rend compte* d'un ensemble de phénomènes, en associant à chacun de ces phénomènes (entité ou événement) de cet ensemble une représentation abstraite des aspects du phénomène que la théorie se propose d'étudier. Pour que la description ait un quelconque intérêt scientifique, il est nécessaire que ces représentations satisfassent à certaines contraintes: on ne peut, par exemple, se contenter d'un catalogue (infini) d'occurrences de phénomènes associés aux représentations que l'on prétend leur attribuer. Les représentations des phénomènes doivent constituer un système *génératif* c'est-à-dire une structure contenant des représentations élémentaires et des règles qui permettent de produire des représentations complexes à partir de représentations élémentaires. Qui plus est, ces représentations complexes doivent correspondre, de façon régulière, aux phénomènes auxquels on prétendrait les associer directement.

Pour illustrer cette double exigence, imaginons que l'on s'intéresse à l'ensemble des phénomènes concernant le mouvement d'un petit objet. Supposons que l'on décide d'associer à de tels objets un nombre positif  $M$  (représentant la masse des objets), un vecteur  $X$  (représentant la position de leur centre de gravité dans un repère donné), à leur vitesse un vecteur  $V$ , et un nombre  $t$  au temps qui s'est écoulé depuis le début du phénomène, admettons en outre que  $X$ ,  $V$ ,  $t$  prennent respectivement les valeurs  $X_0$ ,  $V_0$  et  $0$  au début de l'expérience. Ces représentations élémentaires étant admises, le but d'une théorie de la trajectoire des petits objets sera de mettre en relation les valeurs possibles de  $M$ ,  $X$ ,  $V$ , et  $t$  de telle façon que les valeurs satisfaisant cette relation soient précisément celles par lesquelles l'observation empirique nous aurait conduit à vouloir représenter la position, la vitesse et le temps écoulé. Ainsi, supposons que cette relation nous permette, pour tout  $M$ , de calculer  $X$  en fonction de  $t$  et des valeurs  $X_0$  et  $V_0$  de ces paramètres et supposons en outre, par exemple, que pour une valeur  $t_1$  de  $t$ , la relation donne une valeur  $X_1$  de  $X$ . On s'attend alors à ce que l'observation empirique du mouvement "confirme" ce calcul, c'est-à-dire que, au moment représenté par  $t_1$  dans notre système de représentation, la position réelle du petit objet soit *effectivement* représentable par  $X_1$ . C'est en ce sens qu'il faut comprendre que la théorie *prédit* le résultat de l'expérience ou *rend compte* du phénomène.

Une théorie scientifique, pour décrire un ensemble de phénomènes, procède donc, d'une certaine façon, par analogie: si l'on représente les phénomènes  $A_1, A_2, \dots, A_n$  (dans notre exemple, les caractéristiques spatiales initiales du petit objet dont on étudie le mouvement), par  $C_1, C_2, \dots, C_n'$  (dans notre exemple, les valeurs initiales de  $X$ ,  $V$ , et  $t$ ) et les phénomènes  $B_1, B_2, \dots, B_p$  (dans notre exemple, les caractéristiques spatiales de l'objet en cours de mouvement) par  $D_1, D_2, \dots, D_p'$  (dans notre exemple, les valeurs satisfaisant la relation entre  $X$ ,  $V$ ,  $t$ ,  $X_0$  et  $V_0$ ), la relation que l'on suppose exister entre  $A_1, A_2, \dots, A_n$  et  $B_1, B_2, \dots, B_n$  est suggérée par la relation entre  $C_1, C_2, \dots, C_n'$  et  $D_1, D_2, \dots, D_p'$ . Cette analogie, *constitutive* des descriptions scientifiques, est une analogie de structure. La démarche scientifique établit ainsi un homomorphisme entre la structure phénoménologique et la structure du système de représentation (voir schéma n° 1).



## 2 La théorie et son domaine

### 2.1 Pour une théorie des théories scientifiques

#### 2.1.1 Hypothèses externes et hypothèses internes

Si l'on admet qu'une théorie -au moins descriptive- est un ensemble de règles qui engendrent des descriptions de phénomènes, on est amené à préciser, en amont de la théorie, ce qui sera, pour cette théorie, considéré comme un phénomène: il s'agit de délimiter le domaine d'application de la théorie. Cette délimitation, préalable à la construction de la théorie, constitue ce que Ducrot appelle les 'hypothèses externes' (Ducrot 1973), par opposition aux règles de 'calcul' proposées par la théorie, qui, elles constituent les 'hypothèses internes'. La délimitation du domaine se fait selon un certain nombre de critères de pertinence, qui constituent la stratégie de découpage de l'observable. Pour chaque phénomène retenu, il convient de préciser l'aspect du phénomène dont la théorie se propose de rendre compte. Il est intéressant de remarquer que, pour une stratégie de découpage donnée, les hypothèses externes d'une théorie proviennent souvent (peut-être toujours ?) des hypothèses internes, propres à une théorie plus ancienne ou mieux connue.

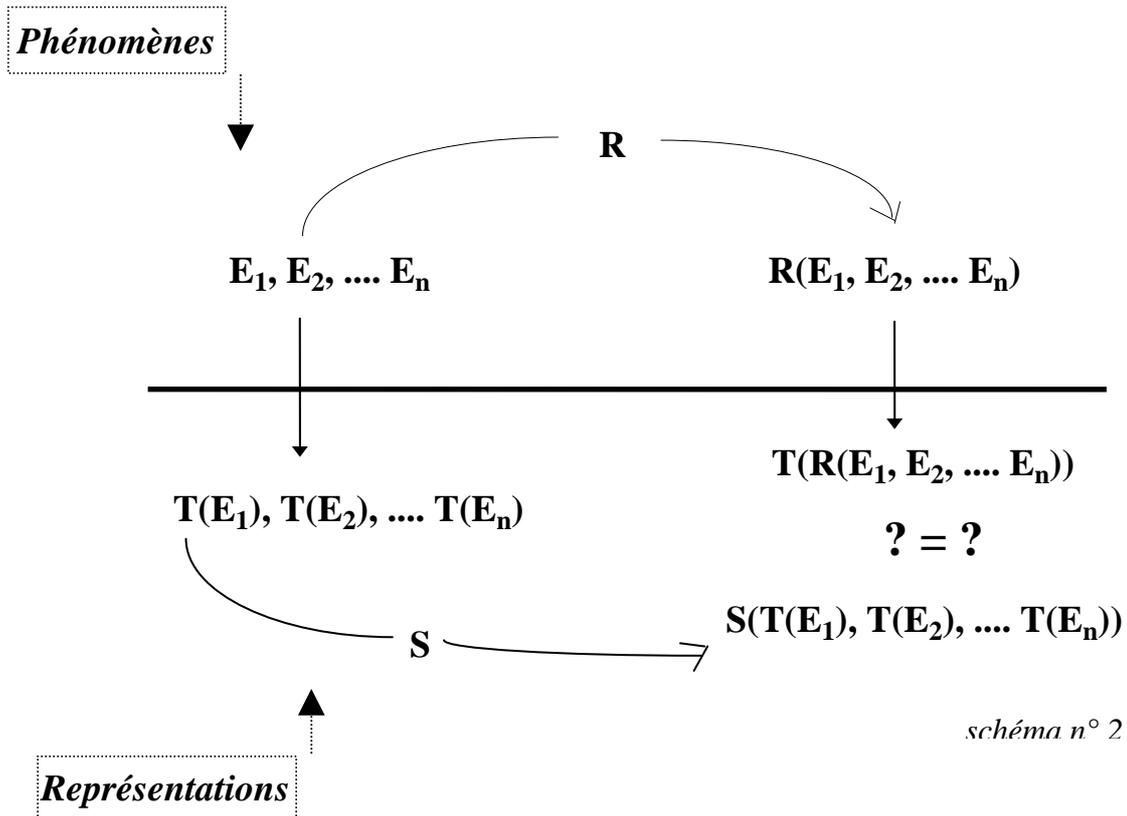
#### 2.1.2 Théories et systèmes de représentation

Ayant ainsi déterminé un ensemble de phénomènes pertinents -ou états-, au moyen d'hypothèses externes, on décrit les aspects de ces phénomènes qui sont jugés pertinents pour la théorie, au moyen d'un système de représentation. Les conditions et les règles de représentation sont externes à la théorie proprement dite: elles constituent une 'théorie de la mesure' pour cette dernière. L'ensemble constituent une 'théorie de la mesure' pour cette dernière. L'ensemble des hypothèses externes définissent des règles structurales reliant des états *simples* à des états *complexes* (voir schéma n[ 2]); la théorie proprement dite *génère* des représentations complexes par l'application d'opérations aux représentations d'états simples. Une théorie *prédit correctement* un phénomène

complexe lorsque la représentation générée par la théorie est identique à la description obtenue au moyen de la *théorie de la mesure*.

Ainsi, étant donné une théorie de la mesure permettant de représenter les aspects pertinents des états  $E_1, E_2, \dots, E_n$  et  $E$ , dans notre système de représentation par  $T(E_1), T(E_2), \dots, T(E_n)$  et  $T(E)$ , et étant donné une règle structurale  $R$ , qui agence une *expérience* permettant de produire l'état  $E$  à partir des états  $E_1, E_2, \dots, E_n$ , le contrôle expérimental de la théorie s'effectue par la comparaison de  $T(E)$  c. à d.:

$T(R(E_1, E_2, \dots, E_n))$ , avec la représentation générée par la théorie, c'est à dire  $S(T(E_1), T(E_2), \dots, T(E_n))$  (voir schéma n° 2).



### Empiricité et accessibilité des phénomènes

Dans le schéma ci-dessus, les phénomènes  $E_1, E_2, \dots, E_n$  et  $E$  interviennent sans que l'on se pose la question de savoir comment ils sont perçus par l'observateur, comme s'ils étaient, pour ainsi dire, directement accessibles à ce dernier. Il s'agit seulement d'une 'simplification didactique' : les phénomènes n'étant accessibles que par l'interprétation que nous faisons de ce que nos sens nous livrent, chacun des  $E_i$  est, lui-même, une représentation théorisée de phénomènes, qui, eux-mêmes, sont des représentations théorisées... comme l'illustre le schéma 3 :

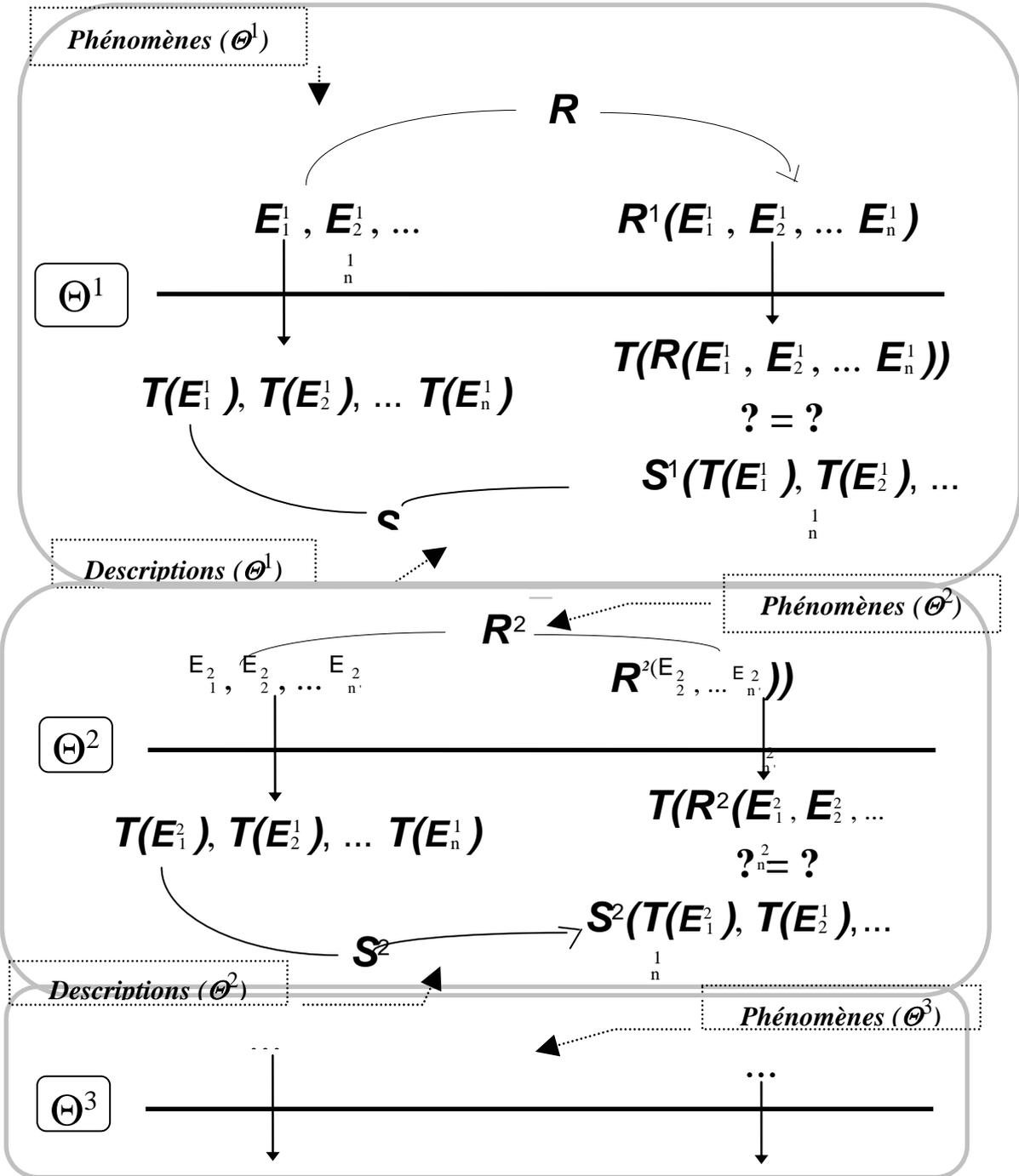


schéma n° 3

## 2.2 Deux exemples

J'illustrerai ce modèle des théories scientifiques au moyen de deux exemples. L'un concerne un petit domaine de la physique, qui a été évoqué plus haut: l'étude du mouvement d'un corps soumis à la pesanteur terrestre; l'autre concerne l'étude de la signification des phrases.

En ce qui concerne le premier exemple, on peut se référer au paragraphe 1.2 à ceci près que la relation entre phénomènes, établie par analogie avec la relation du système de représentation, qui met en rapport leurs descriptions, est ici remplacée par une relation structurale entre phénomènes: nous sommes maintenant dans une situation d'expérience, dans laquelle, en vertu d'une théorie déjà admise, les phénomènes -ou états-  $E_1, E_2, \dots, E_n$  conduisent à l'état  $E$ . La théorie que l'on teste par cette expérience permet d'une part de représenter les phénomènes  $E_1, E_2, \dots, E_n$  et  $E$  dans un système de représentation et, d'autre part, de générer, en vertu des règles du système de représentation, une description de  $E$  en fonction des descriptions de  $E_1, E_2, \dots, E_n$ . L'expérience renforce la théorie si la description générée à partir de  $E_1, E_2, \dots, E_n$ , est la même que celle qui est obtenue directement pour  $E$ . Dans notre exemple, les états initiaux correspondent aux coordonnées spatio-temporelles d'un corps solide et à sa vitesse au début de l'expérience; la relation structurale, à l'absence de forces s'exerçant sur ce corps; l'état résultant, à sa position et sa vitesse après un temps  $t$ . La mécanique classique permet d'une part, de représenter et de mesurer ces différents états et, d'autre part de 'prédire' l'état résultant en fonction de l'état initial, c'est à dire, de générer une description de l'état résultant en fonction de la description de l'état initial. En l'occurrence, la règle qui permet de générer cette description est l'équation classique du mouvement rectiligne uniformément accéléré. La comparaison des mesures de coordonnées et de vitesse à l'état résultant avec la description générée permet, dans ce cas, de renforcer la théorie puisqu'ici, la mesure et la 'prédiction' coïncident.

Je propose maintenant un exemple d'application de ce modèle aux sciences humaines: l'exemple concerne les théories épistémiques de la signification (voir au paragraphe 3 ce que recouvre l'opposition épistémique - procédural).

Supposons que l'on dispose d'une syntaxe convenable, que l'on ne remettra pas en question. Cette syntaxe fournit des règles permettant de construire des syntagmes à partir d'autres syntagmes -dits plus simples ou moins complexes-; certains de ces syntagmes sont des mots, d'autres sont des phrases. Supposons que notre syntaxe fournisse une règle  $R$ , qui permet de construire le syntagme  $E$  à partir des syntagmes  $E_1, E_2, \dots, E_n$ . Une théorie de la signification, fondée sur une telle syntaxe, devra disposer d'une part, des moyens de représenter la signification de  $E_1, E_2, \dots, E_n$  et  $E$  et, d'autre part de générer une représentation de la signification de  $E$  en fonction des représentations des significations de  $E_1, E_2, \dots, E_n$ . C'est ainsi -et seulement ainsi- que notre théorie de la signification permettra de 'calculer' la signification des phrases; les tests empiriques de ces prédictions consistent à comparer les résultats de ces calculs avec la signification que la théorie attribue directement. C'est ici que la comparaison avec la physique s'arrête: les phrases sont des entités abstraites: les seuls objets d'expérience (comparables aux quantités mesurables de la physique) sont les occurrences d'énoncés, c'est à dire les productions réelles des sujets parlants. Or, ce n'est pas de ces objets qu'une théorie épistémique de la signification s'occupe, mais bien des phrases (voir les précisions du paragraphe 3.2). Il reste donc à la linguistique une étape cruciale à franchir pour qu'elle puisse proposer des théories épistémiques de la signification qui soient réellement empiriques: fonder un analogue de la théorie de la

mesure en physique, qui permette de 'mesurer' la signification d'une phrase à partir des observations d'énoncés...

### 2.3 Fonctionnalisme et découpage de l'observable

Un critère de découpage de l'observable en domaines est issu de la thèse fonctionnaliste selon laquelle toute *explication* d'un phénomène peut et doit être formulée en terme du rôle que ce phénomène joue dans un système dont il est un élément.

La thèse fonctionnaliste est, on le voit, destinée à expliciter une des façons de concevoir ce qui vaut comme *explication* des faits. Si l'on développe le point de vue plus prudent esquissé plus haut, selon lequel les faits ne *s'expliquent* pas mais se *décrivent*, et qui précise que le rôle d'une théorie est d'engendrer des *descriptions* au moyen de *règles* (auxquelles on peut conférer une valeur explicative), on re-formulera la thèse fonctionnaliste en disant que la description d'un phénomène peut et doit être engendrée par des règles faisant intervenir le rôle que ce phénomène joue à l'intérieur d'un système plus complexe, qui constitue le cadre général de la théorie.

Cette thèse ne constitue qu'un critère de découpage: elle ne justifie pas à elle seule le "contenu" des domaines d'étude, mais constitue plutôt une méthode. En linguistique, on parlera ainsi de fonctionnalisme social (cf. Searle) ou de fonctionnalisme mental (cf. Harman) selon que l'on prendra en considération le rôle du langage dans la communication ou son rôle dans la construction de l'expression de la pensée.

Les stratégies de 'découpage' du réel en domaines, et, en particulier, la thèse fonctionnaliste, sont, le plus souvent implicites. Ce fait est responsable de bon nombre de controverses inutiles, dues à d'in vraisemblables confusions de domaines, dont j'ai analysé quelques exemples concernant les sciences du langage.(cf. Raccah 1984a et 1986).

La position que je souhaite défendre, sans être résolument antifonctionnaliste, s'oppose cependant à l'hégémonie du fonctionnel dans la valeur explicative des règles conduisant aux descriptions. En linguistique, en particulier, même si l'on s'intéresse au rôle des énoncés de la langue dans la communication, ou dans l'expression de la pensée, on ne peut faire l'économie d'une étude de la langue *per se*. Les descriptions ainsi obtenues pourront éventuellement être utilisées dans une description de telle ou telle fonction du langage, mais ne nécessitent nullement ces justifications. On ne peut donc exclure *a priori* d'étudier la langue en tant que code, indépendamment de l'opinion que l'on aura sur d'étudier la langue en tant que code, indépendamment de l'opinion que l'on aura sur ce à quoi elle sert.

## 3 Modèles épistémiques et modèles procéduraux

### 3.1 Analyse de processus vs analyse de structures:

Décrire un ensemble de phénomènes en décomposant cet ensemble en "modules" ou "composantes" dont on propose de faire l'analyse -que l'on se propose de décrire- d'une façon autonome est une pratique scientifique courante. Mais lorsqu'il s'agit d'étudier des domaines très complexes comme le sont deux qui concernent le comportement humain, cette pratique est, souvent à raison, très critiquée. La critique a souvent raison lorsqu'elle vise une position qui prétend décrire les *processus* qui régissent le comportement humain, mais, c'est du moins l'idée que je défendrai, elle est injuste lorsqu'elle s'adresse à une position qui tend à décrire les *structures* abstraites de ce qui est produit par ces comportements. En effet, un modèle procédural doit justifier

empiriquement les décompositions sur lesquelles il repose; si, pour décrire un processus, on le décompose en plusieurs phases, il n'est pas suffisant de disposer d'un modèle descriptif convenable pour chaque phase: il faut aussi montrer que ce processus est effectivement constitué de ces phases, car la décomposition est une hypothèse sur le processus, et pas seulement une commodité d'expression. En revanche, la décision de décomposer la description (épistémique) d'une structure selon différents aspects ne constitue pas une hypothèse sur la structure (et encore moins, bien sûr, sur les processus qui ont produit cette structure) et n'a donc de justification possible que dans la qualité de la description à laquelle elle permet d'aboutir. La tentation est grande d'adopter, sans justification empirique, un découpage qui convient pour la description d'une structure, et que l'on projette en quelque sorte sur les processus ayant conduit à cette structure, comme si la manière dont je perçois une structure était directement liée à la manière dont elle a été produite. Les erreurs méthodologiques liées à ce glissement sont fréquentes et conduisent parfois à des situations presque absurdes.

Prenons l'exemple d'une théorie linguistique connue: la grammaire transformationnelle. Selon cette théorie<sup>8</sup>, la syntaxe d'une langue naturelle est constituée d'une composante de structure syntagmatique ("Phrase Structure Grammar") et d'une composante transformationnelle, la première génère la "structure profonde" des phrases à partir des éléments lexicaux et la deuxième la "structure de surface" à partir de la structure profonde. Ce modèle, à l'origine épistémique, a conduit à deux erreurs méthodologiques analogues bien qu'utilisées dans des argumentations opposées: certains transformationnalistes ont pris ce modèle épistémique pour un modèle procédural et sont allés jusqu'à susciter des recherches neuro-linguistiques pour localiser "la composante transformationnelle" au niveau du cortex. Des expériences de mesure de temps de réaction ont été faites dans le but de montrer que deux "transformées" d'une même structure profonde étaient "plus proches" l'une de l'autre - d'un point de vue neuropsychologique - que des phrases de même signification ayant des structures profondes différentes. Aucune conclusion en ce sens n'a pu être tirée de ces expériences, certaines ayant donné des résultats positifs, mais d'autres, des résultats négatifs. Heureusement: la situation était déjà assez embrouillée sans que le hasard y mette son grain de sel. J'évoque le hasard car il eût pu se faire que pour des raisons liées ou non à celles invoquées par la grammaire transformationnelle, les processus d'interprétation des phrases fussent de nature à rendre certaines de ces expériences positives et aucune d'entre elles négatives; j'affirme que c'eût été un hasard, non que je critique la grammaire transformationnelle - en fait c'est probablement le meilleur modèle épistémique de la syntaxe - mais parce que ce modèle ne fait (et, en tant que modèle épistémique, ne *prétend faire*) aucune prédiction sur la *nature* des processus d'interprétation.

L'autre aspect de cette même erreur est illustré par la critique violente adressée à Chomsky et aux autres transformationnalistes (y compris ceux qui n'avaient pas de prétentions neurolinguistiques - et, heureusement, ils sont assez nombreux ...), selon laquelle le modèle transformationnaliste est faux parce qu'il n'a pas été trouvé de localisation corticale ni de structures psychologiques correspondant à la dichotomie structure profonde - structure de surface ou à l'opposition règle de ré-écriture -

---

<sup>8</sup> Il s'agit là d'une version ancienne de la théorie (remontant à 1965); les versions plus récentes proposent des analyses plus complexes, mais, en ce qui nous concerne ici, mon analyse reste la même: cette version de la théorie n'a été choisie que par souci de simplification.

transformation. Les présupposés sont les mêmes que pour les 'neuro-chomskyiens': la théorie originale subit une "transformation" qui la fait passer d'épistémique à procédurale. Toute théorie, même présentée comme devant influencer les recherches en neuropsychologie, n'est pourtant pas nécessairement une théorie du cerveau (ni de l'esprit ...).

### 3.2 Qu'attendre d'une théorie du langage?

#### 3.2.1 *Y a-t-il des langues?*

On pourrait multiplier les exemples de ce genre: il me semble cependant que l'idée est suffisamment illustrée. Il reste à savoir si le type de distinction que je propose est compatible avec le domaine de la linguistique. La question revient à celle de savoir si les sciences du langage se réduisent à l'étude de son usage ou si les langues peuvent être étudiées en tant que telles. En dernière analyse, il s'agit de se prononcer sur le concept même de langue. En effet, si l'on retient la première hypothèse, en excluant la possibilité d'étudier les structures des langues indépendamment de leur usage, on ne peut maintenir une entité qu'on appellerait la langue et dont la caractéristique est précisément de se démarquer de l'activité langagière, aussi bien conçue comme processus psycho-quelque-chose que comme fonction socio-quelque-chose ou comme complexe neuro-quelque-chose. Dans cette perspective, on est amené à adopter une position selon laquelle il n'y aurait pas de linguistique à proprement parler: tout discours sur la langue ne serait en fait qu'un discours sur son usage. Indépendamment des motifs de type corporatiste qui peuvent me faire condamner *a priori* -en tant que linguiste- cette position (il faudrait une certaine dose de cynisme à un linguiste pour qu'il soutienne que la linguistique n'a pas d'objet....) il me semble que l'on peut dire assez de choses utiles sur la langue pour que le fardeau de la preuve de son inexistence<sup>9</sup> incombe à ceux qui pensent que les sciences du langage ne sont en fait que des sciences du comportement langagier.

#### 3.2.2 *L'image du miroir et le mirage de l'image*

Il est indéniable, cependant, que l'on peut, sans attendre de disposer d'une théorie épistémique du langage qui soit satisfaisante, se servir des langues comme d'outils pour étudier les processus humains ou sociaux mis en jeu dans leur utilisation. En particulier, la langue, avec toutes ses structures abstraites, est le produit de l'homme et surtout, est constamment utilisée par celui-ci: rien ne peut être plus intéressant que de découvrir l'homme à travers sa langue. Mais il est indispensable d'éviter la confusion entre le reflet et son support: si l'on veut observer l'homme en se servant de son langage comme instrument d'observation (lunette ou miroir...), il faut se garder de prendre les dessins du support pour des tatouages<sup>10</sup> ! Une étude de l'instrument permet d'éviter cette confusion.

Une erreur méthodologique au moins aussi grave serait la confusion converse: penser décrire l'instrument alors qu'on exprime ce que l'on observe par son intermédiaire. Ainsi, si c'est en tant que linguiste que l'on prétend étudier le langage, toute description qui fait intervenir des connaissances sur le comportement humain (du biologique au social) doit être considérée comme entachée d'une erreur aussi évidente

---

<sup>9</sup> Il s'agit là de son inexistence en tant qu'entité abstraite et non pas, bien sûr, en tant que réalité psychologique ou sociologique.

<sup>10</sup> Puisque l'instrument est l'oeuvre de l'homme, on peut se demander pourquoi il y a apposé tel ou tel dessin, mais c'est encore un autre problème.

que celle qui consiste à inclure dans la description d'un miroir celle des objets dont il reflète l'image.

### 3.2.3 Pour une sémantique épistémique

Dans le cadre méthodologique, que je viens d'esquisser, je caractérise le domaine de la sémantique comme celui de l'étude des significations potentielles des phrases de la langue (considérée comme l'ensemble de toutes les phrases bien formées -engendrables par une syntaxe adéquate-). Une description sémantique d'une phrase est une représentation de sa signification. Il ne s'agit pas de dire que la fonction de la phrase P est de coder ou de véhiculer la signification X (la *fonction* de "Avez-vous l'heure ?" n'est pas de véhiculer par exemple un doute du locuteur sur le fait que son interlocuteur a ou n'a pas l'heure...): la phrase P avec sa signification X peut être utilisée pour de nombreuses fonctions. Une théorie sémantique est un ensemble de règles qui conduisent, pour chaque phrase P de la langue, à une description (que j'appelle *représentation*) de sa signification. On en retiendra les caractéristiques suivantes:

- compositionnalité de la description sémantique;
- autonomie de la sémantique par rapport à la connaissance de la situation et à la connaissance du monde: la représentation de la signification des phrases est suffisamment abstraite pour couvrir tous les *énoncés*;
- modularité de la sémantique: la représentation de la signification associée à l'information sur la situation et sur le monde permet l'interprétation du *sens* de l'énoncé;
- partialité: les modèles de représentation de la signification ne contiennent que les indications directement véhiculées par la phrase. Ce ne sont pas des modèles au sens de la théorie des modèles en logique.

Je n'ai pas encore examiné les répercussions possibles de cette 'théorie des théories scientifiques' sur l'enseignement des sciences, ni du point de vue expérimental ni du point de vue théorique. J'ai cependant la conviction, et les travaux de Martine Meheut (Meheut et al. 1987) me confortent dans cette idée, que l'acquisition des outils nécessaires à l'assimilation des théories scientifiques peut être grandement favorisée par des exercices mettant l'accent sur le maniement de systèmes formels de représentation, aussi bien en ce qui concerne les aspects *représentatifs* des systèmes (y compris la problématique de la mesure), qu'en ce qui concerne leurs aspects *génératifs*.

## REFERENCES

Chomsky, Noam (1965):

*Aspects of the theory of syntax*, Cambridge, MIT press.

Ducrot, Oswald (1973): *La preuve et le dire*, Paris, Mame.

van Fraassen, Bas (1985):

"Semantic approaches to laws of Nature", communication au colloque *Theories of Meaning* Florence, 4 Juin 1985.

Harman, Gilbert (1975):

"Language, thought and communication", in Gunderson, K (ed.): *Language, mind and knowledge*, University of Minnesota Press 1975.

Meheut, Martine et al. (1987):

"L'enseignement des modèles particuliers au collège permet-elle de développer des activités de modélisation?", communication aux IXèmes Journées internationales sur l'éducation scientifique, Chamonix, Février 1987.

Raccah, Pierre-Yves:

- (1984) "Où voulez-vous en venir?", communication au colloque de neuro-psycho-linguistique et pathologie du langage, St Riquier, Mai 1984.
- (1986) "Sémantique épistémique et loi de prédominance de l'argumentation", *Cahiers de linguistique française*, vol. 7, 1986 (actes du 3ème colloque de pragmatique de Genève).

Searle, John (1972): "Chomsky's revolution in linguistics", *The New York Review of Books*.