

- 19ème séance -

On a vu que la démarche utilisée ici, est une démarche qui tend à cumuler des observations de plus en plus précises de façon à vérifier les hypothèses de départ, ou à les transformer, ou bien encore à faire apparaître certains autres problèmes.

On a vu que cette procédure consiste à poser un certain nombre d'axiomes qui donnent par une série de règles déductives un certain nombre de représentations sous forme de formules sur lesquelles on applique des règles de réalisation. Ces règles sont des règles spécifiques à chaque langue et produisent un énoncé dans une langue donnée.

Et, cette procédure est une procédure qui se boucle dans la mesure où c'est l'étude extrêmement rigoureuse des langues qui permet d'enrichir les règles déductives d'une part, et d'autre part c'est à partir d'observations sur le langage qu'on peut poser des axiomes de départ ou termes primitifs.

Les règles déductives sont des règles qui permettent de tirer, par le calcul, des conséquences sous forme de formules d'un certain nombre d'axiomes, sans qu'on ait à faire appel à l'intuition pour rectifier l'imprécision.

Et, lorsque la partie axiomatique a atteint un certain degré de stabilité, elle doit pouvoir être enseignée; mais en ce qui concerne les règles, on se heurte toujours au fait qu'elles sont toujours, pour une langue donnée, pour une partie, des règles spécifiques et pour une autre partie, des règles généralisables.

Le généralisable, c'est ce qui peut être construit ou reconstruit à la demande selon les relations impliquées, par exemple lorsqu'il s'agit de certaines opérations concernant la relation qui existe entre la quantification, la qualification, l'aspect...; tandis que l'universel, c'est ce qu'on est obligé de poser, soit comme axiome, soit comme principe pour pouvoir rendre compte d'un certain nombre de phénomènes, par exemple lorsqu'il s'agit de certaines conditions sur l'énonciation comme la nécessité d'avoir deux énonciateurs.

Ici, on ne peut pas travailler comme en physique où on apprend à utiliser certaines procédures d'évaluation ou de mesure, certaines techniques, avec des domaines d'application immédiate; ici, on est obligé d'établir une relation entre les différents domaines qu'on étudie de manière approfondie, et le problème est d'autant plus complexe qu'on ne sait pas comment les problèmes s'articulent; par exemple, si on veut étudier de manière très restreinte le problème des tournures locatives dans des langues données, l'expérience montre que commençant par là, on tombe très rapidement sur des problèmes de modalité, d'aspect, de quantification... qu'il faut résoudre au moins en partie avant de continuer. C'est pour cela qu'une partie de la démarche consiste à dégager la forme que doit avoir le problème pour qu'il soit traitable, parce qu'il n'y a pas, en linguistique, toute une connaissance accumulée au cours des millénaires et qui serait matérialisée sous la forme d'une technique.

Toutes les connaissances se trouvent dans la faculté de langage sous la forme d'une glose épilinguistique, parce qu'on peut toujours gloser les textes, Et, par cette glose, on est dans une relation de langage extrêmement complexe parce que

- d'une part, la glose est un premier discours métalinguistique qui est intérieur à la langue dans laquelle on parle.
- d'autre part, dès qu'on reprend du texte à des fins d'observation on est de plain-pied dans le domaine métalinguistique.

Le logicien travaille uniquement sur le métatexte, tandis que HARRIS, par exemple, travaille à partir de gloses, c'est-à-dire de relations paraphrastiques non contrôlées, qu'il

essaie de contrôler, mais qu'il n'est pas toujours possible de contrôler. C'est que, en ce qui concerne les données en linguistique, il faut en distinguer plusieurs types qui correspondent à plusieurs domaines (voir "Considérations sur un Programme de Traitement Automatique des Langues et du Langage").

Ainsi, on distinguera:

. Les unités lexicales ("à propos de", "lapin", "tremblement de terre", "psyché"...) dont on fait l'analyse distributionnelle, c'est-à-dire qu'on trie et qu'on classe ces unités de façon à construire des classes d'équivalences qui ont certaines propriétés comme par exemple les tables de verbes construites par l'équipe de GROSS.

Mais, les propriétés ainsi dégagées ne peuvent pas être celles des classes d'équivalences que sont les familles paraphrastiques; c'est-à-dire qu'on ne peut pas faire l'analyse distributionnelle d'un énoncé, pas même d'une phrase. Déjà, avec les unités lexicales, lorsqu'on veut étudier les variations de plus d'une unité, l'analyse distributionnelle commence à poser des problèmes, parce que dans certains cas, suivant qu'on a telle conjonction, tel verbe, tel temps ou tel mode, cela donne des relations extrêmement complexes.

C'est d'ailleurs pour toutes ces raisons qu'on a intérêt à considérer qu'on a affaire à une lexis par rapport à une situation où on peut faire varier la valeur d'un certain nombre de paramètres, et dans ce cas on peut établir des relations; sinon, si on veut tout faire varier à la fois, on n'obtient aucun résultat; mais il arrive qu'on fasse des découvertes intéressantes, par exemple, si on prend un verbe comme "pleuvoir" et qu'on le combine avec l'imparfait, on sait que si on peut avoir:

*"Il pleuvait hier"*

on ne peut pas avoir

*"Il pleuvait demain"*

dont on dit que c'est un énoncé mal formé tout en étant un énoncé complet. Mais on s'aperçoit qu'avec "si", on va avoir :

*"S'il pleuvait demain"*

c'est-à-dire un énoncé dont on dit qu'il est bien formé, mais incomplet; "incomplet" parce qu'on oublie de dire que cela peut être un souhait.

Et par là, on se branche sur toute l'étude des modes et à ce moment on sort de ce qui est le fondement de l'analyse distributionnelle, qui est qu'avec telle unité physiquement représentative, on a ou on n'a pas telle autre unité.

Pour prendre un autre exemple de cette nature, par exemple les réalisations possibles ou non de "entrer et voir un revolver sur la table" avec "comme" ou "au moment où" ou "quand" en combinaison temps, on verra qu'on ne peut rien conclure quant à la distribution des termes "comme", "quand", "au moment où".

- Avec "comme" on pourra avoir

-L'imparfait : *"Comme il entrait dans la pièce, il vit un revolver sur la table"*

-Le présent historique : *"Comme il entre dans la pièce, il voit un revolver sur la table"*

Mais on n'aura pas :

-Le passé simple: *"Comme il entra dans la pièce, il vit"*

-Ni le futur: *"Comme il entrera dans la pièce, il verra"*

- Avec "au moment où" on aura:

-L'imparfait : *"Au moment où il entrait dans la pièce, il vit un revolver sur la table"*

-Le présent historique : *"Au moment où il entre dans la pièce, il voit"*

-Le passé simple : *"Au moment où il entra dans la pièce, il vit un revolver sur la table"*

-Le futur : *"Au moment où il entrera dans la pièce, il verra"*

- Avec "quand" on aura:

-Le passé simple : *"Quand il entra dans la pièce, il vit"*

-Le présent historique : *"Quand il entre dans la pièce, il voit un revolver sur la table"*

-Le futur : *"Quand il entrera dans la pièce, il verra"*

mais on n'aura pas :

-L'imparfait : *"Quand il entrait dans la pièce, il vit"*

C'est-à-dire qu'aucun de ces termes n'a la même distribution. Et, si on prend un verbe dont le mode de procès est différent, on aura encore autre chose, par exemple avec "marcher" qui est un prédicat qui n'indique pas l'inchoation, il faut en français marquer dans cet énoncé la borne gauche du procès par "se mettre à", et, on aura à ce moment-là des contraintes de "mettre" qui se combineront avec celles de "marcher".

Dans ce cas l'analyse distributionnelle va donner des milliers d'observations; mais il est difficile de décrire des phénomènes à l'aide de propriétés d'absence ou de présence, c'est-à-dire en termes d'opposition binaire.

Mais il reste qu'une description extrêmement bien faite est précieuse à tous les niveaux et permet de toute façon d'aboutir à un résultat qui permettra de vérifier, confirmer ou infirmer telle ou telle partie de l'ensemble théorique.

. Ensuite, il y a le métatexte.

On a affaire là, à du texte fabriqué en vue de rechercher un certain nombre d'opérations intéressantes concernant l'activité de langage.

C'est par excellence le texte que fabrique le logicien lorsqu'il pose des questions sur l'activité cognitive, les valeurs de vérité, la déduction, l'inférence, le fonctionnement de la logique naturelle. C'est aussi très souvent le genre de texte qu'on trouve dans les manuels, textes du genre :

*"Caroline regarde la fumée du calorifère"*

qui ne peut être autre chose que du texte fabriqué.

. Puis, il y a une glose épilinguistique.

La glose est une propriété très intéressante du langage humain. C'est en fait, ce qui fait proliférer le langage sur lui-même en prédisant sur du prédiqué, soit par la création de métaphores, soit par la définition; et JAKOBSON avait attiré l'attention sur cette propriété du langage qui est de définir des mots au moyen d'autres mots.

La glose est liée à l'activité épilinguistique des locuteurs (activité métalinguistique non contrôlée). Si on

prend par exemple un énoncé comme :

*"L'homme qui est venu hier portait un imper beige"*

c'est un énoncé qui a été glosé par

*"Eh bien, y a quelqu'un qui est venu hier, et ce quelqu'un, c'est un fait qu'il portait effectivement un imper beige"*

Ou encore si on considère :

*"Quiconque fera cela sera puni"*

on peut en glosant fabriquer tous les indéfinis qu'on trouvera dans différentes langues et on aura :

*"Tu peux prendre n'importe qui, celui que tu voudras, le premier venu ... de toute façon s'il fait ça, il est puni"*.

Ce genre de considérations, on ne le trouve pas en linguistique mais si on veut, on le trouve chez FREUD.

Et, la paraphrase se distingue de la glose en ce sens qu'elle renvoie à une activité réglée, donc contrôlée par l'observateur : "La linguistique construit des familles paraphrastiques, c'est-à-dire des classes d'équivalences. Pour cela on a besoin d'un système métalinguistique qui permette de représenter la dérivation paraphrastique. Ainsi, si on accepte l'équivalence de:

*"Il s'est vu condamné : à trois ans de prison ferme"*

*"Il s'est vu condamner à trois ans de prison ferme"*

*"Il s'est fait condamner à trois ans de prison ferme"*

*"Il a été condamné à trois ans de prison ferme"*

*"On l'a condamné à trois ans de prison ferme"*

il faudra, grâce à un système de représentation adéquat, marquer formellement l'équivalence, c'est-à-dire dégager les règles qui permettent de passer d'un agencement à l'autre et d'expliquer pourquoi ces agencements particuliers ont des valeurs référentielles (significations) équivalentes, à la modulation près dont il faut se rendre compte."

Enfin, il y a le corpus, qui est bien sûr un corpus contraint puisque sinon il n'a pas de réalité.

En effet, dans le cas des unités lexicales, on fait une analyse distributionnelle et on ne s'occupe pas de savoir

si on a affaire à un échantillon représentatif ou pas, puisqu'on peut l'enrichir au fur et à mesure des besoins.

Dans le cas du métatexte, le corpus n'a pas de raison d'être, puisque, s'il s'agit du logicien travaillant avec des valeurs de vérité, il se fabrique du texte tel que, par rapport aux lois qu'il s'est données, il travaille à l'intérieur de ce texte.

Dans le cas de la glose et de la paraphrase on est dans une situation expérimentale d'où la notion de corpus est exclue puisqu'on peut, par la construction faite à partir de ces textes, reproduire du texte.

Avec le corpus contraint, "l'expérimentation, au sens d'une reproduction, c'est-à-dire un passage réglé d'un système métalinguistique à une suite textuelle et vice versa, paraît difficile ou, du moins d'un autre ordre."

Ce seront soit des enregistrements d'un échantillon de langue qui n'a jamais été décrite et pour lesquels on dispose d'aucun moyen pour affiner et enrichir; soit des échantillons de langages d'enfants ou de psychotiques, puisqu'on n'a là, que les productions à disposition et aucun moyen d'induire des schémas qui permettraient de dériver des familles paraphrastiques et rien ne nous permet de dire qu'on a affaire à du texte homogène; soit du texte littéraire, puisque du point de vue strictement linguistique, on ne peut pas en faire grand chose. On peut Jouer sur la lecture en masquant ou en déformant le texte pour provoquer du texte de manière à étudier le premier texte. On peut aussi à titre proprement expérimental construire le système générateur du texte de façon à pasticher le texte. On peut encore analyser le texte d'un point de vue philologique ou stylistique. Mais c'est tout.

Et, en général, ces quatre ordres de problèmes sont, à cause de l'étiquette "linguistique", ramenés au même plan, comme si dans tous les cas il s'agissait de données.