

L&G ②

Une technique si simple...

*Quand on lui demandait son avis de quelque mot français, il renvoyait ordinairement aux crocheteurs du Port au foin
(RACAN, Vie de Malherbe)*

La linguistique (synchronique) telle qu'elle est (ou devrait être pratiquée) doit éliminer le plus possible les procédures sémantiques ou référentielles et simplifier les procédures d'analyse et de description. Il est des ouvrages de linguistique qui exigent l'adhésion à des constructions complexes – qui découlent **donc** de la langue dans laquelle elles sont formulées.

Or il faut rendre l'analyse des langues aussi peu arbitraire ou artificielle que la compétence du locuteur natif naïf.

Le plus compliqué en linguistique – BLOOMFIELD ¹l'a bien noté – est le commencement :

2. 1. The most difficult step in the study of language is the first step. Again and again, scholarship has approached the study of language without actually entering upon it. Linguistic science arose from relatively practical preoccupations, such as the use of writing, the study of literature and especially of older records, and the prescription of elegant speech, but people can spend any amount of time on these things without actually entering upon linguistic study. As the individual student is likely to repeat the delays of history, we may do well to speak of these matters, so as to distinguish them from the subject of our study.

Passons par-dessus tout « ça » ; supposons un ensemble d'item (de fragments détachés ou isolés) d'une (supposée) langue : un corpus brut à analyser ; supposons encore que tout ça est « écrit » de gauche à droite; imaginons ce (très petit !) corpus, trouvé (par un Martien ?) dans une poubelle juste après la fin de l'année scolaire, dans une petite classe d'un petit village !

¹ Language, 1933, by Holt, Rinehardt and Wiston, Inc.

1.	Le chien mange le chat.
2.	Le chien mange son os
3.	Le chat mange la souris
4.	Il regarde son os.
5.	Le chien mangera son os
6.	Le chat regarde la souris
7.	Le chien regarde le chat et l'os.
8.	Il mange le chat.
9.	Il mange la souris.
10.	Il regarde le chat.

En superposant 1. et 2 ; on isole trois unités (temporaires !) :

a) *Le chien mange*, b) *le chat*, c) *son os*.

Les combinaisons validées sont : a) + b) ou c)

3. permet de séparer a) en trois unités : *Le* mange, *chien* et *chat*,
ou *Le chien*, *Le chat* et *mange* d'une part, *la souris* de l'autre.

En 1. et 3. *le chat* n'occupe pas une place identique par rapport à *mange* ; la fréquence de *le*, supérieure encore à celle de *mange* permet de penser qu'il s'agit d'une forme « grammaticale » et non seulement lexicale.

4. (par superposition avec 2.) apporte une unité : *Il regarde*.

5. permet, par superposition avec 2. d'isoler *-ra*, *fragment* de *mangera* sur *mange*, et *variante libre* (qui commute avec ∅ (ensemble vide))

6. permet de segmenter 4. *Il regarde* : *Il* commute avec 4., mais il pourrait exister une combinaison (croisée) que 9. écarte .

1. et 8., 3. et 9. permettent de poser *Il* comme variante de *Le chien* ou de *Le chat* (ou de *Le + chien* ou *chat*).

8., 9., 10., par exemple, permettent de confirmer que *mange* commute avec *regarde* et *le chat* avec *la souris*, et 6. permet d'avancer que *Il regarde la souris* est un item possible, et donc que *la souris* est une variante libre de *le chat*, etc.

En fin d'analyse nous aurions les unités suivantes :

Le chien, Le chat, Il (qui a une grande fréquence d'apparition) (« à gauche »)

son/l' os, la souris, le chat (qui a une fréquence d'apparition, en deux positions, très grande) (« à droite »)

mange, regarde

-⊙/-ra

et

On peut déjà esquisser la combinatoire (la grammaire de ce corpus) :

Le ou *Il* « commencent » tous les item.

Si *chien* ou *chat* sont « avant » *regarde* ou *mange*, c'est *le* ;

chien et *chat* sont-ils remplacés par ⊙ ? *Il* commence l'item.

A la suite de *regarde* ou *mange* (+ -⊙ ou ra) on peut rencontrer *le chat* ou *la souris* ou *l'/son os*. Et suivi de *l'os* est une variante libre de ⊙ après *le chat* : *il commute donc à ce point avec ⊙*

Evidemment l'analyse même sert à formuler des hypothèses que de nouveaux corpus (corpora ?) viendront vérifier ou invalider. Par exemple, la constatation que *son* commute avec *l'* devant *os* suggère l'hypothèse de *son* devant *chat*, *chien* ou *souris*. Cette hypothèse sera vérifiée pour *chat* et *chien* et infirmé pour *souris*, mais une nouvelle segmentation *s*+*on* ou *a*, permettra de lier *-on* à *chat* et *chien*, *-a* à *souris* : *s* + *on/a*. Et, plus tard, en effectuant la même opération sur *le/la* : *l* + *e/a*, on établira *l'/s+a* . . . !

Il est évident aussi que de nombreux traits de ce corpus limitent son exemplarité, mais, par contre, rendent possible son exploitation rapide et pédagogique. Globalement nous constatons avoir pu isoler et assembler des unités qui ont un « contenu » (sémantique) et une « expression » (graphique ici) : c'est une analyse de « première articulation ». Rien ne nous interdit de poursuivre cette analyse pour arriver à isoler et analyser les unités plus petites – que nous savons ici être des « lettres » - mais dont nous pourrions étudier la distribution, la fréquence, les agrégats, etc., pour la description du corpus au niveau d'une « deuxième articulation »

On lit souvent que la langue est un système de signes (signifiants et signifiés : S^a/e) analysables à leur tour en signes pourvus d'un signifiant et non-pourvus de signifié propre, mais qui permettent de passer d'un signifié x à un autre signifié (z) : $S^a/éy \Rightarrow éz$