

## L&G ①- Linguistique

On va essayer ici de montrer, pas à pas, *comment la / linguistique / peut analyser des corpus écrits et oraux* : la formulation adoptée ici est – volontairement – modeste, semble-t-il ; elle nous semble, au contraire, très ambitieuse, et loin « en avant » des réalités universitaires.

La « linguistique », en fait, est un ensemble de pratiques et de théories, qui prétendent s'opposer ... ou ignorer la « grammaire » ; elle (la linguistique) affirme regretter le caractère normatif des grammaires, alors qu'elle serait descriptive, elle rejette des questions comme celles de l'origine des langues, ou de leur unité, s'attachant – ainsi que le veut la tradition – à la diversité et à l'égalité des langues (égalité : une revendication toute théorique quand on constate que les linguistes autoproclamés (et/ou universitairement « reconnus » ! ) ne publient qu'en anglais et, secondairement, dans *quelques langues-égales-aux-autres...* mais bien supérieures (en rentabilité) aux langues africaines orientales, ou autres.

Nous n'allons pas faire, aujourd'hui l'histoire ni le réquisitoire adéquats.

Nous avons écrit */linguistique/* : nous voulons montrer comment se présente la technique « pure » de description qu'est (ou que devrait être) la linguistique, et le minimum (indispensable) de conventions préalables (et générales) à toute analyse (linguistique). C'est ce que signifie – stricto sensu – *linguistique*.

Schématiquement, l'analyse vise à recenser, dans un corpus, des segments minimaux (et à en faire la liste) et à en décrire les combinaisons rencontrées (avant de les présenter exhaustivement (et hiérarchiquement)).

Nous allons prendre des exemples faciles et tenter de les exploiter, tout en notant les présupposés que nous formulons au fur et en mesure.

■ Soit le « corpus suivant », ensemble d'item séparés par un silence : #

(1) #lemuresthaut#

(2) #lemurdelamaisonesthaut#

En superposant les deux item (1) et (2), on constate

que ce corpus est composé de deux segments : #lemuresthaut#

et &delamaison& marquent l'absence de silence

& et & sont des « suppressions » de # et #

Dans la suite (de gauche à droite #lemurdelamaisonesthaut#, on peut

ajouter & (ou &&, mais nous accepterons que ## équivaut à # et && à &)

#lemurdelamaisonesthaut# est – du fait de #lemuresthaut #

#lemur&delamaison&esthaut# et

#lemuresthaut # est #lemur&esthaut#,

Il existe donc dans ce corpus deux unités et trois segments :

deux unités, donc &delamaison& et #lemur&esthaut # et trois

segments : (1<sup>er</sup>) #lemur& puis (2<sup>ème</sup>) &delamaison& puis (3<sup>ème</sup>)

&esthaut# (de gauche à droite).

La combinatoire s'énonce donc ainsi : le premier segment peut être suivi du second mais doit être suivi du troisième ou les deux item se terminent par le troisième . 1<sup>er</sup> et 2<sup>ème</sup> ou « rien » et 3<sup>ème</sup>

A partir de ces deux possibles : 1<sup>er</sup> + 2<sup>ème</sup> + 3<sup>ème</sup> et 1<sup>er</sup> + 3<sup>ème</sup> on peut reconstituer le corpus : 2<sup>ème</sup> commute avec « rien ».

■ Soit une suite du corpus :

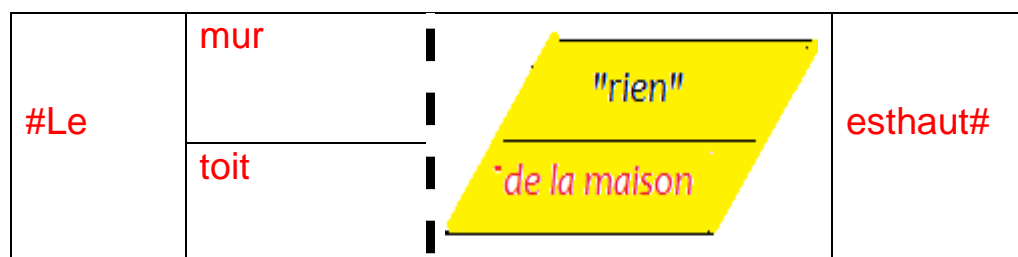
(3) #letoitesthaut#

(4) #letoitdelamaisonesthaut#

En superposant (1) et (3) et (2) et (4) on note que **mur** et **toit** commutent, et il faut donc écrire **#le&toit&esthaut#** et **#le&toit&delamaison&esthaut#** et, de ce fait **#le&mur&delamaison&esthaut#**.

Mais cette présentation masque le fait que si **&toit&** commute avec **&mur&**, **&esthaut#** est constant (comme, d'ailleurs **#Le**.

**&** note une coupure-un collage, mais non la hiérarchie des possibles que l'on va tenter de représenter ici



De gauche à droite, nous avons

#le&mur&esthaut#  
 #le&mur&delamaison&esthaut#  
 #le&toit&esthaut#  
 #le&toit&delamaison&esthaut#

**#le... ..esthaut#** est présent dans tous les item

**&mur&** commute avec **&toit&** et, *ensuite*, **&delamaison&** commute avec « rien » ... ou **&** !

Chaque variante (**mur** et **toit**) a une variante « à droite » : **delamaison** ou « rien ». Les éléments inséparables ou insécables sont

**#le&**, **&mur&**, **&toit&**, **&delamaison&**, **&esthaut#** : **&mur&** et **&toit&** sont commutables (donc incompatibles dans la suite nommée « item »).

A la suite (*à droite*) de **&mur&** ou **&toit&**, **&delamaison&** est facultatif (ou commute avec « rien »).

- Ajoutons deux nouveaux item, rencontrés ... par hasard ( ?!) dans le corpus soumis à analyse.

#les&murs&delamaison&sonhauts#

#les&toits&delamaison&sonhauts#

Si on superpose ces item aux item précédents

#le&mur&delamaison&esthaut#

#le&toit&delamaison&esthaut# on remarque

#les&murs&delamaison&sonhauts#

#les&toits&delamaison&sonhauts#

	&mur&s&			
#le&s&		&delamaison&	&est/sont	haut&s#
	&toit&s&			

&s apparait comme une variante de « rien » dans les colonnes 1, 2 et 5. La présence de &s est corrélée au le passage de **est** à **sont** (ou, pour être plus rigoureux **s..t** étant constant **&e** passe à **&.on** ). Ce que nous connaissons comme le « pluriel » n'est que le nom grammatical donné à ce type de phénomène, illustré dans ces exemples extrêmement nombreux ; dans un vaste corpus, la variété corrélée en &s de segments distincts : &s ... &s &s (+d'autres variétés de segments) apparaissent dans des item dans lesquels ils ne constituent qu'un choix unique (soit « rien » soit &s ...&s .....&s. C'est un signe unique (mais réalisé fractionné) manifesté à des points différents de l'item (on dit aussi, monème ou morphème disjoint)

L'analyse linguistique est aussi - stricto sensu - d'abord *distributionnelle*. Elle admet un certain nombre de *présupposés "logiques"*,

⇒ il existe une **linéarité** ("ordinaire") de minima séparables ("une ligne formée de points"):

a, b, c, d,... (de a à d..) ≠ a, c, d, b, ≠ c, a, b, d,...

⇒ il existe des minima "**semblables**" qui s'opposent aux "**différents**"

a = a ≠ b ≠ c ...

⇒ il existe une possibilité de **mettre en équivalence** une unité et un ensemble d'autres unités:

a, b, c, ... ≡ P où P - a ≡ x

...

Nous les ferons remarquer en cours d'analyses.

Ce que nous venons de présenter est un exemple de présentation possible : nous pouvons déjà, montrer « là » où nous avons dépassé, en cours de route, les limites de ... l'évidence.

Il s'agit de

#les&murs&delamaison&sonhauts#

#les&toits&delamaison&sonhauts#

Si on superpose ces item aux item précédents

#le&mur&delamaison&esthaut#

#le&toit&delamaison&esthaut# on remarque

#les&murs&delamaison&sonhauts#

#les&toits&delamaison&sonhauts#

Nous faisons figurer en surlignage jaune ce qui pose problème, que nous sommes tentés de régler à cause de notre connaissance de la grammaire.

Nous voyons effectivement le passage de « rien » à &s& (#les&murs,

hauts#, mais &est « passe » à &ont... parce que nous le savons : en fait, s restant identique, c'est &e t qui passe & ont

Et donc, linéairement &e passe à &« rien » puis s puis & « rien » passe à &on puis t.

Il n'est pas facile de se débarrasser de vieilles habitudes : au moins faut-il signaler ces sauts dans le supposé « connu », qui ne sont que l'application à des faits nouveaux de recettes anciennes. En fait c'est sur la base d'un paradigme que nous avons mémorisé que s'effectue le passage &est > &ont ! Ce paradigme est celui de la « conjugaison régulière » (en fait la flexion de « *noms conjugués* » comme *chanter*) : 3<sup>ème</sup> singulier / 3<sup>ème</sup> pluriel, soit –e / -ent ; une fois établi (avec *danser, monter, nager, farder, passer, cintrer*, etc), ce paradigme « épaulé » par les « pronoms personnels » (*je, tu, il, elle, nous...* etc., qui ne sont pas des « pronoms », pour la plupart, d'ailleurs !) englobe *est / sont* dans son filet.

Pour « éliminer » cette compétence, nous pourrions choisir un corpus ...  
*de cartes.*

Soit : (*de gauche à droite*)

#♠♣♥#

#♠♣♦#

#♠♣♥♦#

#♠♣#

En superposant ces item, on note que tout item est au minimum #♠♣

& suivi de &♥# ou de &♦# ; et, si &♥ on peut (&♥)&♦#.

L'application de ces « lois » recrée les item.

Il n'y a pas qu'une formulation de la description !

Nous pouvons choisir aussi un corpus de langue « étrangère » (de gauche à droite) :

#umhest#, #fráhestum#, #fráhesti#, #tilhests#, #umhesta#, #tilhesta#.

En superposant ces items, nous vérifions que #&hest&# est constant (& pouvant être suivi de « rien ») nous le représentons par ○

#um○#, #frá○um#, #frá○i#, #til○s#, #um○a#, #til○a#.

- à #um& (« à gauche ») « correspondent » à &# ou à &a# (à droite)
- à #frá& (« à gauche ») « correspondent » &#frá& um# ou &i# (à droite)
- à #til& (« à gauche ») « correspondent » &s# ou &a# (à droite)
- à #um& et #til& peut correspondre (« ou ») &a# : #um& et #til& sont donc des « variantes » de &a# ; en résulte que &# et &s# sont en distribution complémentaire.

Cette analyse considère implicitement que #um& est distinct de &um#, ce qui est une possibilité envisageable, mais ... à noter !

L'analyse isole donc

une constante (présente dans tous les items) &hest&

et (« + »)

ou (« / ») un « possible gauche » #frá& couplé à

soit &um#

soit &i#

ou (« / ») un « possible gauche » #um& couplé à

soit &#

soit &a#

ou (« / ») un « possible gauche » #til& couplé à

soit &s#

soit &a#

Cette analyse décrit les combinaisons attestées et, à partir des éléments isolés, ici au nombre de 7 (hest, frá, um, til, frá, um, til) il est facile de reconstituer les items en « supprimant » les suites & par insertion de &O, tous les items « commençant » et « finissant » par #.

On voit déjà les difficultés de formulation si le souci de rester le plus rigoureux possible persiste !

Dans les manuels de langue ou les ouvrages de grammaire, la formulation semble plus aisée et accomplie ; ce « corpus » est issu de la première leçon d'un cours d'islandais <sup>1</sup>

Icelandic nouns have, then, four cases: nominative, accusative, dative and genitive. In order to show these endings it is convenient to use the following system, taking as our example the noun *hestur* (horse):

<i>nom.</i>	<i>hér er hestur</i> (here is (a) horse)
<i>acc.</i>	<i>um hest</i> (about a horse)
<i>dat.</i>	<i>frá hesti</i> (from a horse)
<i>gen.</i>	<i>til hests</i> (to a horse, on horseback)
<i>nom.</i>	<i>hér eru hestar</i> (here are horses)
<i>acc.</i>	<i>um hesta</i> (about some horses)
<i>dat.</i>	<i>frá hestum</i> (from some horses)
<i>gen.</i>	<i>til hesta</i> (to some horses)

Notice that the prepositions *um*, *frá* and *til* take the accusative case, the dative case and the genitive case respectively.

I

Les « supposés connus » sont énormes (*noms, cas, nominatif, accusatif, datif, génitif, désinences*) et peut-être pas exempts d'erreurs grossières dans leur exploitation...

---

<sup>1</sup> Glendening, P., J., T., *Teach Yourself Icelandic*, 1961, E.U.P.