



**Linx**

Revue des linguistes de l'université Paris X Nanterre

**60 | 2009**

**Nonne scripta manent ?**

---

## Grammaire et dictionnaire : le parti pris des mots

Annie Bertin

---



### Édition électronique

URL : <http://journals.openedition.org/linx/672>

DOI : 10.4000/linx.672

ISSN : 2118-9692

### Éditeur

Presses universitaires de Paris Nanterre

### Édition imprimée

Date de publication : 1 mai 2009

Pagination : 31-46

ISSN : 0246-8743

### Référence électronique

Annie Bertin, « Grammaire et dictionnaire : le parti pris des mots », *Linx* [En ligne], 60 | 2009, mis en ligne le 30 mars 2012, consulté le 27 mai 2019. URL : <http://journals.openedition.org/linx/672> ; DOI : 10.4000/linx.672

---

Département de Sciences du langage, Université Paris Ouest

# Grammaire et dictionnaire : le parti pris des mots

*Annie Bertin*

*Université Paris Ouest Nanterre La Défense- UMR 7114 MoDyCo*

## Introduction

Grammaire et dictionnaire constituent, aussi bien dans l'héritage académique que dans le champ linguistique du 20<sup>e</sup> siècle, des voies antithétiques ou, à tout le moins, disjointes pour aborder les faits linguistiques. A l'Académie Française dont il fut le premier protecteur, Richelieu « proposa de travailler premièrement à un Dictionnaire de la Langue Française & ensuite à une Grammaire, à une Rhétorique & à une Poétique ». Dans ce programme, le dictionnaire occupe la première place, soit qu'il fût la tâche la plus facile à réaliser, soit qu'il fût prééminent. La linguistique générative qui a dominé la seconde moitié du 20<sup>e</sup> siècle a, de fait, renversé cet ordre, faisant des règles syntaxiques d'une Grammaire Universelle, l'objet privilégié des recherches sur le langage. La lexicographie, relevant de la linguistique appliquée, est doublement dévalorisée : outre sa visée pratique, voire utilitaire, et non théorisante ou explicitement modélisante, son objet (le mot) peut sembler une unité scientifiquement discutable voire non pertinente, condamnant à une étude descriptive d'un certain discours, sans jamais atteindre les mécanismes sous-jacents au plan de la langue. Le problème est redoublé si le dictionnaire inclut les mots grammaticaux, qui semblent défier la définition. De fait, nombre d'entreprises lexicographiques contemporaines renvoient à une seconde étape, qui peut rester hypothétique, le traitement de ces mots. Toutefois l'opposition du lexical et du grammatical est battue en brèche à l'heure actuelle dans la linguistique

cognitive, et tout particulièrement à travers les nombreuses études sur la grammaticalisation, la lexicalisation et la dégrammaticalisation. On voudrait donc ici s'interroger sur la possibilité et l'intérêt qu'il y aurait à traiter les mots grammaticaux sous forme de dictionnaire. On tentera à cette fin de cerner le statut du mot grammatical dans un dictionnaire (1), puis le statut du dictionnaire dans le savoir actuel (2), pour avancer des propositions d'un renouvellement lexicographique prenant pour objet-test les mots grammaticaux (3)<sup>1</sup>.

## 1. Le mot grammatical dans le dictionnaire

La tradition grammaticale occidentale, depuis Aristote, classe les mots en partie du discours, selon une catégorisation qui varie en nombre et, sur certains points en terminologie, mais qui établit invariablement une hiérarchie mettant en avant le substantif et le verbe, au détriment des autres, jugées moins nécessaires à la complétude de la phrase. On est ainsi amené à distinguer des mots pleins et des mots vides, des mots lexicaux et des mots grammaticaux. Si les notions mêmes de grammaticalisation et de lexicalisation interdisent d'établir une frontière absolue entre mots grammaticaux et mots lexicaux (Mélis et Desmet, 1998, Prévost & Fagard, 2007), il existe bien tendanciellement une dichotomie entre les deux catégories, qui a pu être appréhendée dans des cadres théoriques divers. On peut donc se demander si les mots grammaticaux relèvent bien d'une approche lexicographique, à plusieurs titres : 1) par leur vacuité sémantique ou du moins la mobilité de leur sémantisme, 2) par l'impossibilité d'en donner une définition, 3) par l'instabilité de leur unité.

### 1.1. Un sémantisme flou

Confronté à cette problématique à propos du mot *rien*, Martin (1966) pose par exemple : « Le lexique s'édifie à partir de traits significatifs particuliers, la grammaire s'organise autour de catégories générales. Le lexème tient sa substance de sèmes autonomes, le morphème de catégories qui l'englobent. L'un a une substance indépendante, l'autre une substance commune à tout un paradigme. Aussi l'un paraît sémantiquement plein, l'autre sémantiquement vide. » (p. 5) ; tout en reconnaissant le caractère asymptotique de la recherche de « critères que l'on voudrait infaillibles. Pure illusion, mais qui nourrit les rêves d'un grammairien » (*ibid.*), l'auteur dégage cependant des tendances sémantiques et syntaxiques distinctes :

Universalité<sup>2</sup>, spécificité, nécessité, structure binaire et cinétique<sup>3</sup>: tels nous semblent les caractères fondamentaux de la substance grammaticale. Si l'on ajoute, en grammaire,

---

<sup>1</sup> Pour une première approche, notre réflexion prendra ses exemples dans les catégories de la conjonction et des indéfinis.

<sup>2</sup> La notion est entendue au sens combinatoire de la faculté de se combiner avec tout membre d'une partie du discours.

<sup>3</sup> La binarité correspond aux deux tensions dégagées par G. Guillaume, selon un mouvement de pensée qui va de l'avant à l'après ; B. Pottier définit ainsi ce qu'il dénomme des tensèmes : « Le tensème I est un *avant*, il pose, il fonde, il particularise ; le tensème II est un *après*, il dispose, il exploite, il généralise » (Pottier, 1964, p. 113, cité in Martin, 1966, p. 7). Le cinétisme serait selon

la tendance à l'abstraction et à la généralité, manifestée dans le discours par les hautes fréquences ; si l'on considère par ailleurs, les fonctions actualisatrices, relationnelles ou représentatives de la substance grammaticale ; si l'on tient compte enfin d'une structure qui situe le morphème à l'intersection de catégories classificatrices générales, on obtient un ensemble de données qui en autorisent au moins une définition approchée. (*ibid.*, p. 7).

On retiendra la caractérisation du grammatical par la mobilité du sens qui, dans un cadre guillaumien, s'ordonne en un double mouvement du général au plus particulier puis du particulier au général.

Le grammatical semble s'opposer au lexical par sa vacuité ou du moins son flou sémantique. Ainsi la conjonction relève-t-elle des catégories floues (Harris, 1986) à plusieurs titres. Du point de vue sémantique, tout d'abord, définir une conjonction fait question soit qu'on la considère comme un mot vide de sens propre, qui ne fait qu'explicitement une relation logique existant indépendamment, soit qu'on définisse ce sens comme provenant non de la conjonction en elle-même (comme ce serait le cas pour le substantif, le verbe voire l'adverbe) mais de sa combinaison avec d'autres éléments de la phrase. Lorsque la conjonction (*sundesmos*) apparaît dans la *Poétique* d'Aristote, elle est d'emblée marquée de négativité quant au sens :

La conjonction est un constituant de la parole dépourvu de sens qui est sans influence, ni positive, ni négative, sur la formation d'une unité signifiante de la parole résultant de la combinaison de plusieurs constituants, qu'on mette cette conjonction aux extrémités ou au milieu ; de plus, c'est un constituant qui ne peut se tenir seul en tête d'un énoncé(...). Ou encore, c'est un constituant de la parole dépourvu de sens, apte à créer une unité signifiante de la parole à partir d'au moins deux constituants qui soient eux-mêmes signifiants. (Aristote- *Poétique*, 20, 1456b 38-14657a 6)

Dans le chapitre qu'il lui consacre dans son *Traité*, Denys le Thrace la définit comme un terme qui ne fait que signaler un sens préexistant dans le contexte :

La conjonction est un mot qui conjoint la pensée en ordonnant, et qui révèle l'implicite de l'expression. (éd. Lallot, 1989, p. 65)

La grammaire médiévale définit conjonctions et prépositions comme des « syncatégorèmes », des mots qui n'ont pas de sens précis et propre, puisque leur sens dépend non du mot lui-même mais de sa combinaison avec d'autres mots de la phrase. La conjonction, qui est de l'ordre du « surplus », de l'accessoire, ne relève pas dès lors de la syntaxe de la phrase à proprement parler<sup>4</sup> mais tient davantage de mécanismes tout à la fois cognitifs et discursifs : dans la tripartition des arts du langage, plus qu'au champ de la grammaire, son étude appartiendrait à ceux de la logique et de la rhétorique. Liée à des valeurs sémantiques qui semblent bien constituer des universaux cognitifs

---

Guillaume une caractéristique sémantique du mot grammatical : « ...les mots les plus abstraits de la langue, comme l'article et certaines prépositions sont des signifiants liés à un signifié non pas statique mais cinétique, emportant avec lui une image non pas d'immobilité mais de mobilité. » (Guillaume, 1964, p. 176 note 6, cité par Martin, 1966, p. 7).

<sup>4</sup> Le « renouvellement des conjonctions » (Meillet, 1915) relève d'ailleurs des évolutions lexicales et non des évolutions syntaxiques selon R. Posner (1997, p. 348) : « The evolution of new subordinating conjunctions however is, I think, better treated as lexical change ».

(temps, cause, concession), elle est pourtant loin d'avoir l'univocité d'un signe logique, et se caractérise souvent par la pluralité de sens qui se recouvrent plus ou moins distinctement.

On pourrait certes proposer de traiter en conséquence dans le dictionnaire uniquement les mots lexicaux, les mots grammaticaux relevant d'une description des emplois à mener dans une grammaire. Une telle répartition rejoint somme toute le sens commun comme la tradition la plus académique. Pourtant l'approche dictionnaire du mot grammatical peut se défendre en tant qu'elle oblige à s'interroger sur les relations entre des homonymes grammaticaux (*si* adverbe et *si* conjonction introduisant l'hypothèse ou l'interrogation ; *que* adverbe, relatif et conjonctif, par exemple) que la présentation grammaticale disjoint dans des chapitres différents. L'ambivalence sémantique se double dans de nombreux cas d'une polyvalence fonctionnelle, qui met en question la validité d'oppositions comme celles de coordination *vs* subordination, complément nécessaire *vs* ajout, conjonction *vs* préposition/adverbe.

La pluralité sémantique, catégorielle et fonctionnelle de la conjonction amène nécessairement à s'interroger sur la nature de cette pluralité : homonymie d'items disjoints ? polysémie ? On devra s'interroger sur la modélisation possible de cette diversité, qu'on la pense en termes de prototype et d'air de famille, de subduction guillaumienne ou de grammaticalisation faisant correspondre termes source et terme cible, selon des parcours sémantiques réguliers, en allant du moins grammatical au plus grammatical.

## 1.2. Un défi à la définition

La lexicographie, obligeant à penser la polysémie et la polycatégorialité est donc une pratique fructueuse pour une bonne connaissance de la langue. Reste que la représentation du sens sous la forme traditionnelle d'une définition de dictionnaire est inadéquate. L'arbre porphyrien qui procède par subdivision du genre à l'espèce se heurte à la généralité de ces mots : de quoi *quand* ou *tout* peuvent-ils être une spécification ? Comment définir *quand* autrement que par le renvoi très général au domaine du temps, ses divers usages grammaticaux (interrogation, subordination), les sens logiques qui peuvent en dériver par implicature contextuelle ? On peut faire le recensement des contextes syntaxiques qui correspondent à des valeurs différentes, avec une attention particulière aux temps verbaux ; ainsi présentant la méthode lexicographique qui préside à la rédaction du *TLF*, Robert Martin écrit-il :

Les monèmes [mots grammaticaux], supports de catégories bien plus que de sèmes, sont directement appréhendés au moyen de l'analyse formelle. La première tâche du rédacteur grammairien consiste à dégager les types syntaxiques dans lequel fonctionne le vocable étudié. Ainsi, le morphème *quand*, versé dans la catégorie des adverbes interrogatifs, se combine avec tel ou tel tiroir grammatical pour livrer tel ou tel effet de sens ; versé dans la catégorie des adverbes conjonctions de temps, il présente une distribution semblable mais glisse imperceptiblement vers des emplois logiques (causalité, supposition, concession...). (Martin, 1969, p. 54)

On pourrait s'interroger sur les difficultés engendrées par des distinctions catégorielles comme celle posée ici entre adverbe et conjonction de subordination,

dont de nombreuses études ont montré les défauts, de manière générale<sup>5</sup> ou à propos de tel mot particulier<sup>6</sup>. On retiendra surtout la question de l'articulation entre les emplois divers en discours et la valeur générale en langue :

Une fois établi l'inventaire complet des situations syntaxiques, on en vient à formuler une hypothèse sur le contenu en langue du vocable étudié. Pour le mot *quand*, par exemple, elle pourrait être que ce morphème de temps suggère un espace en soi virtuel, mais qui tend à se restreindre à un espace minimum, c'est-à-dire à s'actualiser de manière à situer dans le temps le procès qui s'y inscrit ; tantôt, dans l'emploi adverbial, l'espace verbal de *quand* est l'objet d'une interrogation et ne s'actualise que dans la réponse qui y est donnée, tantôt, dans l'emploi conjonctionnel, cet espace est fourni par le procès subordonné qui coïncide, du moins partiellement, avec l'espace du procès principal. (Martin, 1969, p. 54)

Une définition du mot grammatical qui en subsume les usages est-elle possible en lui attribuant un « signifié de puissance » en termes guillaumiens, plus généralement un sens ? Cette question engage plus largement la question du rapport entre sens et usage, le mot grammatical, par son caractère non référentiel, ne faisant qu'exacerber la question et la rendre incontournable. D'un mot qui « révèle l'implicite de l'expression », selon la formule de Denys le Thrace, que confirmeraient des études de détail, peut-on dire qu'il a un sens ?

Le mot grammatical exhibe ainsi la place prépondérante du contexte et de l'usage dans la construction du sens, mise en lumière, selon des points de vue différents, par la sémantique interprétative de Rastier ou les tenants d'une *corpus-based grammar* (Kemmer & Barlow, 2000). Si l'on peut penser que le dictionnaire doit passer « de la description du sens à celle des usages » (Véronis, 2004, p. 37), on constatera que le mot grammatical, en tant que « syncatégorème » y oblige.

### 1.3. Une unité instable

L'une des différences entre mots lexicaux et mots grammaticaux serait que les seconds appartiennent à un paradigme fermé, l'un des critères du degré de grammaticalisation étant précisément la réduction de la variabilité paradigmatique (Lehmann, 1985, 1995). Du fait même du processus de grammaticalisation qui est à l'origine de nombre de mots grammaticaux, par coalescence voire fusion d'éléments précédemment autonomes, la perception de certains syntagmes en tant que constituant une unité peut faire question : si *beaucoup* n'est certainement plus perçu comme constitué de la séquence épithète+ substantif homonyme de *il a tiré un beau coup de fusil*, le statut de *tout le monde* en français actuel comme syntagme grammaticalisé en pronom indéfini fait question (Schneidecker & Capin, 2010) et l'on peut se demander s'il est encore pertinent d'en rendre compte sous le mot *monde*. L'émergence de mots grammaticaux, souvent réalisée à partir de la coalescence des éléments d'un syntagme, s'observe dans tout état de langue, selon une fréquence plus ou moins grande sans doute selon les époques. Dans un état de langue non encore standardisé comme le moyen français, on

---

<sup>5</sup> Cf. par exemple Brill (2006), Muller (1996).

<sup>6</sup> Cf. par exemple, les études sur le mot *comme* in Moline & Flaux (2008) et Desmets & alii (2008).

constate que décrire la conjonction *pendant que* amène à prendre en compte les formes *ce pendant que/ cependant que/ pendant ce / le/ certain/ lequel/ ledit temps que/ pendant ce que*, toutes attestées dans la *Base de Moyen Français*.

S'agissant des locutions conjonctives dont l'unité est stabilisée dans l'usage, on constate des incohérences de lemmatisation qui marquent la difficulté à décider du choix des unités pertinentes en tant que mot-vedette. Dans le *Robert*, *tandis que* et *alors que* figurent en mot-vedette mais non *pendant que* ou *avant que* ; dans le *TLF*, on constate que *parce que* est traité comme une entrée propre, mais *pour que* l'est sous le mot *pour*, *bien que* est traité sous *bien* alors que *éh (he) bien* dispose d'une entrée distincte. La lemmatisation des mots grammaticaux invite donc à prendre en compte la dynamique sous-jacente au réseau des relations qu'entretiennent ses unités.

#### 1.4. Conclusion

Mettant en cause la possibilité d'une définition, le rapport du sens à l'usage, la nomenclature des termes de la langue, le mot grammatical constitue un objet paradoxal pour le lexicographe, qui met à l'épreuve la rigueur de sa démarche. On doit cependant se demander si plus généralement cette démarche est productive.

## 2. Le dictionnaire dans le savoir actuel

La notion même de dictionnaire peut paraître désuète en ce 21<sup>e</sup> siècle commençant, aussi bien en tant que modèle intellectuel sous-jacent qu'en tant qu'objet producteur de savoir. Pourtant, dans le temps même où le dictionnaire semble voué à être submergé par les connaissances encyclopédiques auquel donne accès la communication informatisée, on peut constater, dans le champ des études linguistiques comme dans celui du traitement informatique, un retour d'intérêt pour les unités lexicales, légitimant une approche lexicographique.

### 2.1. La mise en cause du dictionnaire

Il est certes plus facile de définir les mots *ageot* ou *huître*<sup>7</sup> que *assez* ou *puisque* : dans un cas le mot est associé à une ou des images mentales, dans l'autre, il passe par des notions abstraites de quantité/qualité ou de rapport causal. La dévalorisation des mots grammaticaux dans la catégorisation grammaticale issue de la tradition aristotélicienne a partie liée avec une conception de l'intellection dépendante de l'image : « Quelle image y-a-t-il sous l'intellection de conjonctions comme *si* ou *et* ? » dans un tel cadre, se demande un spécialiste de la philosophie médiévale du langage (Jolivet, 1982, p. 54-55).

Ce primat du visuel dans la conception de la connaissance, qui augmente la difficulté de définition des mots grammaticaux, mais aussi de toute lexie non

---

<sup>7</sup> On reconnaîtra dans ces exemples une allusion à la poésie de Francis Ponge, à laquelle Jacques Anis consacra plusieurs articles d'inspiration sémiotique et génétique. *Le parti pris des choses* témoigne que le poète est peut-être le plus efficace des lexicographes, du moins pour les mots désignant des objets concrets, qui posent des problèmes de définition inverses de ceux des mots grammaticaux, mais également redoutables.

descriptive, est sans doute pour une bonne part responsable de la mise en cause du dictionnaire comme modélisation du savoir, au bénéfice d'une représentation encyclopédique, qui permet non seulement de classer par subdivision *a priori* du genre à l'espèce, mais de reconnaître, de distinguer, en retenant les différences significatives (Wilson, 1967, Eco, 1984).

Cette remise en cause du dictionnaire en tant que modèle, opérée dans les années 60 et 70 du XX<sup>e</sup> siècle, s'est doublée, avec l'informatisation des ressources de connaissances<sup>8</sup>, d'une remise en cause du produit. Les organisateurs d'une Journée d'étude consacrée aux « Perspectives nouvelles en lexicographie » constataient fin 2009 :

Les dictionnaires papier ont de plus en plus de mal à trouver un public et leurs versions électroniques sont loin d'être des succès commerciaux. Pour s'informer sur un mot, ce n'est plus le *Petit Robert* mais Wikipédia ou Google qui deviennent le premier choix d'un nombre croissant de personnes. On pourrait penser qu'il ne s'agit que d'une question financière mais le problème semble plus complexe ; le *TLFi* qui est consultable gratuitement n'est pas devenu pour autant une ressource incontournable pour le grand public.

La lexicographie traditionnelle n'a peut-être pas su réaliser son changement de support, les dictionnaires électroniques n'étant parfois guère plus qu'une version numérisée des dictionnaires papier.

Corbin et Casiglia (2009) font le même constat quant à l'atonie des projets lexicographiques d'envergure, qu'ils soient publics ou commerciaux, particulièrement en France, la lexicographie britannique se montrant un peu plus innovante, par la prise en compte de modèles linguistiques récents, sans toutefois adapter véritablement le dictionnaire au nouveau mode de communication informatisé.

Hors du domaine français, le Symposium international organisé à l'initiative du *Center for Lexicography* de l'Université d'Aarhus, réunissant des spécialistes d'origines géographiques variées (danois, hongrois, chinois, brésiliens) part également du constat d'une crise d'identité, liée à l'écart croissant entre les modélisations de la langue par la linguistique théorique et la description lexicographique disponible, aux bouleversements induits par le passage au support électronique et plus largement aux mutations amenées par la « société de l'information et de la connaissance » dans les formes d'accès à l'information (Tarp, 2009, p. 17). La lexicographie se trouve de ce fait, selon l'équipe d'Aarhus, à un carrefour, si bien que les orientations qui seront prises auront des conséquences sinon irrémédiables du moins décisives. Dans la perspective fonctionnelle que défendent ces chercheurs, il est tout d'abord nécessaire de prendre conscience de l'autonomie de l'activité lexicographique par rapport à la linguistique, en tant que le dictionnaire, quel que soit le champ qu'il couvre, est spécifiquement un outil, devant donc satisfaire des besoins qu'il convient d'identifier, en termes de situations dans des activités sociales. La lexicographie se justifierait alors pleinement en tant que s'intégrant aux recherches visant à maîtriser la pléthore d'informations, en termes de rapidité et de pertinence d'accès d'une part, d'intégration raisonnée des informations pour constituer un savoir d'autre part.

---

<sup>8</sup> Cf. également Béjoint (2007).

Parallèlement l'évolution des méthodes et des conceptions linguistiques semble de nature à remettre au premier plan le mot, ou du moins, les unités lexicales pertinentes.

## 2.2. Le retour du lexique

Si la langue se conçoit comme une combinaison d'unités signifiantes, sa description par le linguiste peut mettre l'accent soit sur les signes soit sur la dimension combinatoire, ce qui revient à accorder une place prééminente soit au lexique soit à la syntaxe.

Au primat de la règle syntaxique, dont ils déplorent qu'elle domine le champ linguistique, Mel'čuk *et al.* (1995), dans le cadre de la théorie Sens-Texte, opposent le rôle premier du lexique :

Les lexies ainsi conçues forment la partie primordiale de la langue. En exagérant quelque peu, on pourrait même dire que l'ensemble des lexies EST la langue. En effet, une langue est constituée de lexies et de règles servant à la manipulation de ces dernières. Les règles qui réunissent les lexies en syntagmes, les syntagmes en phrases, et les phrases en discours sont donc nettement secondaires par rapport aux lexies- en ce sens que leur nature et leur forme sont déterminées par les lexies.

On peut résumer ces affirmations par la comparaison suivante. L'ensemble des règles d'une langue L, c'est-à-dire sa grammaire (au sens large du terme), est aux lexies de L ce que l'ensemble des instructions d'assemblage d'un meuble en pièces détachées est à ces pièces. Les instructions d'assemblage sont obligatoirement écrites en fonction des pièces à assembler. Il en est de même pour les règles d'une grammaire : étant donné qu'elles nous permettent d'assembler des lexies, elles doivent être formulées en fonction des lexies ! Il s'ensuit que LE LEXIQUE D'UNE LANGUE PRIME LOGIQUEMENT SA GRAMMAIRE. Rappelons cependant que ce n'est nullement le point de vue dominant dans la linguistique théorique actuelle, qui met un accent tout particulier sur la syntaxe et relègue le lexique au rôle d'un accessoire négligé, qui doit assurer les « interactions lexicales » après la construction des arbres syntaxiques. (Mel'čuk, Clas, Polguère, 1995, p. 17).

Cette revalorisation du lexical rejoint une revalorisation des unités au détriment de la règle, qui informe les recherches actuelles dans divers secteurs de la linguistique. Dans le domaine du TAL, par exemple, l'ambition pratique première de traduction automatique a été supplantée, dans les années 60, par un retour à la recherche fondamentale sur la modélisation, puisque l'ordinateur, premier objet susceptible de manipuler des symboles en dehors du cerveau humain, est considéré comme devant changer le niveau d'analyse des langues, tout comme le microscope a révolutionné la biologie (ALPAC, 1966, p. 121, in Cori, 2008, p. 96). Le TAL a donc été centré, des années 60 à 90, sur l'analyse syntaxique automatique, en relation avec les grammaires syntagmatiques. Or, le développement, à partir des années 90 de la micro-informatique et d'internet a changé la donne, amenant au développement d'un TAL dit robuste, qui ne prétend pas constituer une description linguistique de la langue ou du langage, mais vise à traiter, avec le minimum d'erreurs, des occurrences attestées. Le TAL n'a plus pour but l'écriture de règles mais il s'attache au traitement de segments de discours selon des règles contextuelles ou statistiques. La réévaluation de l'unité et du contexte

qui en découle se retrouve, selon un parallélisme qui est sans doute loin d'être fortuit, dans des modèles linguistiques basées sur l'usage, comme celui développé par Langacker (1988, 1990, 2000), Kemmer et Israël (1994), Kemmer et Barlow (2000). Contrairement à la division saussurienne entre langue (abstraite) et parole (concrète par ses attestations), dans laquelle les énoncés sont l'actualisation de règles qui les précèdent, la langue, dans ces modèles, est constamment (re)construite dans l'immanence des énoncés et de leur production. De ce fait, la linguistique ne peut qu'être descriptive, sensible au contextuel qui modèle l'usage, aux constructions spécifiques de certains items lexicaux, qui, par extension et analogie, modèlent la syntaxe. La description d'occurrences attestées, en relation avec les traits discursifs et culturels qui façonnent leur énonciation, ainsi que les réorganisations catégorielles et sémantiques (re)deviennent l'accès obligé aux mécanismes linguistiques dans leur généralité.

### **2.3. Conclusion**

On peut conclure des points précédents que, dans le contexte épistémologique actuel, la lexicographie, à condition de prendre en compte les mutations technologiques, peut constituer une méthode privilégiée de description du langage. Il reste à définir les objectifs et les contenus d'un tel renouvellement. On ne fera dans le cadre de cet article qu'esquisser quelques pistes.

### **3. Pistes pour un renouvellement de la lexicographie**

La lexicographie actuelle, dans la mesure où elle traite du langage sur la base de corpus informatisés, amène à engager le débat sur la place respective que doivent y occuper les sciences du langage d'une part, les sciences et techniques de l'information de l'autre. On a pu défendre une lexicographie totalement indépendante de toute contrainte utilitaire et visant à une description exhaustive d'un côté (Mel'uk *et al.* 1995) ou bien, à l'inverse, une lexicographie totalement émancipée de la linguistique, conçue en termes fonctionnalistes, comme un outil devant satisfaire les besoins d'utilisateurs qu'il convient d'identifier préalablement, ces besoins étant eux-mêmes variables selon les situations, ce qu'il convient également d'identifier. Sans laisser l'ingénierie linguistique la reléguer au rang de prétexte, ce type de lexicographie, défendue par l'équipe d'Aarhus, met l'accent sur le caractère intrinsèquement utilitaire de la discipline, ce qui permettrait de l'intégrer dans un ensemble de disciplines dont le but commun est l'accès efficace à l'information.

Pour notre part, suivant une voie moyenne, et quels que soient les jugements qui s'attachent à cette posture, nous défendons une activité lexicographique qui se constitue en fonction d'un public à définir et qui participe de la constitution du savoir, mais qui se nourrit de la recherche linguistique qu'elle peut en retour enrichir.

Participer à la constitution du savoir dans une société dite de l'information et de la connaissance suppose, nous semble-t-il, de satisfaire à une double exigence : la maîtrise d'une surabondance d'informations et le souci primordial du dynamisme de l'objet langue et de l'outil qui vise à en rendre compte.

### 3.1. La maîtrise des corpus

L'informatique met à disposition des hommes contemporains une masse d'informations jamais égalée, imposant un double défi : l'accès rapide à l'information pertinente, la validation de la qualité de l'information. Dans le cas des mots grammaticaux, dont bon nombre sont de haute fréquence et d'emplois diversifiés, donnant lieu à des articles complexes et étendus, la consultation sur écran rend sans doute encore plus inopérante la longueur des pages consacrées aux prépositions *à* et *de* dans le dictionnaire traditionnel. On perçoit bien alors la nécessité de repenser l'article de dictionnaire pour en accroître l'efficacité, par une aide à l'extraction d'information.

Certes le savoir ne se confond pas et ne se réduit pas à l'information. C'est un point très utilement mis en exergue dans la *Déclaration de Heidelberg*<sup>9</sup> qui fixe les objectifs de qualité auxquels s'engagent bon nombre de lexicographes européens, engagés dans un travail d'ordre historique ; l'article 7 (*Savoir versus information*) pose :

Le savoir présuppose la connaissance et l'intelligence du cadre dans lequel prennent place les informations collectées ; c'est par l'intégration dans un système scientifique que les informations sont transformées en savoir. L'utilité des informations, augmentées par la mise en œuvre des réseaux électroniques, est garantie seulement par l'accroissement du savoir.

La lexicographie a un rôle de validation, d'organisation, d'intégration de l'information en un savoir. Ceci ne nous paraît pas contradictoire avec la prise en compte des innovations que peuvent apporter les recherches portant sur *l'information and data accessology* aussi bien dans le traitement de la documentation<sup>10</sup> que dans la gestion des descriptions que l'on en tire<sup>11</sup>.

### 3.2. La langue comme objet dynamique

La lecture structuraliste du *Cours de linguistique générale (CLG)*, ou plutôt sa « scolarisation » (Anis, Chiss, Puech, 1988) qui, particulièrement en France, a marqué la pratique institutionnelle des études linguistiques, a tendu à mettre en avant le point de vue synchronique sur la langue, en tout cas à distinguer radicalement les points de vue synchronique et diachronique dans l'étude linguistique. La notion de grammaticalisation, développée depuis les années 1980, même si elle ne constitue pas un dispositif radicalement original, a mis en valeur, par sa dénomination même, la porosité de l'opposition entre le lexical et le grammatical mais aussi celle des plans synchronique et diachronique (Bat-Zeev Shyldkrot, 1995). Ce cadre est particulièrement apte à rendre compte des faits d'auxiliarisation, de l'évolution des morphèmes de négation et des indéfinis et du renouvellement des articulateurs temporels, logiques ou discursifs (Traugott & Heine 1991, Lehmann, 1995, Marchello-Nizia, 2006 *inter alia*). Les travaux menés autour de la grammaticalisation se doublent plus récemment de l'étude des mécanismes de dégrammaticalisation (Norde, 2009), par quoi le morphème grammatical

---

<sup>9</sup> Cf. La *Déclaration de Heidelberg*, à l'occasion du colloque « L'héritage culturel européen et la lexicologie du XXI<sup>e</sup> siècle », sur le site du DEAF (<http://www.deaf-page.de/deaff.htm>).

<sup>10</sup> On peut penser à des traitements automatiques comparatifs des constructions de deux mots comme aide à la rédaction d'article du type du *sketch engine* développé par Kilgariff *et alii* (2004).

<sup>11</sup> Sur ce point cf. Martin (2007) et *infra* 3.3.

revient au lexical, selon un processus qui n'est pas si exceptionnel qu'on l'a dit, dans un premier temps (Ramat, 1992), en affirmant l'unidirectionnalité de la grammaticalisation. De même, les études récentes portant sur le développement de la prédication complexe, dans le cadre de la linguistique fonctionnaliste américaine, tendent à dégager un mouvement cyclique du lexical au grammatical puis au lexical, qui aboutirait à la création de verbes lexicaux complexes, selon un mouvement général parataxe-syntaxe-lexique (Givon, 2009)<sup>12</sup>. Si l'on conjoint cette position récente de Givon à la formulation d'un état antérieur de sa recherche, souvent reprise : « Today's morphology is yesterday's syntax » (Givon, 1971, p. 413), on constate le lien indissoluble de la grammaire et du lexique, et par là même l'utilité d'une approche lexicographique des items grammaticaux, qui invite à modéliser le continuum polycatégoriel.

On note au demeurant que ce double mouvement de grammaticalisation et de lexicalisation, qui est à la base de nombreux travaux en Europe et dans le monde, sur le modèle du fonctionnalisme américain, a été précédemment étudié tant dans le cadre de la grammaire comparée, par les études indo-européennes, que dans le cadre de la psychomécanique de Gustave Guillaume<sup>13</sup>.

Il appartient à la lexicographie de développer, à partir des outils nouveaux offerts par l'informatique, une technique capable de s'adapter au dynamisme de l'objet linguistique.

### 3.3. Le dictionnaire comme outil dynamique

Le dictionnaire, s'il doit survivre aux mutations culturelles induites par les mutations technologiques, pour rester le garant d'une information maîtrisée en savoir, doit certainement changer plus radicalement qu'il ne l'a fait ces dernières décennies<sup>14</sup>.

Les remarques présentées ici ne peuvent constituer que des esquisses de pistes pour opérer cette mutation. On notera toutefois que dès les années 60 du vingtième siècle, un linguiste lexicographe comme Robert Martin souhaitait une réalisation matérielle du dictionnaire qui pût rendre compte, par des procédés typographiques, du rapport de l'évolution diachronique à l'ordonnement de la polysémie en synchronie<sup>15</sup>, et

---

<sup>12</sup> La prédication complexe serait issue de la condensation sous une même intonation de constructions à verbe fini ou non fini, qui antérieurement disposaient de contours intonatifs autonomes, pouvant amener la création d'un lexème : « Both types can then go on to create morphologically complex lexical verbs. Both thus share the general diachronic trend of parataxis-to-syntaxis to lexis, albeit with somewhat different synchronic properties of both the syntactic and lexical product. » (Givon, 2009).

<sup>13</sup> Cf. la grammaticalisation de *rem* en indéfini négatif *rien* et la lexicalisation de celui-ci, telles que décrites par Martin (1966).

<sup>14</sup> On rejoint là les conclusions de Gasiglia (2009).

<sup>15</sup> Cf. Martin (1966, p. 289) : « Ce parallélisme de la synchronie et de la diachronie rassure le lexicographe qui, par des artifices typographiques, le jeu des caractères et des indications sur les dates limites d'emploi peut réussir à combiner le profil historique avec les exigences du mécanisme en synchronie ». L'auteur note toutefois que la similitude entre le cinétisme du mot dans la perspective de la psychomécanique et l'évolution historique n'est pas universelle, le rapport pouvant être inversé : « On imagine aussi bien une progression historique qui recule de proche en proche le point initial du mouvement cinétique au lieu de pousser le dynamisme de plus en plus en avant » (*ibid.*).

L'on ne s'étonnera pas dès lors que le même recommande, quarante ans plus tard, que l'informatisation du dictionnaire, après avoir concerné sa base documentaire puis son support de diffusion consiste désormais en une nouvelle conception, modulable de l'article (Martin, 2007). Ce caractère modulable résultera en particulier de l'exploitation de bases de données dynamiques et de systèmes d'annotation et de conceptualisation plurielle<sup>16</sup>.

La modélisation de la polysémie et de la polycatégorialité des items grammaticaux peut varier (signifié de puissance, sémantique prototypique, air de famille, *inter alia*), mais le dictionnaire est précisément le lieu d'exposer voire de comparer ces modélisations : on peut imaginer que soient présentées plusieurs conceptualisations grâce à une visualisation qui permette de confronter puis de choisir, selon les options théoriques de celui qui consulte l'outil, mais aussi selon les besoins de telle requête, qui peut être différente d'une requête de la même personne dans d'autres circonstances.

Une visualisation plurielle devrait également permettre de mettre en regard divers modes de structuration de l'article, selon un ordonnancement logique (d'un sens propre à un sens figuré), chronologique, par fréquence croissante et par fréquence décroissante (les emplois les moins fréquents pouvant être ceux dont on a le plus besoin dans une recherche d'ordre de compréhension), en optant pour une structuration sémantique ou syntaxique.

Ces nouveaux modes d'affichage doivent permettre de disjoindre la macrostructure de l'ordre alphabétique et offrir une consultation multicritère. Dans le cas de conjonctions, par exemple, on peut imaginer une consultation à partir des constituants : la préposition (le *par* de *parce que*, mais auparavant de *par (seul) (i)tant que*), l'adverbe (le *si* de *si...que*, mais également celui de *aussi... que*, *ainsi que*, *si que*, *issi que* et de leurs concurrents en *com(me)* en français ancien), le phorique *ce* (de *à ce que* et *de ce que* mais aussi celui de *parce que* et auparavant celui de *avant/après, dès/depuis ce que*) ou *tant (tant que et tant que ne*, mais aussi *autant que* et auparavant *autresi tant que, tant ...que, par tant que, por tant que*). Mais on devrait aussi pouvoir interroger par catégorie sémantique (temps, cause, condition, concession, comparaison, *etc.*) ou par structure caractéristique (construction elliptique, emplois des modes et des temps, type de verbe). On peut attendre du croisement de ces critères des avancées sur la connaissance du langage, grâce à une comparaison longitudinale des différents états d'une langue et à une comparaison typologique entre langues. Ce type de dictionnaire aurait aussi bien une fonction cognitive (productive de savoir) qu'interprétative (donnant accès à une compréhension exacte).

---

<sup>16</sup> A titre d'exemple d'un logiciel proposant annotation et conceptualisation multiples, cf. Mazziotta, 2010.

## **Conclusion**

Le présent article laisse ouvertes beaucoup de questions quant à la mise en pratique des propositions énoncées. On ne pourrait apporter de réponses aux questions soulevées ici que par la mise en œuvre concrète d'une entreprise qui nécessiterait la collaboration de linguistes, de philologues et de spécialistes de l'ingénierie linguistique, à partir d'une réflexion commune.

Jacques Anis, qui conjugait les compétences d'un philologue, d'un linguiste et d'un praticien avant-gardiste des technologies de l'information, nous invite à croire, par le dynamisme souriant dont témoignent ses multiples travaux, qu'il ne faut pas renoncer.

## RÉFÉRENCES BIBLIOGRAPHIQUES

- ANIS, J., & CHISS J.-L., PUECH C. (collab.), 1988, *L'écriture. Théories et description*, Bruxelles, De Boeck-Université, collection Prisme.
- BAT-ZEEV SHYLDKROT, H., 1995, éd., *Langue Française*, 107.
- BÉJOINT, H., 2007, « Informatique et lexicographie de corpus : les nouveaux dictionnaires », *Revue Française de Linguistique Appliquée*, XII/1, « Corpus : état des lieux et perspectives », p. 7-23.
- BRIL, I., & G. REBUSCHI, (éds), 2006. *Coordination et subordination : typologie et modélisation. Faits de Langue*, 28, Paris, Ophrys.
- CORBLIN P. & N. GASIGLIA, 2009, « Changer les dictionnaires ? Une pluralité d'approches », Introduction à *Lexique* 19, p. 7-38.
- CORI, M., 2008, « Des méthodes de traitement automatique des langues aux linguistiques fondées sur les corpus », *Langages* 171, p. 95-110.
- DESMETS, M., GAUTIER, A. & VERJANS, T., éd., 2008, *Linx*, 58.
- ECO, U., 1984, *Semiotica e filosofia del linguaggio*, Torino, Einaudi (1988 trad. fr.).
- GASIGLIA, N., 2009, « Evolutions informatiques en lexicographie: ce qui a changé et ce qui pourrait émerger », *Lexique* 19, p. 235-298.
- GIVON T., 1971, « Historical Syntax and Synchronic Morphology : an Archeologist's Field Trip », *Chicago Linguistic Society* 7, p. 394-415.
- GIVÓN, T., 2009, « Multiple routes to clause union: The diachrony of complex verb phrases » in Givón, T. and Masayoshi Shibatani (eds.), 2009, *Syntactic Complexity: Diachrony, acquisition, neuro-cognition, evolution*. vi, pp. 81-118.
- GUILLAUME, G., 1964, *Langage et science du langage*, Paris, Nizet- Québec, Presses de l'Université Laval.
- HARRIS, M., 1986, « Aspects of subordination in English and other languages », *Bulletin of the John Rylands University Library of Manchester*, 69, 1, p. 123-126.
- JOLIVET, J., 1982, *Arts du langage et théologie chez Abélard*, 2<sup>de</sup> éd. augmentée, Vrin, Paris, 1982.
- KEMMER, S. & ISRAËL, M., 1994, « Variation and the usage-based model », in Beals, K., et al. (eds), *CLS 30: Papers from the 30<sup>th</sup> Regional Meeting of the Chicago Linguistic Society*, Vol. 2, p. 165-179, Chicago, CLS.
- KEMMER, S. & BARLOW, M., 2000, (eds), *A Usage-Based conception of Language*, Stanford, Center for the Study of Language and Information Publications.
- KILGARIFF, A., RYCHLY, P., SMRZ, P. & TUGWELL, D., 2004, « The Sketch Engine », in Williams, G. & Vessier, S. (éds.), *Proceedings of the Eleventh EURALEX International Congress*, Lorient, Université de Bretagne-Sud, p. 105-115.
- LALLOT, J., 1989, *La grammaire de Denys le Thrace*, Traduction annotée, Éditions du CNRS.

- LANGACKER, R., 1988, « A usage-based model », in Rudzka-Ostyn, B., (ed.), *Topics in Cognitive Linguistics*, Current Issues in Linguistic Theory 50, Amsterdam, Benjamins, p. 127-161.
- LANGACKER, R., 1990, « The rule controversy: A Cognitive Grammar perspective », *Center for Research in Language Newsletter* 4(3), p. 4-15.
- LANGACKER, R., 2000, « A Dynamic Usage-Based Model », in Kemmer, S. & Barlow, M., *A Usage-Based conception of Language*, Stanford, Center for the Study of Language and Information Publications, p. 1-64.
- LEHMANN, C., 1985, « Grammaticalization : Synchronic Variation and Diachronic Change », *Lingua e Stile*, XX/3, p. 303-318.
- LEHMANN, C., 1982/1985/1995, *Thoughts on grammaticalization*, Munich, Lincom-Europa.
- MARCHELLO-NIZIA, Ch., 2001, « Grammaticalisation et évolution des systèmes grammaticaux », *Langue française*, 135, p. 33-41.
- MARCHELLO-NIZIA Ch., 2006, *Grammaticalisation et changement linguistique*, Bruxelles, De Boeck.
- MARTIN, R., 1966, *Le mot « rien » et ses concurrents en français (du XIV<sup>e</sup> siècle à l'époque contemporaine)*, Paris, Klincksieck, 1966.
- MARTIN, R., 1969, « Le trésor de la langue française et la méthode lexicographique », *Langue française*, 2.
- MARTIN, R., 2007, « Sur la nature du 'signifié de langue'. Réflexions d'un lexicographe », *Bulletin de la Société de Linguistique de Paris*, 102-1, p. 17-33.
- MAZZIOTTA, N., 2010, « Logiciel NotaBene pour l'annotation linguistique. Annotations et conceptualisations multiples », *Recherches qualitatives*, n°9, Hors série *Logiciels pour l'analyse qualitative : innovations techniques et sociales*, 83-94, ISSN 1715-8702 - <http://www.recherche-qualitative.qc.ca/Revue.html>.
- MEILLET, A. 1915, [1948], « Le renouvellement des conjonctions », *Annuaire de l'Ecole pratique des Hautes Etudes, 1915-1916*, repris in *Linguistique historique et linguistique générale*, Paris, Champion, 1948, p. 159-174.
- MELČUK, I.A., CLAS, A., POLGUERE, A., 1995, *Introduction à la lexicologie explicative et combinatoire*, Duculot.
- MELIS, L. & DESMET, P., 1998, « La grammaticalisation: réflexions sur la spécificité de la notion ». *Travaux de linguistique* 36, p. 13-26.
- MOLINE, E. & FLAUX, N., 2008, (éd.), *Langue Française*, 159.
- MULLER, Cl., 1996. *La subordination en français, le schème corrélatif*, Paris, Colin.
- NORDE, M., 2009, *Degrammaticalization*, Oxford, Oxford University Press.
- POSNER, R., 1997, *Linguistic Change in French*, Oxford, Clarendon Press.
- POTTIER, B., 1964, « Vers une sémantique moderne », *Travaux de linguistique de Strasbourg*, II,1, p. 107-137.
- PRÉVOST, S., & FAGARD, B., 2007, « Présentation : grammaticalisation et lexicalisation : la formation d'expressions complexes », *Langue française*, 156, p. 3-9.
- RAMAT, P., 1992, « Thoughts on degrammaticalization », *Linguistics*, 30, p. 549-560.

*Annie Bertin*

- SCHNEDECKER, C. & D. CAPIN, 2010, « *Tout le monde* : étude d'une pronominalisation très particulière », *Colloque dia sur la variation et le changement en langues*, Université de Gand, 13-15 septembre 2010.
- TARP, S., 2009, « Beyond Lexicography: New Visions and Challenges in the Information Age », in *Lexicography at a Crossroads. Dictionaries and encyclopedias Today, Lexicographical Tools Tomorrow*, Bergenholz, H., Nielsen, S. & Tarp S. (eds), Bern, Peter Lang.
- TRAUGOTT, E.C., & HEINE, B., (éds), 1991, *Approaches to grammaticalization*, Amsterdam-Philadelphie, John Benjamins.
- VERONIS, J., 2004, « Quels dictionnaires pour l'étiquetage sémantique ? », *Le français moderne* 72-1 : p. 27-38.
- WILSON, N.L., 1967, « Linguistical Butter and Philosophical Parsnips », *Journal of Philosophy*, p. 64.

Base de Moyen Français: [www.atilf.fr](http://www.atilf.fr)