



**HAL**  
open science

## Quelle place pour la syntaxe ?

Sophie Prévost

► **To cite this version:**

Sophie Prévost. Quelle place pour la syntaxe ?. Frédéric Duval; Céline Guillot-Barbance; Fabio Zinelli. Les introductions linguistiques aux éditions de textes, Classiques Garnier, pp.255-273, 2019, Histoire et évolution du français, 978-2-406-08578-2. 10.15122/isbn.978-2-406-08580-5.p.0255 . halshs-03986730

**HAL Id: halshs-03986730**

**<https://shs.hal.science/halshs-03986730>**

Submitted on 13 Feb 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Quelle place pour la syntaxe ?

Sophie Prévost

Restée pendant longtemps le parent pauvre des introductions aux éditions de textes<sup>1</sup>, la linguistique a vu son importance croître au fil des années, et elle occupe désormais une place souvent non négligeable, à côté de la présentation des principes qui ont présidé à l'édition et des aspects littéraires. Mais, il faut bien le constater, la syntaxe se voit souvent réserver la part congrue de la présentation linguistique, les développements linguistiques concernant en premier lieu, encore bien souvent, la phonétique et la morphologie (ou la morpho-syntaxe). Les raisons de cet état de fait sont diverses. On peut ainsi supposer que cette place moindre (même si croissante) de la syntaxe traduit, malgré le décalage dans le temps, l'intérêt plus tardif qui a été porté à la syntaxe dans le champ de la linguistique en général (même si la syntaxe occupe au contraire depuis plusieurs décennies le devant de la scène), intérêt qui semble encore modéré dans le champ des études philologiques, en comparaison avec la phonétique et la morphologie. Cela tient peut-être aussi au fait qu'il est moins évident de pointer, en syntaxe, les faits qui méritent mention, au-delà de la mise en évidence de quelques constructions « marquées », l'univers de référence étant plus difficile à délimiter. Il faut ajouter à cela que l'une des préoccupations majeures d'un éditeur est la localisation d'un texte, de son manuscrit de base, et qu'il n'est pas simple d'attribuer une valeur diatopique à un phénomène de syntaxe<sup>2</sup>.

S'interroger sur la place de la syntaxe dans les introductions linguistiques conduit à une double question. Il s'agit d'une part d'essayer d'évaluer la place qui est faite à la syntaxe, non seulement dans l'introduction linguistique, mais aussi dans l'édition même du texte, et dans l'apparat critique dans son ensemble. Il s'agit par ailleurs de s'interroger sur la place qui pourrait, sinon devrait, y être faite, au regard des attentes que l'on peut supposer chez les linguistes syntacticiens. Nous proposerons dans un premier temps quelques éléments de réflexion sur les bouleversements qu'a apportés la linguistique de corpus, en particulier dans le rapport qu'entretient le linguiste avec les éditions de textes. Nous nous pencherons ensuite sur les principes d'édition et les variantes, pour aborder au final, la question de la place de la syntaxe dans les introductions linguistiques.

### 1. « Linguistique de corpus » : une nouvelle donne

Ce que l'on a coutume désormais d'appeler la « linguistique de corpus » a bouleversé le paysage des recherches linguistiques depuis bientôt deux décennies, et a notablement modifié le rapport que bon nombre de linguistes entretiennent avec les éditions de textes. Dans la perspective qui nous intéresse ici, on peut donc se poser la double question suivante : Que cherchent les linguistes syntacticiens dans les introductions aux éditions de textes ? Que cherchent, dans ces mêmes introductions, les linguistes de corpus, et question préalable, vont-ils y chercher quelque chose ? Ces deux questions mettent en évidence des problématiques qui sont un peu différentes. Comme le faisaient justement remarquer les organisateurs de ce

---

<sup>1</sup>Nous nous intéresserons ici aux textes médiévaux.

<sup>2</sup> Je remercie l'un des évaluateurs de cet article d'avoir attiré mon attention sur ce dernier point.

colloque dédié aux introductions linguistiques des éditions de textes :

Par ailleurs, la multiplication des corpus textuels modifie les comportements de recherche des linguistes, qui préfèrent obtenir des résultats chiffrés à des requêtes qu'ils auront préalablement définies, plutôt que de reprendre les données hétérogènes des introductions.

De toute évidence, le développement de la « linguistique de corpus » a entraîné un déplacement très net du rapport au texte. On pourrait, de manière quelque peu simplifiée, définir ici ce type d'approche, voire cette nouvelle sous-discipline, comme le fait de travailler sur des données langagières nombreuses et, les deux étant souvent liés, sur un nombre élevé de textes. Numérisées, ces données sont exploitables par des outils de traitement automatique, à l'aide de requêtes de différents ordres. On a donc désormais la possibilité de travailler sur des données langagières sans parcourir linéairement les textes dont elles proviennent. À une lecture horizontale, syntagmatique, tend ainsi à se substituer une lecture verticale, paradigmatique, le mode de visualisation des données collectées sous forme de concordances illustrant pleinement cette lecture par classes d'occurrences, quelles qu'elles soient.

Certains de ces textes sont connus du linguiste, parce qu'il les a par ailleurs lus, voire étudiés, mais il n'y a là aucune nécessité, ni aucune règle. Une telle approche, pour ainsi dire « en aveugle » peut être perçue comme une facilité, rendue possible par les outils de traitement qui permettent de collecter automatiquement les données souhaitées (sous réserve que le texte présente l'enrichissement adéquat), mais elle aussi rendue obligatoire par la masse des données traitées : il n'est en effet souvent pas possible, pour de simples raisons de temps, de parcourir l'ensemble du corpus que l'on a choisi d'exploiter pour une étude spécifique. Par exemple la seule partie ancien français (IX<sup>e</sup>-XIII<sup>e</sup> siècles) de la Base de Français Médiéval<sup>3</sup> comprend près de 2.4 millions de mots (ponctuation incluse), soit près de 20 fois la taille de la *Queste del Saint Graal*, corpus qu'il est difficile à parcourir dans son intégralité... mais qu'il est cependant tentant d'exploiter dans son intégralité. Un compromis acceptable réside dans une lecture partielle, un « survol » des textes, démarche que la communauté des médiévistes, et plus généralement de ceux qui travaillent sur des données langagières non contemporaines, est sans doute plus encline à mener, en raison d'une certaine habitude à fréquenter les textes.

Le mouvement est de toute évidence croissant, même si des études sont toujours menées sur des corpus plus restreints qui autorisent le parcours intégral des textes, et que, par ailleurs, certaines études, en particulier celles qui relèvent du domaine sémantique, ne permettent pas (encore ?) la collecte automatique des faits langagiers étudiés, faute d'encodage adéquat de l'information.

La massification des données exploitées, que permet leur numérisation et leur traitement automatique présente d'indéniables avantages. Elle permet en particulier de mieux rendre compte d'un état de langue, en dépassant la variation idiolectale, ainsi que des changements, puisque l'accroissement des données considérées autorise un repérage plus précis des « lieux » du changement. Les inconvénients n'en existent pas moins, et nous les avons esquissés ci-dessus. Il s'agit en premier lieu d'une tendance croissante à travailler « en aveugle » sur les textes. Souvent, on collecte des structures langagières coupées de leur contexte linguistique et textuel, parfois réduites à des étiquettes linguistiques, et, enfin, déconnectées de leur source textuelle (on ignore même parfois le texte dont elles sont issues)

---

<sup>3</sup> BFM - Base de Français Médiéval [En ligne]. Lyon : ENS de Lyon, Laboratoire ICAR, 2012, <<http://bfm.ens-lyon.fr>>.

et de ses caractéristiques. On sait que ce dernier point est important – certains faits langagiers sont sur- ou au contraire sous-représentés dans certaines formes de textes, dans certains genres. Il existe de ce point de vue une certaine tension, entre le désir de dégager des tendances et celui de ne pas écraser la spécificité des textes.

Certains outils d'exploration des textes permettent néanmoins de « garder le contact » avec la source textuelle. C'est le cas du logiciel TXM<sup>4</sup>, qui permet, à l'aide de requêtes portant sur des séquences de caractères ou sur des étiquettes morpho-syntaxiques, d'interroger les textes de la BFM. Les résultats de la requête sont produits sous forme de concordances, avec des contextes gauche et droit paramétrables, et point essentiel, un simple clic sur l'une des occurrences permet d'accéder au texte et de visualiser ainsi ladite occurrence dans son contexte textuel, avec la mise en forme de l'édition d'origine.

Tous les moteurs d'exploration de corpus n'autorisent pas un tel accès au contexte textuel large, ou ne le permettent pas aussi facilement, et la démarche reste de toute façon optionnelle, au gré du linguiste.

Le fait est là : les nouvelles modalités d'exploration des textes permettent de travailler sur une langue qu'on ne connaît pas ou qu'on connaît mal (une telle situation n'est pas rare dans le cadre des études typologiques, mais l'on s'appuie souvent dans ce cas sur des informateurs locuteurs natifs des langues que l'on étudie). De manière assez symptomatique, on voit désormais émerger des études sur gros corpus qui émanent de linguistes très peu familiers (voire pas du tout) avec le français médiéval. On ne peut que se réjouir de ce regain d'intérêt porté à la langue ancienne, mais l'on peut, en revanche, déplorer la présence, parfois, de conclusions qui dénotent une méconnaissance de la langue, et ne parviennent pas à resituer les faits étudiés dans le contexte plus large de celle-ci, ce qui peut conduire à certaines incongruités.

La linguistique de corpus a produit de nouveaux usages, et de nouveaux usagers. Pour autant, l'appréhension des textes via le support numérique et les outils d'exploration automatique ne rend pas caduque l'apparat critique des éditions, au contraire.

D'une part, la plupart des versions numérisées des textes donne à voir un texte privé des modifications qu'il a subies, et de ses variantes. Toute édition est le résultat de choix éditoriaux, auxquels nous n'avons pas directement accès lorsque nous consultons/explorons la version numérisée d'un texte : l'apparat critique est en effet disjoint du support numérique du texte, sauf lorsqu'il s'agit d'une « édition numérique ».

D'autre part, et c'est en partie lié, bon nombre de linguistes de corpus recourent conjointement à l'apparat critique des éditions papier, qui leur permet précisément d'accéder aux principes d'édition, aux variantes et à l'introduction linguistique du texte. Cette dernière est susceptible de fournir des informations essentielles, d'autant plus importantes que l'outil informatique, s'il est infaillible, n'est pas intelligent et n'a pas d'initiative : il ne collectera que les constructions qui correspondent strictement à la requête formulée. Toute construction quelque peu déviante, à laquelle le linguiste n'aura pas nécessairement pensé, parce qu'elle est rare, inhabituelle, inattendue dans ce texte, lui échappera. L'introduction linguistique d'un texte recense, justement, ce genre de constructions.

C'est d'un double point de vue que je poursuivrai mon propos : celui d'une linguiste usagère des bases textuelles, conduite à « brasser » de grandes quantités de données, et celui d'une linguiste qui s'attache aussi, plus spécifiquement, à un ou plusieurs texte(s), dans le

---

<sup>4</sup> <http://textometrie.ens-lyon.fr/>

détail.

C'est plus précisément au domaine de la syntaxe que je m'attacherai, en envisageant sa place tant dans les principes d'édition que dans les variantes et les introductions linguistiques. Il ne s'agit certainement pas de distribuer de bons et de mauvais points à telle ou telle édition : ce serait fort malvenu de la part d'une linguiste qui ne s'est elle-même jamais risquée au difficile exercice de l'édition de textes. Il s'agit simplement d'énoncer quelques remarques, et de formuler quelques suggestions.

## 2. Les principes d'édition

En 2006, en introduction aux actes du colloque sur les pratiques d'édition en Europe, Frédéric Duval écrivait:

Le défi qui s'offre aux philologues consiste en fait à satisfaire deux demandes opposées, celle d'un public qu'il faut persuader que les textes du Moyen Age valent encore la peine d'être lus en version originale, et celle des linguistes, qui désirent un texte présentant le plus d'affinités possible avec les manuscrits. Sur ce point rien n'est joué.

Le défi est en effet de taille, tant les attentes des uns et des autres peuvent être divergentes. Ce sont avant tout celles du linguiste que nous évoquerons ici.

De toute évidence, le linguiste souhaite une édition non-interventionniste sur le plan syntaxique : il lui importe davantage d'avoir la « version usagée » d'un copiste que la version reconstruite et supposée proche de celle de l'auteur. C'est d'autant plus important pour les usagers des bases textuelles, qui n'ont accès qu'au texte, et non à l'apparat critique (sauf s'ils font la démarche d'aller consulter l'édition papier) et ne peuvent donc évaluer le degré d'intervention de l'éditeur.

La question de l'interventionnisme se pose d'une part pour les « modernisations ». Il semble cependant que la situation soit moins critique pour les textes médiévaux que pour ceux du XVII<sup>e</sup> siècle, l'interventionnisme étant bien moindre dans les premiers (voir les remarques à ce propos dans Duval (2015)). Dans deux domaines les interventions restent cependant assez fréquentes : la ponctuation (voir l'article d'A. Lavrentiev dans ce volume) et la segmentation (auxquels on pourrait d'ailleurs ajouter les abréviations et les signes diacritiques).

Il est cependant vrai que le domaine de la syntaxe est moins sujet aux « corrections », hormis pour ce qui relève de la syntaxe d'accord, c'est-à-dire de la morpho-syntaxe, tels que les accords « fautifs » de déclinaison. Pour ce qui touche à la syntaxe de phrase, les interventions sont bien moindres, parfois inexistantes. Cela s'explique, nous semble-t-il, par différentes raisons. D'une part, la variation syntaxique est davantage perçue comme intrinsèque au système langagier, beaucoup moins en termes de « fautes » ; d'autre part, et c'est en partie lié, les règles et les normes sont moins clairement établies dans ce champ ; enfin, et cet aspect est à vrai dire premier, la syntaxe, comme le lexique d'ailleurs, semblent être des lieux de moindre intervention du copiste.

On ne saurait discuter la question de l'interventionnisme éditorial sans discuter la notion de faute. Il existe de toute évidence une tension entre la « faute », orientée vers un original idéal, et l'intelligibilité, orientée vers le lecteur, contemporain ou moderne (cf. F. Duval, 2006).

Cette notion de faute se révèle problématique à différents égards. En premier lieu, elle relève d'une approche anachronique de la langue (voir Duval (2006 et 2015) pour l'historicité de

cette notion): la prétendue incohérence du système correspond en effet à une vision moderne, ce dont témoigne le caractère relativement récent du concept de norme. De plus, la notion de faute contribue à construire une vision lissée et erronée d'une langue intrinsèquement complexe et soumise à la variation (il existe de ce point de vue un effet circulaire par rapport aux grammaires : se fondant en partie sur des versions « corrigées » des textes, celles-ci donnent parfois à voir une image « régulière », et fautive, de la langue, à l'aune de laquelle des textes « non corrigés » seront jugés « fautifs »<sup>5</sup>). Par ailleurs, cet interventionnisme met en cause les études diachroniques : comment repérer les changements si toute déviance par rapport à une norme supposée est gommée ? On le sait, tout fait minoritaire, voire tout hapax, est digne d'intérêt, parce qu'il peut signaler les prémices d'un changement. Enfin, et ce point est lié au précédent, toute correction invalide en partie les recherches quantitatives.

Le linguiste diachronicien ne peut que défendre une approche la moins interventionniste possible, et souhaiter que, *a minima*, les corrections soient explicitées.

Nous formulons ci-dessous quelques remarques sur les lieux majeurs de modernisation syntaxique : la ponctuation et la segmentation.

## 2.1. La ponctuation

Nous ne développerons pas ce point, par ailleurs traité ici-même de manière plus détaillée – et bien plus compétente – par A. Lavrentiev.

Il ne s'agit pas, dans le cadre d'une édition « critique », de conserver le système du copiste, peu compréhensible pour la grande majorité des lecteurs modernes, mais d'essayer de le restituer au mieux. Dans une perspective syntaxique, un point nous semble particulièrement important. Il s'agit de la délimitation des unités, qui, si elle respecte le/les manuscrit(s), permet de restituer la spécificité du français médiéval. Comme l'a bien montré B. Combettes (2002 et à paraître), le système de hiérarchisation et de rattachement des unités était à cette époque fort différent de celui du français moderne : il n'est pas rare de rencontrer des unités que le lecteur/locuteur moderne considère comme « en l'air », ou qui ne sont pas clairement attachées à l'amont ou à l'aval textuel. Il ne s'agit d'ailleurs peut-être pas tant d'une incapacité du lecteur moderne à établir le bon rattachement que d'une indétermination intrinsèque de celui-ci. C'est un fait connu : la « phrase » telle que nous la concevons aujourd'hui est un concept relativement moderne (XVIII<sup>e</sup> siècle, cf. Seguin 1993) et donc largement anachronique pour la période médiévale. Il n'en demeure pas moins tentant de chercher à insérer le texte médiéval dans un tel moule, tant il structure le rapport que le locuteur moderne entretient au texte et à la langue. S'en extraire, et parvenir à restituer la segmentation réelle du texte, ou au moins s'en approcher, nous permet cependant d'appréhender une dimension essentielle du texte médiéval, susceptible d'ailleurs de conduire à une partielle révision d'études syntaxiques antérieures, en particulier pour ce qui touche à la « phrase » complexe.

---

<sup>5</sup> On ne saurait cependant négliger le fait que la notion de faute est un outil de travail complexe, qui sert à s'orienter dans les relations existant entre les manuscrits. De ce point de vue, elle constitue davantage un moyen qu'une fin. Je remercie l'un des évaluateurs d'avoir attiré mon attention sur ce point.

## 2.2. La segmentation en mots/unités lexicales

La segmentation en mots/unités lexicales constitue assurément un lieu d'intervention majeure. L'interventionnisme, en ce domaine, semble le plus souvent résulter du dilemme suivant : il s'agit d'un côté de faciliter la lecture, ainsi que les requêtes automatisées, il s'agit de l'autre de respecter les segmentations du manuscrit. Dans l'ouvrage *Conseils pour l'édition des textes médiévaux* (2014) F. Vieilliard et O. Guyotjeannin déclarent ainsi « le principe général est de ne transcrire en un seul mot que les expressions composées usuelles ou celles qui sont passées soudées dans la langue moderne ».

Ils mettent cependant à part les cas où l'analyse grammaticale rend perceptible l'individualité ou la solidarité de chaque composant : ainsi de *des ore en avant* versus *desorenavant* (on peut dès lors s'interroger sur le traitement à réserver à *ma dame...*).

Il nous semble que l'on peut distinguer deux cas de figure, et nous nous appuyons largement pour cela sur l'introduction de l'édition numérique de la *Queste del Saint Graal* (Marchello-Nizia), et des exemples qui y sont fournis.

Tout d'abord, lorsqu'une seule unité graphique est utilisée pour plusieurs mots, comme cela n'est pas rare quand ceux-ci forment une seule unité syntaxique (syntagme nominal, groupe verbal), C. Marchello-Nizia suggère de dégroupier ces mots en suivant les normes de la graphie moderne, afin de faciliter la lecture, comme dans les exemples suivants : *ladamoisele* > *la damoisele* ; *illiot* > *il li ot* ; *silidist* > *si li dist*. Pour le même raison, l'emploi de l'apostrophe est souhaitable en cas d'élision : *senaloit* > *s'en aloit*.

C. Marchello-Nizia propose un interventionnisme raisonné et faisant preuve de discernement. Ainsi *tres*, lorsqu'il est suivi d'un adverbe ou d'un adjectif, peut être interprété comme un adverbe en devenir, et il doit de ce fait être séparé : *tresgrant* > *tres grant*. En revanche, lorsqu'il est suivi d'un verbe, il correspond à un préfixe, et doit demeurer attaché : *trespasser*. Le traitement des unités en cours de figement est quant à lui différent, et C. Marchello-Nizia suggère de respecter l'usage du copiste, y compris dans sa variation, qui rend compte du changement en cours (ainsi que des incertitudes que le copiste pouvait avoir vis-à-vis de l'usage de son modèle). Seul le respect des graphies peut en effet permettre de quantifier et de dater précisément le phénomène.

Cette variation complexifie cependant le traitement automatique des textes, et risque de faire omettre des séquences. Le caractère numérique de l'édition de la *Queste del Saint Graal*, réalisée par C. Marchello-Nizia et A. Lavrentiev, permet de pallier cette difficulté. En effet, en surface, la graphie du manuscrit est préservée, tandis qu'à un niveau abstrait est produite une graphie unifiée : ainsi une requête portant sur 'parmi' permettra de collecter les occurrences de *parmi* comme celles de *par mi*. L'introduction fournit d'ailleurs la liste des occurrences concernées, ainsi que la quantification de leurs différentes graphies.

## 3. Les variantes

Les variantes sont assurément les grandes perdantes des bases textuelles, desquelles elles sont le plus souvent absentes.

D'une manière générale, se pose pour l'éditeur de texte une double question. En premier lieu, il faut décider des variantes à retenir, sachant que littéraires et linguistes ne s'intéressent pas nécessairement aux mêmes. Par ailleurs, faut-il retenir toutes les variantes, dans tous les manuscrits ?

Les choix sont de ce double point de vue fort variés. En effet, alors que certaines éditions font un choix très sélectif, d'autres optent au contraire pour une présentation très complète des variantes. Ainsi dans *La chevalerie Ogier : Enfances* (éd. M. Ott, 2012), les variantes occupent 100 pages.

Les variantes présentent un intérêt considérable pour le linguiste. Elles permettent non seulement d'appréhender la variation synchronique intrinsèque à la langue, mais aussi, la variation diachronique : si les manuscrits sont de dates différentes, la présentation complète des variantes permet de repérer des changements.

Il s'agit donc d'informations essentielles, mais malheureusement souvent difficiles à exploiter pour le linguiste. En effet, une variante fait certes sens par rapport à une autre, mais elle fait sens aussi (et tout autant) par rapport au système linguistique dans lequel elle s'inscrit. Il nous semble qu'il serait donc souhaitable de grouper les variantes selon les manuscrits, en adoptant cependant une vision dynamique qui permette de ne pas perdre le lien entre les variantes de différents manuscrits, et de voir comment, par exemple, la naissance de la variante d'un manuscrit A a été provoquée par la variante d'un manuscrit B. Le support numérique devrait permettre une telle présentation bidimensionnelle des variantes.

Par ailleurs, il serait fort utile que les variantes soient classées d'un point de vue linguistique : phonétique / graphique, morphologique et syntaxique.

Cette double classification en ferait un outil assurément fort précieux, et davantage exploitable qu'il ne l'est souvent actuellement.

#### **4. Les introductions linguistiques**

Nous l'avons rappelé en introduction : la linguistique est souvent le parent pauvre des introductions, et l'on peut par conséquent se poser la question suivante : à quoi servent les introductions linguistiques ?

F. Duval (2006) retient quatre fonctions pour les introductions : contribuer à l'identification de l'œuvre, favoriser l'établissement du texte, favoriser sa compréhension, et nourrir la connaissance générale de la langue en signalant les faits linguistiques jugés intéressants.

Les deux derniers points sont liés : fait difficulté (gêne la compréhension) ce qui est inattendu, rare ; et c'est sans doute aussi le plus intéressant. Il s'agit donc de mettre en lumière ce qui déroge par rapport à un attendu.

Si la linguistique est le parent pauvre des introductions, la syntaxe est elle-même, souvent, le parent pauvre de ces mêmes introductions : pauvre parmi les pauvres, sa place ne semble guère enviable (place qu'elle partage cependant souvent avec le lexique). Cet état de fait nous semble avoir plusieurs explications.

On peut en premier lieu établir un parallèle avec la place de la syntaxe dans les grammaires, et dans l'histoire de la linguistique, où la syntaxe prend place derrière la phonétique et la morphologie. La situation s'explique par ailleurs en partie par la formation des éditeurs : beaucoup sont littéraires, avec une formation non spécifiquement linguistique (et les cours d'ancien français que certains ont pu suivre sont souvent davantage centrés sur la phonétique et sur la morpho-syntaxe). Enfin, et c'est peut-être là le point essentiel, la syntaxe est un lieu de variation plus important et/ou un lieu où les régularités – et les règles qui en sont tirées – sont plus difficiles à déterminer.

Pour la déclinaison, il existe un modèle bien défini (même si imparfaitement respecté), modèle par rapport auquel on repère et on évalue les « déviations ». C'est bien moins le cas pour la syntaxe.

Si la linguistique occupe encore parfois une place réduite dans les introductions, on ne peut cependant que constater que celle-ci tend à croître. Il est vrai que l'on part de bien peu. En effet, lorsque l'on consulte les éditions anciennes, force est de constater une présence le plus souvent limitée à quelques remarques phonético-graphiques, lesquelles sont d'ailleurs généralement subordonnées à la datation et à la localisation du/des manuscrit(s). Pour le reste, il faut se contenter de quelques remarques en notes (« notes et éclaircissements » destinées à éclairer les difficultés)<sup>6</sup>.

La place de la linguistique a donc crû, et une place, certes encore souvent réduite, est désormais réservée à la syntaxe. L'objectif est, explicitement, de pointer les faits remarquables et /ou intéressants. Déterminer par rapport à quoi ils le sont n'est cependant pas une affaire simple.

Il nous semble que l'on peut établir différents types de faits « remarquables ».

#### **4.1. Les faits remarquables**

##### ***Les faits langagiers qui marquent un écart par rapport à une norme***

Typiquement, il s'agit du non-respect de la déclinaison, qui relève de la syntaxe d'accord, les faits étant variablement reportés sous la rubrique « morphologie » ou « morpho-syntaxe ». Cette rubrique est largement partagée par l'ensemble des éditions, et bien souvent la liste des occurrences est fournie, avec ses références.

Il arrive parfois – bonus pour le linguiste – que soient indiqués les contextes linguistiques propices aux formes « déviantes ». C. Marchello-Nizia signale ainsi, dans l'introduction de la *Queste del Saint Graal*, que l'on rencontre la forme *Sire* en apostrophe mais la forme *sires* (marquage casuel redondant) au cas sujet et pour l'attribut. De la même manière, M. Ott, dans l'introduction de l'édition de *Guibert d'Andrenas* (2004) recense les contextes semblant favoriser le non-respect du cas sujet. Autre exemple, dans l'introduction de l'édition de *Le Dit de la Fleur de Lis* (2014), texte du XIV<sup>e</sup> siècle, F. Duval identifie les îlots de résistance du cas sujet, à une époque où le système casuel est déjà fortement dégradé.

##### ***Les traits dialectaux***

Il est rare de rencontrer des remarques dialectales de nature syntaxique. Il est indéniable que la variation diatopique est bien moindre dans ce domaine en comparaison de celui de la morphologie. Par exemple, dans *l'Estoire de la guerre sainte* (2014), C. Croizy-Naquet formule quelques remarques syntaxiques, mais qui ne sont pas spécifiquement liées à l'anglo-normand : elle souligne ainsi que la syntaxe est, dans ce texte, beaucoup moins spécifique, dialectalement, que la phonétique ou la morphologie.

##### ***Les écarts par rapport à l'état de langue attendu***

Les écarts par rapport à l'état de langue attendu constituent une source potentielle de remarques, ce qui suppose cependant que la langue de l'époque du texte est perçue comme

---

<sup>6</sup> Les articles de Bédier (1928) et de Roques (1926) sont d'ailleurs silencieux sur cette question.

étant dans un état stable.

En priorité ce sont les archaïsmes, ou au contraire les traits précurseurs (possibles prémices à un changement), qui sont signalés.

La comparaison entre deux manuscrits est à cet égard fort enrichissante. F. Duval note ainsi, à propos du *Dit de la Fleur de Lis*, que le manuscrit A, plus tardif, utilise davantage le pronom personnel sujet que le manuscrit B, signe que le copiste a modernisé la langue de Guillaume de Digulleville. La démarche de M. Ott, dans l'introduction de *Guibert d'Andrenas*, est analogue. Soulignant que « cette brève étude est simplement destinée à évaluer le degré de rajeunissement de la langue dans B1 (et B2) par rapport à d'autres manuscrits », elle examine ainsi le sujet (expression, situations prédicatives, avec impératif), la position des pronoms personnels conjoints, les compléments de détermination, les présentatifs, les subordonnants, et l'adverbe *si*.

### ***Le repérage de constructions minoritaires, rares ou d'usage inhabituel***

Il s'agit probablement de la démarche la plus complexe, et pour laquelle il est difficile de dégager des tendances : en effet, que retenir ? L'approche est nécessairement subjective, et sélective.

## **4.2. Remarques sur la structuration des introductions**

L'observation des introductions aux éditions, qu'elles soient anciennes ou plus récentes, laisse paraître une structuration assez immuable, avec une division selon les sections suivantes : phonétique/graphisme; morphologie ou morpho-syntaxe; syntaxe ; lexique.

Les sections concernant la phonétique et la morphologie (ou morpho-syntaxe) sont généralement très structurées. La première traite du vocalisme puis du consonantisme (ou dans l'ordre inverse), de manière souvent assez détaillée. Pour la seconde, se succèdent généralement des remarques sur la morphologie nominale (avec différentes catégories : noms, adjectifs... et bon nombre de remarques sur les faits d'accord) puis sur la morphologie verbale (remarques sur les radicaux et les désinences, déclinées selon les différents temps utilisés dans le texte).

L'organisation du propos suit, dans les deux cas, une grille d'analyse prédéfinie, dont il suffit, d'une certaine manière, de renseigner les différentes sections, lorsque cela s'avère pertinent.

Le traitement de la syntaxe est plus complexe. Il convient tout d'abord de signaler que la limite avec la morpho-syntaxe est assez floue. On constate ainsi que de mêmes questions sont traitées dans l'une ou dans l'autre section, selon les éditions (« morpho-syntaxe » et « syntaxe » sont d'ailleurs des sections le plus souvent exclusives l'une de l'autre, mais pas toujours). C'est typiquement le cas pour les phénomènes d'accord ou de désaccord (cas, nombre et genre), et pour l'expression du sujet (dans le *Dit de la Fleur de Lis*, cette question est traitée sous la rubrique « morpho-syntaxe », alors qu'elle apparaît sous le titre « syntaxe » dans d'autres textes).

La difficulté à laquelle sont confrontés les éditeurs est l'absence d'une grille d'analyse prédéfinie, telle qu'il en existe pour la phonétique ou la morphologie. Les remarques sont parfois organisées selon des rubriques. Par exemple, dans les *Triumphes de Pétrarque* (traduction de Bourgouin (Parussa, 2012), on trouve les rubriques suivantes : Ordre des mots, Modes non personnels, Coordination, Reprise pronominale et pléonasmе, *Et* explétif, Ruptures de construction, bien différentes de celles qui se trouvent dans la 6<sup>e</sup> partie de *Perceforest* (Roussineau, 1987) : Adjectifs, Pronoms (sujet et régime), Verbe (accord, modes

et temps), Adverbes, Conjonctions, Syntaxe de phrase (anacoluthes, relatives expressives, « autres »).

Ces deux exemples illustrent la grande variabilité qui se rencontre d'un texte à l'autre, qui va de la syntaxe du groupe à la syntaxe de phrase, y compris de la phrase complexe (faits de parataxe), avec des considérations qui relèvent parfois davantage de la sémantique grammaticale (en particulier pour ce qui touche au système temporel). Dans *La chevalerie Ogier : enfances* (Ott, 2012) plusieurs remarques sont ainsi regroupées sous le label « syntaxe de proposition » (faits de parataxe, subordonnées en *ou*, imminence contrecarrée, système temporel, subordonnées paratactiques), mais l'on trouve aussi des remarques éparses sur la structures V1 (verbe en 1<sup>ère</sup> position), *Que* causal, les démonstratifs épiques, les impératifs avec sujet, la place du pronom sujet, les cas de désaccord entre sujet et verbe, les constructions disloquées...

La présence régulière de liste de remarques non regroupées sous des rubriques nous semble assez symptomatique de la difficulté à structurer rigoureusement cette rubrique « syntaxe ». Ces listes peuvent en outre être parfois assez longues. *La Chevalerie Ogier* comprend 23 remarques, la *Continuation de Perceval* par G. de Montreuil (Le Nau, 2014) en contient 39, qui vont de la suppression du pronom personnel régime direct, au passage du discours direct au discours indirect et aux changements de temps, en passant par la fonction de reprise de *si*, l'emploi du sujet avec impératif...

Il est intéressant de ce point de vue de considérer les éditions successives des différentes parties de *Perceforest*. L'évolution est nette entre la première édition, celle de la 4<sup>e</sup> partie en 1987, et la dernière, celle de la 6<sup>e</sup> partie, en 2015. On passe ainsi d'une présentation structurée de la syntaxe à une liste de remarques (cela vaut aussi d'ailleurs pour la phonétique et la morpho-syntaxe). La 3<sup>e</sup> partie (1988) ne contient que 14 remarques, tandis que la 2<sup>e</sup> partie (1999) en inclut 67, nombre qui augmente encore dans la 1<sup>ère</sup> partie (2007) et atteint 106 remarques, pour redescendre dans les deux dernières éditions : 81 remarques dans la 5<sup>e</sup> partie (2012) et 63 remarques dans la 6<sup>e</sup> partie (2015). D'une manière générale, ces remarques ne font pas l'objet d'une classification explicite, et la cohérence de leur organisation ainsi que la progression de leur présentation ne sont pas toujours évidentes à saisir. Certains regroupements (accords, emploi des pronoms, phrase complexe, auxiliaires...) semblent en outre parfois quelque peu discutables.

Ce mode de présentation énumératif n'est pas problématique en soi, dans la mesure où, à l'inverse, toute classification donne – voire force – à voir la langue d'une manière qui n'est pas nécessairement celle que l'on souhaite. Il témoigne en revanche de l'absence d'une grille d'analyse implicite, et laisse deviner la difficile question qu'a pu se poser l'éditeur ou l'éditrice : que faut-il traiter ?

L'intérêt de ce que l'on trouve dans ces remarques n'est pas en cause, même si l'on peut parfois être surpris. Dans l'introduction à *La Chevalerie Ogier*, M. Ott mentionne ainsi les constructions disloquées : pourquoi ce choix ? Aussi intéressantes que soient ces constructions, on sait qu'elles ne sont pas rares en ancien français (la *Chanson de Roland* en présente de nombreuses occurrences). Pourquoi à l'inverse n'est-il pas question des tournures en « c'est... que » ? Sont-elles absentes du texte ? Il serait sinon intéressant de les mettre en relation avec les constructions disloquées.

Une question demeure souvent, en particulier pour le linguiste : que ne trouve-t-on pas que l'on devrait trouver ? La réponse est assurément complexe (et nous le répétons, il ne s'agit nullement de jeter la pierre aux auteurs des introductions), et elle se pose en outre

différemment, selon que le texte est supposé appartenir à un état de langue « stable » (même si l'on sait qu'aucun état n'est jamais parfaitement stable), ou au contraire en évolution.

Chaque éditeur reste évidemment le mieux placé pour décider de ce qui lui semble intéressant, mais il serait envisageable de recenser les faits que l'on estime mériter mention selon une grille d'analyse, comme cela se fait le plus souvent pour la phonétique et la morphologie. Nous proposons ci-dessous une possible grille d'analyse des faits de syntaxe.

### **4.3. Une proposition de grille d'analyse des faits de syntaxe**

La grille proposée ci-dessous n'a rien de révolutionnaire – il s'agit avant tout d'essayer de recenser l'ensemble des faits susceptibles d'être abordés, en les classant quelque peu – et elle reste évidemment indicative, cela d'autant plus que toutes les « cases » ne méritent pas nécessairement d'être remplies.

La classification qui nous semble la plus simple, la plus neutre, et qui rejoint une certaine tradition grammaticale, consiste à organiser les remarques selon le type d'unité syntaxique : syntaxe interne des groupes, syntaxe de la phrase simple et syntaxe de la phrase complexe.

#### **Syntaxe interne des groupes :**

Les sous-rubriques suivantes peuvent être envisagées :

- groupe nominal : ordre, faits d'accord,...
- groupe verbal : répartition des auxiliaires, accord du participe passé, usage des modaux, auxiliaires d'aspect,...
- constructions du verbe : transitivité, constructions factitives, et infinitives
- clitiques compléments : ordre

#### **Syntaxe de la phrase simple :**

Dans cette rubrique pourraient être abordés les points suivants :

- position du verbe (et pas seulement les structures avec le verbe en 1<sup>ère</sup> position)
- ordre des constituants majeurs (y compris l'expression du sujet)
- spécificités des propositions non déclaratives (interrogatives, impératives...)
- formes de phrase : impersonnelles, clivées, disloquées...

#### **Syntaxe de la phrase complexe :**

- subordination, coordination, juxtaposition, faits de parataxe
- rapport entre la principale et sa/ses subordonnée(s), systèmes corrélatifs
- attraction modale

Ce sont là de simples suggestions : la liste n'est pas exhaustive, et à l'inverse, tous les points mentionnés ci-dessus n'ont pas nécessairement à être traités, cela dépend évidemment des spécificités du texte.

De plus, il demeure deux questions.

Tout d'abord, l'éditeur devrait-il fournir une liste exhaustive des structures remarquables ? La linguiste que je suis a envie de répondre positivement, tout en ayant conscience de la possible ampleur de la tâche. Il serait en tout cas souhaitable que soit explicité le choix qui a été fait, quel qu'il soit. Il n'est en effet pas toujours précisé si les occurrences qui sont mentionnées pour illustrer une construction remarquable le sont dans leur totalité ou non.

Par ailleurs, on peut se demander où s'arrête la fonction de l'éditeur/éditrice : doit-il/elle interpréter les phénomènes repérés ? Cela ne nous paraît pas nécessaire, et ce n'est de toute façon pas sa tâche.

L'éditeur nous semble avoir avant tout la mission d'attirer l'attention du linguiste sur les aspects à explorer. Et cela s'avère d'autant plus important pour les linguistes de corpus.

## 5. Conclusion

Objet de rarissimes remarques dans les éditions les plus anciennes (ce qui s'explique largement par le contexte linguistique général de l'époque), la syntaxe a vu sa place progressivement croître au fil du temps. Celle-ci reste néanmoins bien moindre que celles des domaines plus traditionnels que sont la phonétique et la morphologie. De plus, on ne peut que constater la variabilité des points abordés (ce qui n'est pas le cas pour la phonétique et la morphologie) et la fréquente absence d'une véritable structuration des remarques, les deux étant en partie liés. Cette situation tient probablement à la plus grande difficulté à déterminer, dans le champ syntaxique, ce qui fait écart par rapport à une norme supposée – cette norme étant elle-même malaisée à appréhender – et cela tient conjointement à l'absence d'un modèle, d'une grille d'analyse, auxquels se référer, d'autant plus nécessaire que les faits susceptibles d'être présentés couvrent un spectre plus large que ce n'est le cas pour la phonétique et la morphologie. Au-delà de la seule introduction linguistique, il nous semble que les faits syntaxiques pourraient gagner en visibilité, et leur mention en utilité pour le linguiste, par un traitement plus structuré des variantes, fort précieuses mais malheureusement souvent malaisées à exploiter. Nous avons par ailleurs rappelé combien le linguiste est attaché à un interventionnisme minimal, les « modernisations » opérées tant dans le domaine de la ponctuation que de la segmentation lui masquant des informations essentielles pour appréhender la structuration des unités textuelles et l'évolution de certaines constructions.

On ne saurait évidemment prescrire ce qui doit se trouver dans une introduction, et ce n'était certainement pas l'objectif de ces quelques pages, qui ont simplement essayé, à travers quelques suggestions, d'indiquer comment les introductions linguistiques pourraient devenir un outil plus utilisé par les linguistes syntacticiens.

## Références :

- BEDIER, Joseph, « La tradition manuscrite du *Lai de l'Ombre*. Réflexions sur l'art d'éditer les anciens textes », *Romania* 54, 1928, p. 161-196 et p. 321-356.
- COMBETTES, Bernard, « Aspects de la grammaticalisation de la phrase complexe en ancien et en moyen français », *Verbum* 24, 2002, p. 109-128.
- COMBETTES, Bernard, « De la parataxe à l'hypotaxe : le cas des subordonnées circonstancielles en ancien et en moyen français », Conférence à la Société de Linguistique de Paris, novembre 2014, à paraître dans *Bulletin de la Société de Linguistique de Paris*.
- DUVAL, Frédéric, « Introduction », *Pratiques philologiques en Europe*, éd. F. Duval, Paris, Ecole des Chartes, 2006, p. 5-20.
- DUVAL, Frédéric, « La philologie française, pragmatique avant tout ? », *Pratiques philologiques en Europe*, éd. F. Duval, Paris, Ecole des Chartes, 2006, p. 115-150.

- DUVAL, Frédéric, « les éditions du XVII<sup>e</sup> siècle », *Manuel de la philologie de l'édition*, éd. D. Trotter, Berlin : De Gruyter Mouton, *Manuals of Romance Linguistics 4*, 2015, p. 369-394.
- DUVAL, Frédéric, *Les mots de l'édition de textes*, Paris, Ecole nationale des chartes, 2015
- MARCHELLO-NIZIA, Chrisitane, *Introduction à l'édition numérique de la Queste del Saint Graal*, publiée en ligne, <http://portal.textometrie.org/txm/texte/intro>, 2013.
- ROQUES, Mario, « Etablissement de règles pratiques pour l'édition des anciens textes français et provençaux », *Romania 52*, 1926, p. 243-249.
- SEGUIN, Jean-Pierre, *L'invention de la phrase au XVIII<sup>e</sup> siècle : contribution à l'histoire du sentiment linguistique français*, Paris, Bibliothèque de l'Information grammaticale, Éditions Peeters, 1993.
- VIEILLIARD, Françoise et GUYOTJEANNIN, Olivier, éd., *Conseils pour l'édition des textes médiévaux, tome 1. Conseils généraux*, Paris, Éditions du CTHS : École nationale des chartes, 2014.

#### Textes cités :

- Gerbert de Montreuil, *Continuation de Perceval* Quatrième continuation, édité par Frédérique LE NAN, Genève, Droz (Textes Littéraires Français 627), 2014.
- Guibert d'Andrenas*, édité par Muriel OTT, Paris, Champion (Les classiques français du Moyen Âge, 147), 2004.
- Guillaume de Digulleville, *Le dit de la fleur de lis*, édité par Frédéric DUVAL, Paris, École des chartes (Mémoires et documents de l'École des chartes, 95), 2014.
- La chevalerie Ogier : enfances*, édité par Muriel OTT, Paris, Champion (Les classiques français du Moyen Âge 170), 2012.
- Pétrarque, *Les Triomphes*, Traduction française de Simon Bourgouin, édité par Gabriella PARUSSA et Elina SUOMELA-HÄRMÄ, Genève : Droz (Travaux d'Humanisme et Renaissance, 495), 2012.
- L'Estoire de la guerre sainte*, édité par Catherine CROIZY-NAQUET, Paris, Champion (Les classiques français du Moyen Âge, 174), 2014.
- Perceforest*, 4<sup>e</sup> partie, tome 1, édité par Gilles ROUSSINEAU, Genève, Droz (Textes Littéraires Français 343), 1987.
- Perceforest*, 3<sup>e</sup> partie, tome 1, édité par Gilles ROUSSINEAU, Genève, Droz (Textes Littéraires Français 365), 1988.
- Perceforest*, 2<sup>e</sup> partie, tome 1, édité par Gilles ROUSSINEAU, Genève, Droz (Textes Littéraires Français 506), 1999.
- Perceforest*, 1<sup>ère</sup> partie, édité par Gilles ROUSSINEAU, Genève, Droz (Textes Littéraires Français 592), 2007.
- Perceforest*, 5<sup>e</sup> partie, édité par Gilles ROUSSINEAU, Genève, Droz (Textes Littéraires Français 615), 2012.
- Perceforest*, 6<sup>e</sup> partie, tome 1, édité par Gilles ROUSSINEAU, Genève, Droz (Textes Littéraires Français 631), 2014.
- Queste del saint Graal*, édité par C. MARCHELLO-NIZIA et A. LAVRENTIEV, Lyon, Equipe BFM, 2013. Publié en ligne par la Base de français médiéval, [http://catalog.bfm-corpus.org/qgraal\\_cm](http://catalog.bfm-corpus.org/qgraal_cm). Dernière révision le 2013-07-19.

Sophie PREVOST  
LaTTiCe (CNRS/ENS/Université Sorbonne Nouvelle)