

# ÉCRIRE EN CRITIQUE : EXPLORATION MORPHO-SYNTAXIQUE SUR CORPUS

Driss ABLALI  
LASELDI, Université de Franche-Comté

« C'est le langage qui parle, ce n'est pas l'auteur ». Mallarmé

## Introduction

Différentes études ont mis en évidence des variations systématiques des catégories morphosyntaxiques pour caractériser et discriminer les discours et les genres<sup>1</sup>. Cet article explore, à travers la comparaison de deux corpus de critique, critique littéraire et critique journalistique, les indices de divergence et de convergence entre discours que permettent d'observer les méthodes exploratoires de données textuelles. À partir d'une réflexion sur le genre, on cherche à mettre en évidence les traits discriminants permettant d'identifier le genre de la critique loin du carcan littéraire où il a longtemps été cantonné : existe-t-il une identité morphosyntaxique du genre de la critique ? Existe-t-il un style, des catégories grammaticales, des signes de ponctuation qui caractérisent l'article du journaliste et celui du littéraire et qui définissent une posture discursive originale ? Est-ce que ce qui fonde un genre est de l'ordre de l'autoproclamation de ce genre, comme cela a pu être le cas dans les théories littéraires, ou de l'ordre de la situation des discours ? Notre propos n'est évidemment pas de réhabiliter le critique, nous interrogerons le corpus pour découvrir les spécificités du discours dans lequel le même genre, à savoir la critique, s'inscrit, le discours des études littéraires et le discours journalistique. Il est temps en effet de remettre en cause la démarche littéraire qui, en disqualifiant la notion de discours, fait des genres le niveau le plus englobant de sa typologie.

Dans la perspective de la sémantique interprétative de F. Rastier<sup>2</sup>, nous distinguons quatre niveaux hiérarchiques supérieurs : les *discours* (ex. juridique vs littéraire vs essayiste vs scientifique), les *champs génériques* (ex. à l'intérieur du discours littéraire : théâtre, poésie, genres narratifs), les *genres* proprement dits (ex. comédie, roman « sérieux », roman policier, nouvelles, contes, mémoires et récits de voyage), les *sous-genres* (ex. roman par lettres). Au niveau inférieur de la classification, nous trouvons les textes de même genre et d'un même auteur.

En questionnant les modalités grammaticales et syntaxiques du genre de la critique, le but est d'identifier un régime singulier du critique littéraire et du critique journaliste. Quelles en sont les caractéristiques et sous quelles formes se manifeste-t-il ? Interroger le même genre au sein de deux discours différents, c'est donc analyser les liens qu'il tisse entre morphologie et syntaxe.

Notre travail portant sur le genre de l'article de critique, il importait, dans la perspective d'une étude contrastive en corpus, de pouvoir à la fois opposer la critique littéraire à la critique journalistique et en même temps de pouvoir montrer que le genre n'est pas envisageable sans discours. Nous explorerons d'abord la longueur du mot, facteur déterminant quant au rythme du texte. Un autre facteur révélateur du style d'un genre est la longueur de la phrase et le nombre de mots par proposition. Nous nous intéresserons par la suite à la segmentation à l'intérieur de la phrase en analysant les signes qui la représentent : la virgule, le point-virgule, les deux-points ainsi que les signes de parenthèse. Nous terminerons sur la question des catégories grammaticales, pronoms personnels et temps verbaux.

## Les paramètres du corpus

Le corpus compte 5 214 615 occurrences. Il comprend uniquement des articles intégraux et non des extraits. Il se répartit sur deux discours : littéraire et journalistique. 2 041 875 en critique littéraire, contre 3 116 740 en critique journalistique. Le premier comprend des articles tirés de revues françaises ou francophones relevant tous du domaine de la critique littéraire. Il contient 261 articles, publiés entre 1980 et 2004, extraits de 11 revues, comme *RITM*, *SEMEN*, *TEXTE*, *CAHIERS DE NARRATOLOGIE*, *CHAMPS DU SIGNE*, *ERITA*, *LECTURE LITTÉRAIRE*, *LOXIAS*. Le second comporte les articles de deux quotidiens français *LIBERATION* et *LE MONDE*, publiés

<sup>1</sup> Cf., entre autres, Rastier & Malrieu 2001, Beauvisage 2001, Loiseau, Poudat & Ablali 2006.

<sup>2</sup> Cf. Rastier 2001.

entre 2002 et 2003 dans la rubrique « critique », soit 1950 articles. Quantitativement, le corpus se présente ainsi :

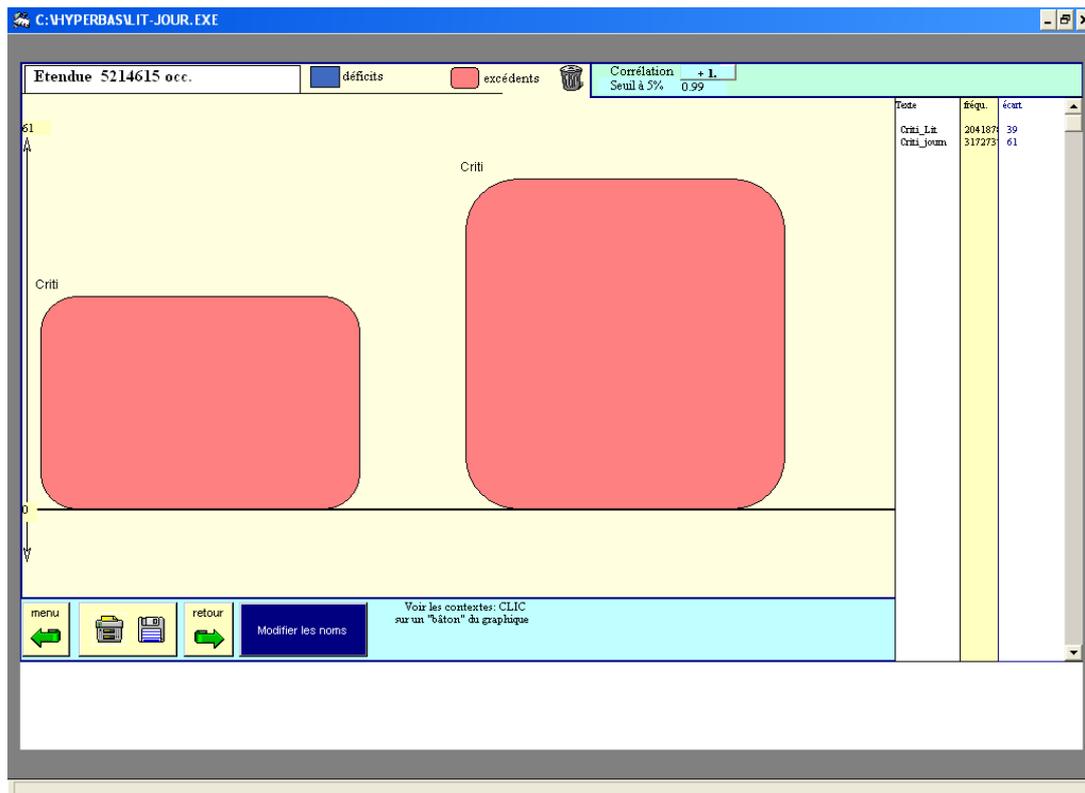


Fig. 1 : Étendue relative du corpus « Critique »

### Segmentation de la phrase critique

Afin d'opposer critique littéraire et critique journalistique, une première démarche consiste à comparer l'écart entre les deux pour chaque variable, en se basant sur les moyennes des valeurs des deux discours constituant les deux ensembles. Ici il s'agit de faire une exploration morphosyntaxique du corpus pour voir la manière dont sont agencées les phrases et les propositions, ainsi que l'examen de la distribution quantitative des mots à l'intérieur des segments, de la longueur des mots, tous rendus possibles par l'application des logiciels adaptés. Les données morphosyntaxiques dont nous disposons pour mener cette étude ont été produites par la société Synapse, à l'aide du logiciel Cordial.

Pour la longueur du mot, nous avons remarqué que le nombre moyen de lettres par mot est de : 4,65 pour le discours journalistique, et 4,89 pour le discours littéraire. Rien d'étonnant : la concision, c'est l'ergonomie du journalisme, c'est obtenir le même résultat informatif en moins de mots, en moins de phrases, moyennant moins de "bruit". « Dans le cas du français moderne, écrit C. Muller, le recours à un vocabulaire technique, savant ou simplement recherché augmente la proportion des mots longs, faisant ainsi monter notre indice ; un style familier ou relâché, sur le plan du lexique, agit évidemment en sens inverse. D'autre part une syntaxe soignée tend à réduire la densité du texte en mots grammaticaux, donc en mots courts, éléments que la langue courante multiplie » (1979 :153). Ainsi, la longueur du mot nous permet de confirmer l'opposition des deux discours, les mots longs se trouvant plutôt dans le discours littéraire, et les mots courts dans le discours journalistique. Et cette opposition générique se manifeste aussi au niveau de la proposition : pour le discours journalistique, le nombre moyen de mots par proposition est de 8,86, alors que pour le discours littéraire, il est de 10,48. Un autre facteur, plus souvent étudié que celui de la longueur du mot, est la longueur de la phrase, son unité supérieure, considérée comme un critère déterminant dans l'analyse stylistique et dans l'étude du rythme des textes. Comme critère stylistique fortement caractéristique d'un auteur, d'un genre ou d'un discours, elle permet ainsi d'appréhender des variations intéressantes, notamment d'un point de vue diachronique, pour étudier l'évolution de la longueur de la phrase dans plusieurs œuvres successives d'un même auteur, ou d'un point de vue contrastif, soit à l'intérieur d'un même discours, opposer par exemple

le portrait au reportage au sein du discours journalistique, soit à partir de discours différents, comme nous le faisons ici, opposer le discours journalistique au discours littéraire. La longueur moyenne de la phrase s'obtient en divisant le nombre d'occurrences du corpus par le nombre de ponctuations fortes. Lorsque l'on compare la longueur des deux discours, encore une fois le discours littéraire dépasse le journalistique. Ce dernier se caractérise par un nombre important de phrases courtes. Il enregistre une moyenne de 18,63 mots par phrase contre 22,02 pour l'article littéraire. Une différence de 3,39 est énorme, surtout lorsqu'il s'agit du même genre. C'est encore une fois la preuve que les caractéristiques stylistiques et syntaxiques ne dépendent pas du genre dans lequel le texte s'inscrit mais de la situation des discours. « À chaque discours, on peut faire correspondre un système ou *symmorie* générique. Chaque groupe de pratiques sociales correspondant à un discours se divise en activités spécifiques », nous dit F. Rastier (2001 : 4 ), qui fait ici le postulat que, par exemple, non seulement le discours juridique se différencie du discours médical ou du discours littéraire, mais qu'il faudrait trouver les variables qui discriminent de façon préférentielle chaque niveau de la hiérarchie : discours, champs génériques, genres, sous-genres. Ces variables, nous les retrouvons aussi au niveau de la proposition. Si l'on regarde l'histogramme ci-dessous, on voit sans équivoque que le discours journalistique est plus riche en propositions indépendantes, 88,23% contre 84,44 pour le discours littéraire. Car l'écriture journalistique requiert des règles et codes spécifiques qui structurent, organisent et codifient son contenu pour optimiser sa transmission au lecteur : comme l'espace est compté, le lecteur est présumé pressé, le style doit donc être concis. Et puisque le lecteur n'a pas de temps pour le verbiage, la chasse aux propositions relatives (1,86% vs 2,60%) et aux propositions subordonnées (8,63% vs 11,11%) devient une priorité. L'histogramme ci-dessous présentant les variations de la critique littéraire par rapport à la critique journalistique, montre l'ensemble des variables des deux discours :

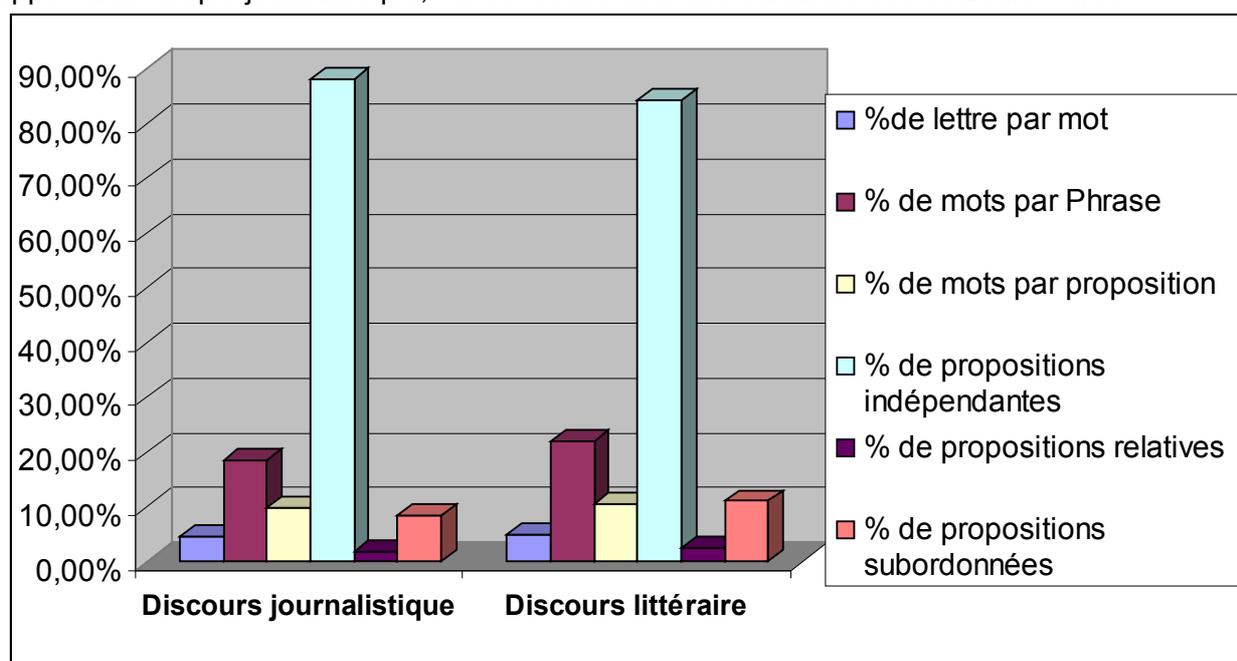


Fig. 2 : Segmentation de la phrase dans les deux discours

Toutefois, la longueur des mots et des phrases ne suffit pas à elle seule à caractériser les textes, ce qui nous amène à prendre en compte aussi les marques de ponctuation. Pour le point, principal marqueur de fin de phrase, nous observons que sa fréquence est plus élevée dans le discours journalistique. Plus de phrases courtes engendre automatiquement plus de points. Le pourcentage de points par rapport à l'ensemble des ponctuations est de 36,56 pour l'article journalistique, et de 22,34 pour l'article littéraire. Quant à la virgule, son effectif suit le cours logique des choses : le texte qui développe l'effectif le plus important en points fait usage de moins de virgules par rapport à l'autre discours. Dans notre corpus, la fréquence des phrases courtes est importante dans le discours journalistique, l'effectif de la virgule par rapport à l'ensemble des signes de ponctuation est de 45,88, alors qu'il est de 52,7 pour le discours littéraire. Or pour les autres marques de ponctuation, toujours fortes, comme le point d'exclamation et le point d'interrogation, c'est dans le

discours littéraire que nous trouvons les plus hautes fréquences, comme on le voit dans le graphique ci-dessous. L'article littéraire fait un usage beaucoup plus important du point d'exclamation que le discours journalistique. Comme le dit Culioli, le point d'exclamation entend signifier « le haut degré d'une propriété ». Le lecteur est lui-même appelé à le prendre en charge. Simplement il faut rappeler aussi que cette haute fréquence est liée à des fins intertextuelles. Le critique littéraire appuie sa thèse en recourant à des extraits des textes analysés. C'est ce qui justifie aussi la présence massive d'un autre signe de ponctuation, comme le tiret, lié à la présence de dialogue pour illustrer des propos, et qui est pratiquement absent du discours journalistique. Quant au point d'interrogation, il est dominant dans le discours littéraire, pas seulement pour la question de l'intertexte, mais surtout pour des raisons liées à la spécificité du genre de la critique littéraire, dont la thèse constitue le parangon. Cette constatation n'a rien d'étonnant, le critique écrit pour répondre à une problématique, développe des hypothèses, qu'il construit lui-même dans l'introduction de son article avec des interrogations soit directes, soit indirectes. Or dans l'article journalistique, on prononce des jugements sur les livres et les films qui paraissent, on tranche entre les bons et les mauvais sans en problématiser le contenu. On trouve aussi d'autres signes de ponctuation, au sein de l'article littéraire, qui semblent ainsi se rattacher à un pôle plus « scientifique », discriminé par une présence plus importante des crochets et accolades ou l'usage des parenthèses.

Le graphique ci-dessous, dont les résultats sont réalisés avec Cordial-Synapse, présente les caractéristiques des signes de ponctuation des deux discours :

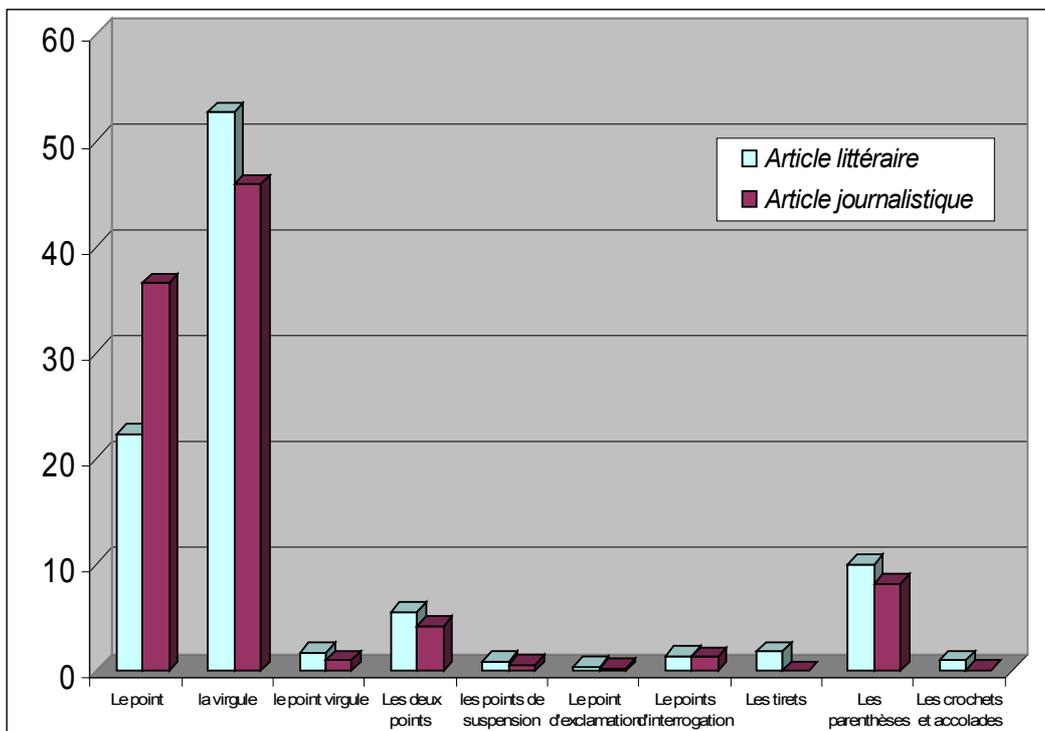


Fig. 3 : La distribution des signes de ponctuation (effectifs)

Sur la base de ces oppositions, on peut d'ores et déjà tenter d'esquisser au niveau phrastique un portrait-type de l'article journalistique par rapport à l'article littéraire. On note que les deux discours sont significativement distincts sur le plan morphosyntaxique : le discours journalistique exploite différentes techniques avec une dominance des mots courts, des phrases concises rejetant la coordination et la subordination, une diminution des virgules au profit des points, qui sont le signe d'une écriture plus incisive, à laquelle il faut sans doute rattacher des catégories grammaticales, des temps verbaux et pronoms personnels plus discriminants.

### Catégories grammaticales, temps et personnes

« C'est bien en effet dans la distribution des catégories grammaticales et principalement des classes nominale et verbale, que se manifeste la distinction des styles, des genres et des écrivains », écrit E. Brunet (1983 : 836), nous rajoutons des discours aussi, car il n'y a pas de genre sans discours. Il n'y a pas, par exemple, un discours épistolaire, il y a un genre épistolaire, comme la lettre, entre autres, qui prend des manifestations discursives différentes, selon qu'elle s'insère dans le discours des écrits ordinaires, comme le cas des lettres adressées par les allocataires à leur caisse d'allocations familiales, ou selon qu'elle est écrite par Valmont à Cécile dans *Les Liaisons Dangereuses*, comme c'est le cas dans le discours littéraire. « Investigating this question, écrivent Douglas Biber et ses collègues, can help to understand how different varieties exploit the grammatical categories of words available to them... Using the corpus, we can analyze the distribution and function of these different categories of words and study the part that they play in fulfilling the communicative function of different registers » (Biber, Conrad, Reppen 1998 : 57). La question à laquelle nous voudrions répondre dans ce dernier point s'inscrit dans le sillage de la citation de Biber : quelles sont les contraintes interprétatives mésosémantiques liées à la structure textuelle interne au genre de la critique ? Existe-t-il des signatures journalistique et littéraire en termes de catégories grammaticales, de pronoms personnels et de temps verbaux qui corroborent ce qui a déjà été discriminant ci-dessus ?

Commençons par l'étude des parties du discours. Quelles sont les particularités grammaticales qui contribuent au style de l'écriture journalistique et littéraire ? Le tableau ci-dessous présente la distribution des parties du discours pour les deux discours :

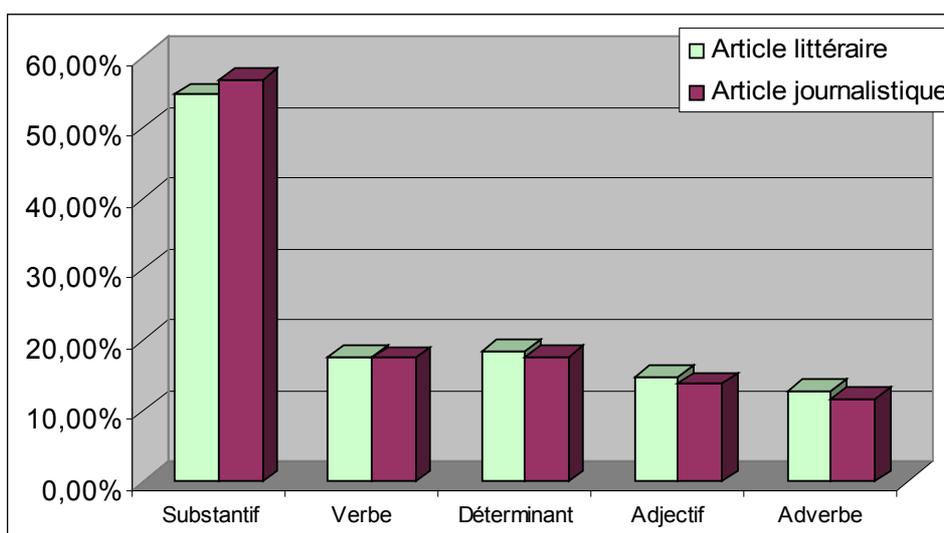


Fig. 4 : Classement hiérarchique de cinq catégories grammaticales

Première remarque : les deux discours sont riches en substantifs, une des caractéristiques des corpus de langue française. Pour le discours journalistique, la structure syntaxique la plus assimilable, la mieux comprise par le lecteur, est la structure habituelle : Sujet-Verbe-Objet. Cela correspond, comme on l'a vu ci-dessus, à la concision de l'idée. Le substantif est la catégorie grammaticale la plus employée dans les deux discours, avec une moyenne plus grande pour le discours journalistique : 56,87 contre 54,61. Pour les trois autres catégories grammaticales, les valeurs associées au discours journalistique sont plus faibles. Nous notons également la corrélation entre la fréquence du substantif et la fréquence des phrases courtes dans le discours journalistique, car, comme on le voit, le substantif journalistique est plus solitaire que son homologue littéraire. Les adjectifs et les adverbes sont plus dominants dans le discours littéraire : 14,83 contre 13,89 pour les premiers, et 12,93 contre 11,71 pour les seconds. Quant aux pronoms, notamment les pronoms personnels, car ce sont eux qui marquent la grande majorité des occurrences par rapport à l'ensemble des pronoms, 58,68 pour le discours littéraire, 60,89 pour le discours journalistique. Au niveau de la première personne aussi bien du singulier que du pluriel, l'article littéraire arrive en première position, contrairement à la troisième personne, qui domine dans le discours journalistique aussi bien avec « il » qu'avec « on » :

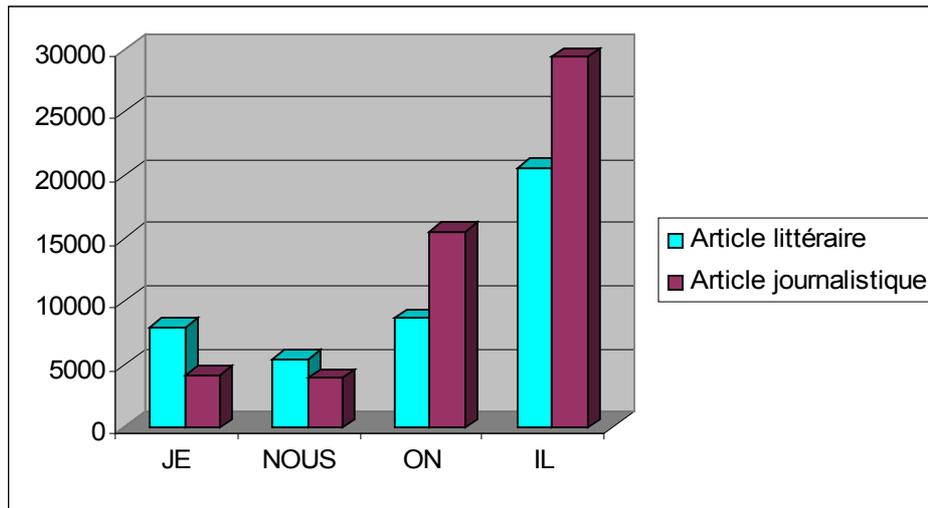


Fig. 5 : La distribution des pronoms dans le corpus

Dans cet histogramme on voit bien que « il » est le pronom personnel le plus dominant. Une seule explication se cache derrière cette haute fréquence : dans le genre de la critique, il est question indifféremment de textes et d'auteurs, c'est-à-dire d'objets ou d'êtres absents, ceux que le critique analyse en ayant recours à la 3<sup>ème</sup> personne, laquelle assure une reprise sémantique d'un précédent. Quant au « on » son statut est plus problématique. Le « on » par son caractère malléable et sujet à changer, permet d'alterner les points de vue. Il peut désigner l'auteur ou le journaliste comme responsable du propos, comme il peut désigner une communauté non déterminée de gens plus ou moins compétents dans le domaine concerné. Le « on » permet de prendre ses distances par rapport à l'événement ou à la problématique développés, de même qu'il permet d'exprimer explicitement une idée quelconque. Or lorsqu'on regarde le même histogramme en se concentrant sur la première personne, aussi bien du singulier que du pluriel, en examinant leur fréquence parmi l'ensemble des pronoms personnels de notre corpus, la différence entre les deux discours est plus aisément observable :

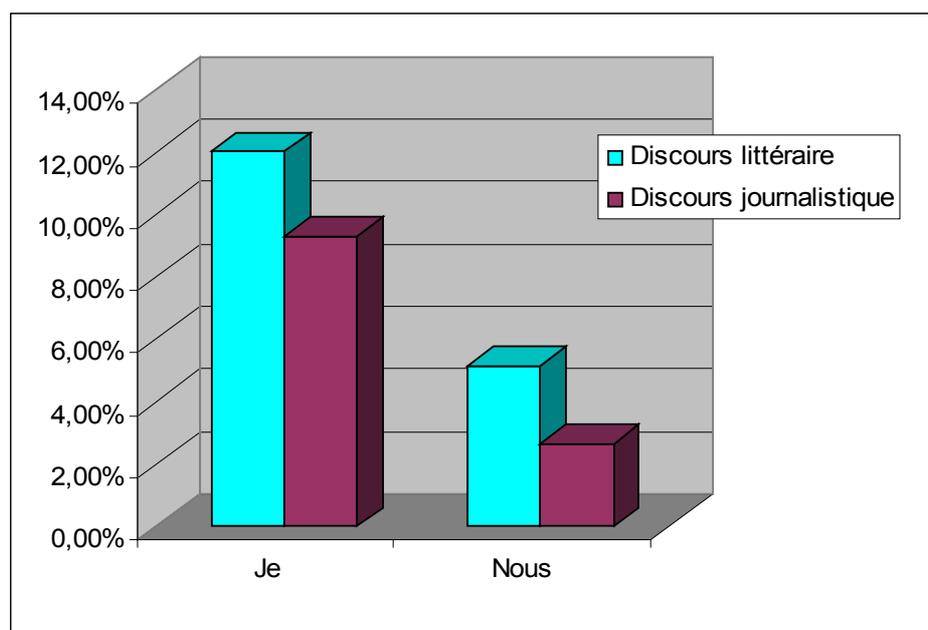


Fig. 6 : % de pronoms personnels « je » et « nous » parmi l'ensemble des pronoms personnels

Nous constatons que la grande majorité des occurrences est donnée par le discours littéraire, qui exploite de façon dominante la première et la deuxième personne. Cela est lié encore une fois à la

spécificité du discours. L'article littéraire s'inscrit dans un discours fortement conditionné par des contraintes normatives. Ce n'est pas une question de choix, comme tout énonciateur le fait lorsqu'il s'approprie le langage, mais une contrainte dictée par la nature du discours. La première personne permet ainsi d'opposer le point de vue critique-chercheur aux autres travaux précédents qui s'intéressent à la question étudiée. Et contrairement à ce qui a été montré par Loffler-Laurian<sup>1</sup> dans son étude sur le discours scientifique en chimie et en physique, le discours des études littéraires fait un usage systématique de la première personne car l'aspect des théories est subjectif et la valeur des hypothèses est personnelle. Or dans l'article journalistique, il n'est question ni de théories ni d'hypothèses. On écrit pour être lu par un large public. Le point de vue du journaliste importe peu, un spécialiste du cinéma n'intéressera que peu de lecteurs, les érudits comme lui, alors que dans le discours littéraire, la démarche est inverse : l'article, qui est l'image d'une école de pensée, expose avant tout la rigueur méthodologique et les préalables à la découverte scientifique. On s'adresse à des spécialistes du même domaine, on cherche à partager avec des collègues les résultats d'une recherche dans un langage spécialisé, on est loin de la vulgarisation.

Chaque discours a donc une forme canonique au niveau des pronoms, et les temps verbaux permettent de mieux comprendre l'usage qui en est fait. Les pronoms sont étroitement liés à une autre catégorie grammaticale, celle des verbes, dont la distribution dans notre corpus est représentée par le graphique suivant :

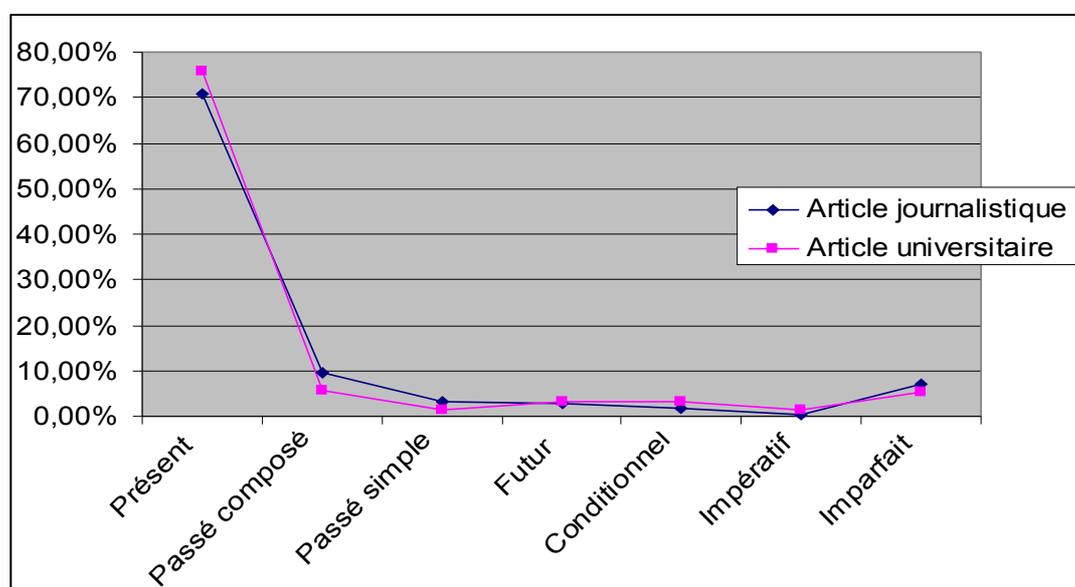


Fig. 7 : La distribution des temps verbaux dans le corpus

Avec l'aide de l'étiquetage de Cordial, nous avons pu effectuer une analyse de l'ensemble des deux discours. La distribution des différents modes de notre corpus montre un indicatif qui domine largement. Quant aux autres modes, ils sont minoritaires. Nous ne nous intéresserons ici qu'à certaines des sous-catégories qui nous ont semblées les plus discriminantes.

Au niveau des temps verbaux, nous observons premièrement une opposition au niveau de l'axe passé vs présent. Le discours journalistique se caractérise par un usage plus important des temps narratifs, comme pour le passé composé, (9,69% contre 5,63%), le passé simple, (3,04% contre 1,58%) et l'imparfait (6,91% contre 5,30%), tous les trois exprimant une temporalité passée. Quant à l'axe du présent, c'est le discours littéraire qui marque les plus grandes fréquences : un mode narratif à visée atemporelle typique des textes scientifiques, et d'ailleurs caractérisé par l'usage du présent de l'indicatif à hauteur de 75,87% contre 70,86% pour le discours journalistique. L'emploi du futur, dominant chez le littéraire, souvent dans l'introduction, avec une fréquence de 3,31%, est lié à des contraintes stylistiques, comme dans des expressions de ce genre : « nous nous attacherons... », « nous verrons ensuite », « nous nous limiterons », qui permettent d'anticiper sur les résultats de la recherche. Un dernier mode nous semble discriminant : l'impératif. Bien que

<sup>1</sup> Cf. l'auteur 1980, p.135.

l'impératif soit un mode très limité dans notre corpus, 0,30% dans le discours journalistique contre 1,35% pour le discours littéraire, sa distribution n'est pas sans intérêt. Le littéraire écrit souvent dans le but d'être compris, et il veut surtout s'assurer que son lecteur le suive. D'où l'emploi fréquent de l'impératif à des fins pédagogiques et démonstratives : « Ajoutons qu'au niveau dialogique... », « Commençons par cette dernière question... », « regardons de plus près la succession de... », « passons aux opérations interprétatives ».

### Conclusion(s)

L'objectif de cette étude n'était pas de dresser un parangon de chacun des deux discours. Elle n'avait pas non plus la prétention de traiter de tous les aspects sous lesquels on pourrait définir le genre de la critique. Mais elle aura montré que l'insertion d'un genre dans un discours donné n'est pas sans influence sur l'aspect morphosyntaxique et stylistique du texte. Car le facteur prédominant de ces divergences semble être celui du discours. En effet, le profil morphosyntaxique qui émerge de nos différentes analyses est celui d'une écriture qui exprime, en fonction de la situation des discours, à la fois les spécificités et la diversité du genre. Un genre oscillant entre une écriture concise et technique, entre un style privilégiant les mots courts et les points, la troisième personne, le passé composé et le passé simple (l'article journalistique), et une écriture pédagogique et normative (l'article littéraire), préférant les mots longs et la virgule, la première personne, le présent et l'impératif. Ces résultats nous permettent de voir le bien fondé des variables morphosyntaxiques dans la définition contrastive d'un profil-type de la critique, qui sont à même de donner une représentation objective des genres en fonction des discours dans lesquels le texte prend place. C'est donc sur une typologie des discours que se fondera une typologie des genres, à travers laquelle nous pourrions regrouper et typer les textes.

### BIBLIOGRAPHIE

- BEAUVISAGE, T. 2001. Exploiter des données morphosyntaxiques pour l'étude statistique des genres - Application au roman policier, *TAL*, 16.
- BIBER, D., CONRAD, S. & REPPEN, R. 1998. *Corpus linguistics, Investigating Language, Structure and Use*, Cambridge, Cambridge Approaches to Linguistics.
- BRUNET, E. (éd.) 1983. *Etude statistique des textes littéraires, Hommage à Pierre Guiraud*, Cumfid n°14, CNRS, Institut national de la langue française, URL 9 - Université de Nice octobre 1983.
- BRUNET, E. 1988. *Le vocabulaire de Victor Hugo*, Paris-Genève, Champion-Slatkine.
- LOFFLER-LAURIAN, A.-M. 1980. L'expression du locuteur dans les discours scientifiques. "JE", "NOUS" et "ON" dans quelques textes de chimie et de physique, *Revue de linguistique romane*, 44, pp. 135-157.
- LOISEAU, S., POUDAT, C. & ABLALI, D. 2006. Exploration contrastive de trois corpus de sciences humaines, *JADT*, Besançon, Les cahiers de la MSH Ledoux, pp. 631-642.
- MALRIEU, D. & RASTIER, F. 2001. Genres et variations morpho-syntaxiques, in Daille, Romary (dir.), *Linguistique de corpus*, *TAL*, vol. 42 n°2, Paris, Atala/Hermès, pp. 547-577.
- MULLER, C. 1979. *Langue française et linguistique quantitative, Recueil d'articles*, Genève, Éditions Slatkine.
- RASTIER, F. 2001. Éléments de théorie des genres, in <http://www.revue-texto.net>.
- RASTIER, F. 2001. *Arts et sciences du texte*, Paris, PUF.