

COMMENT PEUT-ON ANALYSER LES SIGNES TYPOGRAPHIQUES D'UN CORPUS ETENDU ?

KHELIFI Hadria

Université de Lorraine / France
hadriakhelifi@gmail.com

Résumé : Le présent article se veut exploratoire, il vise à montrer une démarche qui s'inscrit dans la linguistique du corpus, et qui pourrait être efficace pour témoigner d'une pratique numérique dans l'enseignement et/ou l'apprentissage du système d'organisation de l'écrit. La notion corpus a permis de rassembler des textes de domaines divers et de tailles volumineuses construisant un échantillonnage représentatif et homogène. L'analyse manuelle ou traditionnelle de ce type de corpus n'est plus efficace vu qu'elle nécessite une quantité de travail et de temps considérables. À cet effet, il est intéressant de s'interroger sur l'intérêt d'une analyse automatique d'un corpus dédié à la recherche des signes typographiques, à l'aide du logiciel approprié au domaine investi.

Mots-clés : analyse logométrique, discipline universitaire, écrit scientifique, genre, ponctuation.

Abstract : This article is meant to be exploratory, it aims to show an approach that fits into the linguistics of the corpus, and that could be effective to testify to a digital practice in the teaching and/ or learning of the organization system of writing. The concept corpus has brought together texts of various fields and large sizes constructing a representative and homogeneous sampling. In addition, the manual or traditional analysis of this type of corpus is no longer effective since it requires a considerable amount of work and time. To this end, it is interesting to wonder about the interest of an automatic analysis of a corpus dedicated to the search for typographic signs, using the appropriate software for the field invested.

Keywords : logometric analysis, academic discipline, scientific writing, genre, punctuation.

* * *

Depuis l'arrivée des nouvelles technologies de l'information et de la communication, les méthodes et les outils se sont diversifiés selon les domaines de recherche et selon les objectifs visés par les usagers, notamment que leur prise en main est devenue aisée. Aujourd'hui, l'enseignement et/ ou l'apprentissage des différents aspects de la langue à travers le numérique foisonnent le monde. Les moyens pédagogiques offerts entre distance et présence, réservés à l'autoformation ou à la formation sont variables en fonction des besoins variés des apprenants et des enseignants.

Le présent texte est purement méthodologique. À travers ce texte, nous révélons le besoin de recourir à la méthode par logiciel pour préparer et analyser un corpus multidisciplinaire de grande taille issu de douze mémoires d'études universitaires en sciences humaines et dans les filières que nous appelons communément « scientifiques ». Il s'agit d'une recherche sur les signes typographiques dans un écrit scientifique. Une approche qui peut ainsi être adressée au public en langue, en littérature, en sciences du langage et en didactique des langues pour explorer et analyser des articles de revues, des livres, des chapitres de livres, des abstracts, etc. E. Brunet souligne que le sens du texte, conditionné notamment par la syntaxe, n'est pas traitable automatiquement contrairement à la ponctuation qui peut tirer parti de ce type d'analyse (Brunet, 2016). Quant à l'interprétation des résultats, c'est la méthode logométrique qui organise le travail du

chercheur en proposant des pistes d'analyses distinctes, en triant le corpus et calculant les différents profils recherchés.

Grâce à cette analyse, nous avons la possibilité d'augmenter la validité des analyses du corpus contrairement à l'étude classique qui se base sur des règles de lecture, de comptage et d'interprétation. *Hyperbase* est un outil qui sert à étudier le style et le genre d'écrit d'auteurs, à créer des bases hypertextuelles avec les textes qu'on lui fournit¹, mais aussi à apprêter un corpus consacré à enseigner ou à apprendre une langue.

1. Articulation du corpus

Le corpus est à l'origine un texte que le chercheur a choisi d'analyser entièrement ou fragmentairement. Il est aussi le terrain privilégié de l'observation des pratiques langagières. Le texte devient un corpus quand le chercheur commence son étude théorique ou empirique. F. Rastier affirme : « Tout texte doit être rapporté à un corpus pour être interprété » (Rastier, 1987 : VI). C'est ce choix du texte à étudier et le début de l'analyse du chercheur qui mènent à l'appellation *corpus*.

Le corpus que nous analysons est un discours écrit, une suite de mots ayant une forme de phrases composées suite à une organisation typographique suggérée par les scripteurs et selon des marqueurs discursifs. Il dispose d'un contexte et de différents éléments assurant sa mise en ordre et sa cohérence.

L'un des premiers aspects à évoquer avant de procéder à l'analyse du corpus est la structure du genre des textes qui composent notre corpus car l'aspect homogénéité est un critère important pour l'élaboration du corpus. F. Rastier souligne que « la représentativité » (Rastier, 1987 : VI) et « l'homogénéité » (idem : VI) sont deux critères permettant de caractériser le corpus. Ce dernier n'est pas seulement une récolte de données sur lesquelles nous réalisons une étude mais un élément qui « doit obéir à des conditions de représentativité » (Ablali, 2009 : 3). Ainsi, afin d'avoir une interprétation convaincante, intéressante et désambigüe, il faut choisir un corpus appartenant à un genre unique. Grâce à l'uniformité et l'harmonie du corpus, le chercheur évite de se trouver face à de nombreuses caractéristiques génériques propres à chaque genre de texte dans le cas où il choisit d'étudier un corpus provenant de plusieurs genres. Selon F. Rastier, vu que le traitement automatique, comme le cas de notre étude, s'applique non à des phrases mais à des textes, la définition de leur genre est une nécessité (Rastier, 2011 : 71). Nous avons choisi ce corpus car il est caractérisé par un seul trait spécifique qui est son appartenance à l'écrit universitaire, à une seule pratique qui est la rédaction du mémoire d'étude, à un discours réalisé par la tranche estudiantine. Aussi, les rédactions de cette masse étudiante sont considérées comme des documents produits dans un seul espace, en l'occurrence, l'université. En terme de quantité, nous avons choisi ce grand corpus pour avoir un résultat satisfaisant. Le même auteur rajoute que pour que ce traitement soit efficace et simple, il faudrait spécifier « les fonctionnements propres aux différents genres » (idem : 72). Connaître le genre du corpus facilite la recherche d'information. Le corpus de référence est général, il se compose des documents non liés à un domaine spécifique. Il s'agit de la banque textuelle *Frantext*².

2. Préparation du corpus

La présente recherche mobilise un seul corpus multidisciplinaire, soit 12 mémoires de magister au total. Ce corpus est destiné prioritairement aux spécialistes des disciplines qui participent, à travers des mémoires, à la constitution de ce corpus. Elle est conduite par une étude de discours universitaire écrit par des étudiants de magister, ayant soutenus

¹ Pour en savoir plus sur ce logiciel, visitez le site : <http://ancilla.unice.fr>

² Est une base textuelle construite vers 1960 pour mettre en place un exemple de documents au Trésor de la langue française. Elle comprend presque mille œuvres des XIX et XX siècles, elle est constituée de textes scientifiques, techniques et littéraires. Ce dernier genre est le mieux représenté dans cette base. URL : <http://www.biu.sorbonne.fr/sid/IMG/pdf/Frantext_mode_d_emplois.pdf>, consulté le 01 mars 2018.

leurs mémoires d'étude entre 2005 et 2013 en Algérie. Ces mémoires, pris tels quels, portent sur les spécialités scientifiques, économiques, techniques, sociales et littéraires. Après avoir éliminé les parties : dédicace, sommaire, notes bibliographiques³, bibliographie, annexes, chaque texte du corpus se compose d'un nombre de page entre 56-224.

Les textes que constitue le corpus sont extraits d'un site Web⁴ sous le format PDF⁵. Le choix d'adopter le format .TXT nous a été imposé par le logiciel *Hyperbase*, utilisé pour un traitement automatique de notre corpus.

3. Choix du logiciel *Hyperbase*

Grâce aux explications fournies et aux exemples étudiés dans le manuel *Hyperbase*, nous avons été attiré spécifiquement par ce logiciel. *Hyperbase* est le plus approprié à notre recherche vu les diverses fonctionnalités que nous avons testé nous-même. Par exemple, il permet de réaliser une comparaison entre les textes du corpus pour engendrer des courbes et des listes de spécificités, établir le lien ou la déconnexion des textes, définir la quantité de l'écrit, calculer la haute fréquence des mots utilisés dans tout le corpus, étudier la distance lexicale entre les textes du corpus, définir la typologie du discours, repérer la terminologie disciplinaire, distinguer la position de l'auteur, exposer la richesse du vocabulaire et l'état de son évolution, etc. En outre, les nombreuses recherches enrichies à l'aide de cet outil comme celles de D. Ablali et B. Wiederspiel⁶, D. Mayaffre⁷, C. Poudat⁸, A. Tutin⁹, F. Grossmann¹⁰, S. Kastberg Sjöblom¹¹, etc. sont des exemples de son efficacité.

4. Traitement automatique des textes

Depuis le passage à l'écrit numérique, le traitement des textes via des logiciels impacte de plus en plus la méthodologie de la recherche dans différentes disciplines des sciences humaines. La raison de l'usage de ces outils vient de la nécessité d'analyser des corpus copieux, qui exigent une quantité considérable de travail. Ces mêmes outils sont également valables pour explorer et analyser des documents tel que : articles de revues ou de journaux, livres, chapitres de livres, abstracts, etc.) et avoir divers résultats, ce qui rend leur utilisation très alléchante.

Le logiciel assure un résultat hautement performant, satisfaisant et beaucoup plus ordonné qu'un travail manuel étant exécuté via une entité informatique intelligente. En effet, les logiciels sont connus par leur efficacité à permettre au chercheur de retourner facilement au texte qui constitue le corpus par un simple clic et à partir de n'importe quelle fonction. Il est ainsi possible de faire appel, dans le cas du logiciel *Hyperbase* de la version 9.0 dont nous nous sommes servis dans notre recherche, à des phrases, des paragraphes pour vérifier l'emplacement des formes, lire aisément et comprendre leur contexte. Le retour

³ Les notes bibliographiques sont à l'origine des bibliographies.

⁴<http://chamilo3.grenet.fr/dokeos/edaf/courses/TELETU/>

⁵ Portable Document Format ou Format de Document Portable

⁶ Ablali D., Wiederspiel B. 2015. Quand le lien numérique maintient le fil narratif des personnes en souffrance. URL : <https://www.cairn.info/revue-communication-et-langages1-2015-4-page-77.htm>, consulté le 2juin 2018.

⁷ Mayaffre D. 2000. *Le poids des mots*, Honoré Champion. Paris.

⁸ Voir la bibliographie, en bas d'article.

⁹ Tutin A. 2013. La phraséologie transdisciplinaire des écrits scientifiques : des collocations aux routines sémantico-rhétoriques, In : Tutin, A., Grossmann, F., *L'écrit scientifique : du lexique au discours*, [Presses universitaires de Rennes](http://www.universitaires.de Rennes), Rennes.

¹⁰ Grossmann F. 2011. Écriture scientifique et positionnement d'auteur. URL : <http://books.openedition.org/pressesenssib/950?lang=fr>, consulté le 2juin 2018.

¹¹ Kastberg Sjöblom M. 2013. Collocations et cooccurrences dans le dictionnaire bilingue : étude lexicométrique. URL : <https://www.cairn.info/revue-ela-2013-2-p-207.htm>, consulté le 2juin 2018.

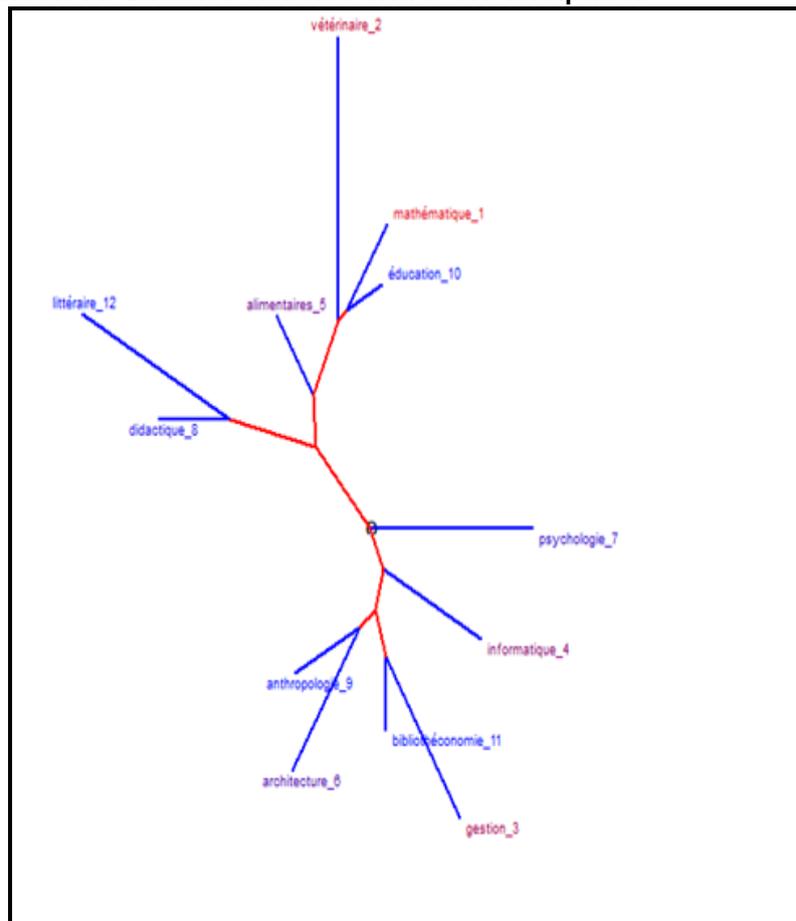
au texte est la seule solution pour l'interprétation du contexte. Les images ci-dessous sont une démontrent l'utilité de cet outil.

Figure 1 : Exemple de spécificités du corpus entier

N°	écart	corpus	texte	mot	N°	écart	corpus	texte	mot
565.97		736	975	formation	-69.21		1252134	136	je
524.44		104	339	processus	-67.24		8010689	20186	,
425.82		103	274	enseignant	-55.28		818177	127	vous
420.87		346	497	information	-50.86		1626354	2438	il
339.65		1540	850	développement	-44.37		868891	878	ne
289.43		93294	5972)	-43.44		987132	1240	qu'
279.28		143	212	objectifs	-42.52		1540412	2858	que
272.45		91853	5597	(-37.04		663862	761	elle
269.74		162	218	activités	-36.31		637831	731	n'
245.76		215	229	variable	-34.72		418479	262	lui

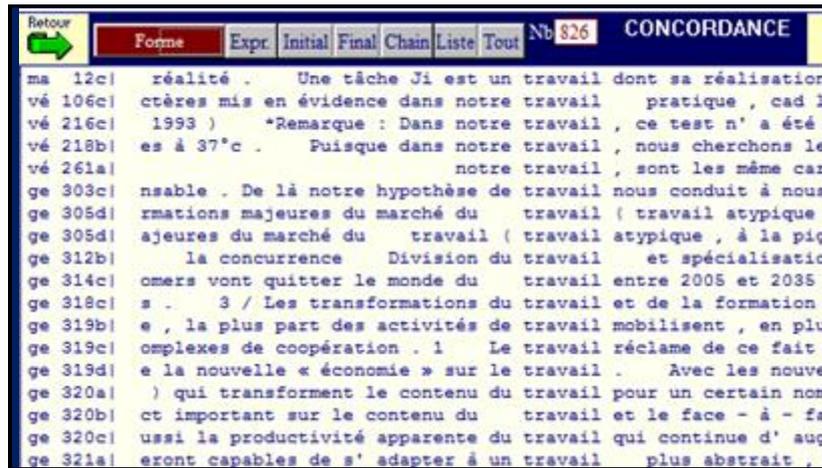
Source : Hadria KHELIFI, thèse de doctorat en sciences du langage en cours

Figure 2 : Connexion ou distance lexicale entre les disciplines



Source : idem.

Figure 3 : Exemple de contextes du mot *travail*



Source : idem.

Naviguer dans le corpus via le logiciel évite une recherche manuelle qui demande un temps conséquent au chercheur surtout si le corpus est très long. Le logiciel offre une flexibilité pour analyser le texte et mettre en évidence les liens existant entre son contenu et les informations recherchées. Mais cette approche par logiciel n'est pas exempte de difficultés telles que la collecte des données qui doivent être à la fois larges et diversifiées, ou les problèmes techniques qui peuvent entraver l'utilisateur.

Le traitement des données via le logiciel constitue une partie nécessaire dans le traitement du corpus mais qui s'avère insuffisante si elle n'est pas complétée par une autre qui est l'interprétation des résultats. C'est le passage d'une étape caractérisée par des formes, des chaînes, des graphiques, etc. vers une autre étape de constitution des idées en corrélats sémantiques.

4. Lecture des résultats

Les résultats fournis par le logiciel orientent l'analyste, notamment s'il est un non-spécialiste. D. Mayaffre, qui analyse le discours politique, n'est, à l'origine, qu'un historien. En effet, c'est le logiciel qui, en offrant diverses pistes de recherche, suite à ses différentes fonctionnalités et selon les composantes du discours (verbe, lemme, nom propre, adjectif, adverbe, ponctuation, mots outils, etc.), orientent le chercheur vers les différentes manières d'interpréter les résultats. C'est la méthode lexicométrique qui organise le travail du chercheur en proposant des analyses distinctes. En d'autres termes, le logiciel apprête le texte au chercheur en le triant, calculant les différents profils recherchés (quantité des signes typographiques, ponctuation basse, ponctuation fréquente, etc.) organisant toutes les données sous forme de tableaux, graphiques ou histogramme. L'exemple qui suit est un extrait d'une analyse que nous avons effectué sur le corpus cité plus haut, et dont l'objectif est de vérifier si la ponctuation employée correspond à celle réservée au genre scientifique.

Dans cette recherche, nous partons de l'hypothèse que la discipline aurait un impact sur la ponctuation du genre du mémoire d'étude. Quel est donc cet impact sur le genre d'écrit réservé à la rédaction universitaire scientifique ? Quel est le rapport qu'entretient le système de ponctuation avec les différentes disciplines qui constituent le corpus ? Quelles sont les des marques de ponctuation les plus présentes dans ce genre d'écrit ?

Selon le logiciel *Hyperbase*, parmi les signes typographiques présents dans le corpus : le point, point-virgule, point, deux points, virgule, parenthèse, tiret, crochets droits et accolades et certaines ponctuations à valeur énonciative comme les points de suspension, le point d'exclamation, le point d'interrogation. Ci-dessous le tableau des signes typographiques. Cette ponctuation peut dépendre de citations ou du style direct, des

éléments que nous n'avons pas éliminés lors de la phase de préparation du corpus. Aussi, les exemples exposés plus bas sont seulement une partie des exemples étudiés.

Figure 4 : Fréquences des signes typographiques dans le corpus

Corpus	Fréquences des signes typographiques par rapport à l'ensemble des ponctuations										
	!	?	;	:	,	()	-	[]	{ }
Anthropologie	865	33	0	24	24	108	1568	159	197	13	0
Architecture	1077	1125	0	61	319	526	2125	270	963	12	0
Bibliothéconomie	1308	16	1	42	19	293	889	133	282	2	0
Didactique des Langues	837	63	73	132	57	450	1303	156	608	0	0
Education physique et Sportive	1336	3	0	16	24	190	640	516	92	5	0
Gestion	3073	117	13	69	160	693	3147	521	1922	0	0
Informatique	1672	39	0	4	15	276	1065	242	135	202	
Mathématique	867	43	1	0	2	136	666	414	158	98	9
Médecine vétérinaire	1654	5	1	2	271	457	2696	847	945	7	0
Psychologie	2287	35	0	35	171	606	3211	1532	1654	3	0
Sciences alimentaires	1910	20	0	1	67	211	1636	691	459	0	0
Textes littéraires	622	64	19	10	90	134	1240	116	214	27	0
Total	17508	1563	108	396	1219	4080	20186	5597	7629	369	9

Source : idem.

Nous remarquons que seulement cinq signes sur onze montrent leur présence. Ces signes sont en ordre décroissant : virgule, point, tiret, parenthèse, deux points. Le nombre des occurrences du reste des ponctèmes (points de suspension, points virgules, points d'interrogation, points d'exclamation, crochets droits et accolades) varie entre 9 et 1563 du chiffre total des signes typographique. Ce résultat n'est pas surprenant, bien au contraire, il nous permet de les négliger ou de les considérer comme éléments « marginaux » (Poudat, 2006 :110). D'ailleurs, à l'exception des crochets et des accolades, ces signes (points de suspension, points virgules, points d'exclamation, points d'interrogation) dont la valeur est basse sont plutôt connus par une participation importante dans les textes littéraire.

5. La ponctuation basse dans le corpus

5.1. Les points de suspension

Marquent une interruption dans la phrase, un inachèvement de cette dernière qui n'influence pas toujours le sens de la phrase. L'auteur recourt aux points de suspension pour éviter l'usage abandonné de l'abréviation « etc. ». À partir du tableau, nous avons extrait deux modèles de ponctuation, en l'occurrence les points de suspension et le point virgule.

- Les points de suspension sont connus par leur emploi pour signaler une ouverture en fin d'un *listing*.

Exemple 1

Le musée d'art et d'ethnologie : où s'expose un ensemble d'œuvres d'art, tableaux, sculptures ■ choisies pour leurs intérêts stylistique, artistique, ou encore montrant les différentes phases de la carrière d'un artiste (musée Picasso par exemple), il concerne également les objets dont la réunion permet de mettre en avant les particularité d'un pays, d'une région, d'une époque (musée du costume, des arts et métiers, musée Bardo, ...)

(Mémoire d'architecture)

- Ce type de points est aussi utilisé pour annoncer une coupure au sein d'une liste. Dans un tel cas, les points de suspension sont bornés par des crochets droits.

Exemple 2

économie par exemple ne peuvent être compris et expliqués qu'en relation avec des phénomènes politiques, religieux, techniques [■], chaque élément prend signification de l'ensemble culturel et social dans lequel il s'insère. Le même ensemble social peut aussi saisi par d'autres disciplines avec lesquelles l'anthropologie entre en complémentarité.

(Mémoire d'anthropologie)

5.2. Le point-virgule : est considéré comme un ponctème marginal, mais d'une manière moindre par rapport aux autres signes typographiques. Il représente faiblement le genre décrit, *i.e.*, le genre scientifique d'un point de vue général. Dans ce corpus, le point-virgule représente des propositions dépendantes juxtaposées pour une explication :

Le point-virgule a une fonction essentiellement introductive et définitoire, comme celui des deux points, après lequel des éléments sont cités :

Exemple 1

Le développement rural met en rapport direct ou indirect une série d'acteurs relevant de catégories variées ■ paysans de statuts divers, jeunes femmes, représentants locaux et des agents de développement de terrain de différentes institutions, ces acteurs de développement,

(Mémoire d'anthropologie)

Le point-virgule renforce une définition précédente, informant de la continuité de l'idée précédente :

Exemple 2

La culture à Constantine revêt un sens patrimonial et historique particulier, riche de la richesse historique de la ville, et diverse de la multiplicité civilisationnelle qu'elle a connu, elle demeure indissociable de ses éminents savants ■ la culture prend ainsi un caractère patrimonial spécifique à la ville, fait de savoirs faire, de traditions, et de biens culturels matériels de valeur inestimable.

(Mémoire d'architecture)

6. Présence significative des signes typographiques dans le corpus

La ponctuation a un impact de sur le sens et la hiérarchisation de ce dernier dans un énoncé et sur la cohérence du texte, car c'est un système parmi d'autres qui indiquent les liaisons et les ruptures entre les différentes unités du texte : phrase, propositions, paragraphes. Ce système influence l'interprétation des phrases dans la mesure où elle dirige le lecteur, grâce à un ensemble de signes, à accomplir et enrichir sa compréhension des faits successifs qu'ils séparent. Il gère aussi, l'activité de lecture notamment en ce qui concerne le temps consacré à cette lecture. Ainsi, il attribue, chez certains écrivains, au style de l'écriture, son relief et sa lisibilité (phrase rythmée ou pas, nombre de propositions équilibré ou pas, etc.). C'est pour ces raisons que nous allons étudier l'effet de ce système plurifonctionnel sur le plan syntaxique et au niveau du sens.

À partir du tableau de la ponctuation entière du corpus qui figure plus haut, nous avons extrait deux signes typographiques les plus représentés. Il s'agit de la virgule et la parenthèse.

6.1. La virgule

Est le deuxième ponctèmes le plus relativement fréquents dans ce corpus. Ce marqueur de pause est représenté dans le corpus entier avec un chiffre de 20186 occurrences par rapport à l'ensemble des signes typographiques du corpus, et se montre dans tous les mémoires d'étude en question avec des valeurs entre 640 en mathématique jusqu'à 3211 occurrences en *psychologie*.

En effet, c'est « L'énumération, l'emploi de subordinées, les connecteurs, les appositions et les incises, qui augmentent la taille des syntagmes et des phrases constituent les lieux privilégiés d'utilisation de la virgule. » (Rinck F., 2006 : 218). La virgule est connue par sa présence élevée dans plusieurs genres (Rinck F., 2006 : 216) vu sa fonction nécessaire soulignée par E. Brunet : « Même le signe le plus léger, la virgule, peut jouer des rôles bien différents, tantôt en liant, tantôt en isolant les termes. » (Brunet, 2016) et selon les exemples suivants :

- La présence d'une virgule entre le sujet et le verbe, entre le verbe et l'adverbe, entre l'adverbe et la préposition, entre la préposition et la conjonction peut poser une ambiguïté, car elle ne doit pas y être. La première virgule dans l'exemple ci-dessous ne doit pas y être :

Exemple 1

Néanmoins cette ouverture sur le monde et d' évolution rapide des sociétés constitue une menace culturelle des plus redoutables car elle entraîne la standardisation culturelle , et la dévalorisation de tout patrimoine , essentiel pour l' affirmation identitaire et la pérennité de la mémoire de la ville .

(Mémoire d'architecture)

- La virgule est un moyen de segmentation alternative du texte en phrases chez certains scripteurs notamment quand leurs phrases sont longues ou s'il s'agit des périodes :

Exemple 2

Sur le terrain les éleveurs , dans toutes les catégories de la filière aviaire (chair , ponte , reproduction) sont fréquemment confrontés à des affections bactériennes , touchant l' intégrité de leur cheptels , en raison du taux de mortalité ou de morbidité entraînés par ces pathologies . Très souvent , il s' agit simplement des bactéries habituelles du tube digestif , principalement les espèces appartenant à la famille des Entérobactériaceae , qualifiées de marqueurs d' un déséquilibre intestinal qui dans des conditions de stress , mauvaise hygiène ou antibiothérapie

(Mémoire de médecine vétérinaire)

6.2.La parenthèse

Un signe typographique qui sert à introduire dans le discours une phrase nécessaire ou additionnelle pour le sens du texte comme il peut porter une idée sur un autre plan. Les parenthèses complètent la segmentation (Brunet, 2016), isolant une réflexion, un point de vue de l'auteur mais aussi elles isolent une demande d'intervention (avis, témoignage) du lecteur. E. Brunet souligne que « les parenthèses servent d'exutoire aux réserves, aux repentirs, aux précisions multiples qui accompagnent la pensée pensante » (Brunet, 2016). Ces précisions se diffèrent selon le besoin de l'auteur comme dans les exemples ci-dessous. Le même auteur rajoute que « La valeur esthétique de la parenthèse est peu appréciée des poètes [...] et peu prisée par la troisième personne » (Brunet, 2016). Dans cette recherche, on entend analyser la parenthèse que ce soit ouvrante ou fermante. Le nombre (5597) mentionné dans le tableau plus haut concerne la parenthèse ouvrante, et est égale au nombre total de la parenthèse fermante.

Les parenthèses sont liées à la mise des références des sources utilisées par l'auteur. C'est dans le cas où elles remplacent l'appel aux notes de bas de page. « Les parenthèses sont nettement associées aux dates et noms propres, eu égard aux pratiques de citation des

textes scientifiques généralement soumises au modèle (Nom propre Prénom, date, page).» (Poudat C., 2006,131). la taille de la parenthèse dépend du nombre des ressources citées à l'intérieur de celle-ci).

Exemple 1

De tout temps le travail a présenté deux visages opposés , l' un favorable , source de plaisir , facteur d' accomplissement , soutien de l' identité et ami de la santé , l' autre défavorable , source de souffrance et de possibles troubles divers (Moulin , 1980) . Source de création , de liberté et d' équilibre ainsi que source de contrainte et de dérives vers la maladie , voila le paradoxe du travail .

(Mémoire de psychologie)

Le contenu de la parenthèse peut être une mise en relief, une accentuation du contenu qu'il a précédé :

Exemple 2

Jusqu' à la seconde guerre mondiale , les structures sociales et la moindre circulation de l' information laissent pratiquement à la seule famille (et surtout la mère de famille) le soin de transmettre les valeurs alimentaires du groupe social . Actuellement , l' éducation nutritionnelle

(Mémoire des sciences alimentaires)

Selon les exemples cités plus haut, les signes de ponctuations les plus représentés sont, comme nous l'avons constaté plus haut : virgule, point, parenthèse, deux points. L'ordre d'apparition des signes de ponctuation repéré dans cette étude est le même ordre qui a été révélé dans l'étude de F. Rinck sur l'article (Rinck F., 2006). Ceci nous permet de confirmer que ce type de ponctuation est réservé non seulement au mémoire d'étude et/ou à l'article mais d'une manière plus générale au genre scientifique.

Conclusion

L'intégration du moyen lexicométrique dans l'analyse du discours universitaire s'avère nécessaire et très utile pour les corpus de taille consistante. Il fournit des résultats immédiats et complets en peu de temps, ce qui permet à l'analyste de mettre en place aussitôt les grands titres des démarches de son cours. C'est un outil heuristique vigoureux repérant diverses pistes de recherche dès le premier bilan.

Habituellement, l'analyse du discours se fait sur des fragments mais avec une analyse logométrique ou automatisée, aucune partie du texte n'est exclue ou oubliée. Aussi, la possibilité de retourner au texte de départ, de naviguer aisément entre le corpus analysé et les données conçues et de régénérer les résultats en cas d'erreur, promulgue cet outil comme un socle méthodologique automatique particulièrement pertinent et de très grand intérêt pour la recherche scientifique.

Sources bibliographiques

- ABLALI D. 2009. *Divorcer de la discipline : vers une nouvelle configuration générique de l'article scientifique*, l'Harmattan. Paris.
- BRUNET E. 2016. La ponctuation et le rythme du discours ; [en ligne] (d'après les données du *Trésor de la langue française*). Volume XXI - n° 1 (2016) URL : <<http://www.revue-texto.net>>, consulté le 01 janvier 2017.
- KHELIFI H. 2016. « La lexicométrie : un outil efficient pour l'analyse du discours ». In : Lachkar A., Diab-Duranton S., Amsidder A. *Médias numériques, langues, discours, pratiques et interculturalité*. Acte de colloque international du 24-26 octobre 2016. Faculté des Lettres et des Sciences humaines. Agadir (Maroc).
- POUDAT C. 2006. Étude contrastive de l'article scientifique de revue linguistique dans une perspective d'analyse des genres, Thèse de doctorat. Sciences du langage. Université d'Orléans. Orléans.
- RASTIER F. 1987. *Sémantique interprétative*. Presses Universitaires de France. Paris.

RASTIER F. 2011. *La mesure et le grain*. Honoré Champion.Paris.

RINCK F. 2006. L'article de recherche en Sciences du langage et en Lettres.Figure de l'auteur et identité disciplinaire du genre, Thèse de doctorat. Sciences du langage. Université Grenoble III. Stendhal Grenoble.