

Quelle place pour la modalisation dans le discours radiophonique religieux? Analyse lexicométrique des échanges discursifs.

What place for modalisation in religious radio discourse? Lexicometric analysis of discursive exchanges.

ZAOUI Meriem, TALEB Souad
Doctorante, zaouimeriem12@gmail.com
MCA, asou-ainsebaa@hotmail.fr

Recue: 2022-03-16

Accepter: 2022-12-01

publier: 29/12/2022

Résumé:	<i>informations sur l'article</i>
<p><i>La structure multimodale et pluridisciplinaire du discours médiatisé et ces mutations fréquentes, poussent à l'étude de l'apport des nouvelles méthodes d'analyse conformes à l'évolution du numérique d'où la lexicométrie comme méthode de traitement des discours. Or, cela peut sembler une voie intéressante de recherche: si une entité linguistique a légitimement une certaine structure, quelle approches accorder face aux nouvelles méthodes d'analyse? Et comment interpréter les données d'analyse? Plusieurs méthodes à travers les quelles le chercheur puise dans ce domaine, à l'instar des logiciels, des techniques statistiques et des nouvelles disciplines tel que la lexicométrie.</i></p>	<p><i>Reçu</i>/...../2020</p> <p><i>Acceptation</i>/...../2020</p>
	<p><u>Mots clés:</u></p> <ul style="list-style-type: none"> ✓ Discours médiatisés ✓ multimodale ✓ lexicométrie
	<p><u>Abstract :</u></p> <p><i>The multimodal and multidisciplinary structure of the mediated discourse and these effective mutations, lead to the study of the contribution of new methods of analysis adapted to the evolution of digital technology, hence lexicometry as a method of processing discourse. However, this may seem like an interesting line of research: if a linguistic entity legitimately has a certain structure, what approaches should be granted in the face of new methods of analysis? And how to interpret the analysis data? Several methods through which researchers draw on this field, such as software, statistical techniques and new disciplines like lexicometry.</i></p>
<p><u>Keywords:</u></p> <ul style="list-style-type: none"> ✓ Mediated discourse ✓ multimodal ✓ lexicometry 	

1. INTRODUCTION

Alors que les manifestations des corpus numériques sont de plus en plus étudiés, on ne sait encore que peu de choses sur les approches et les méthodes d'analyse appropriées. Or, cela peut sembler une voie intéressante de recherche: si une entité linguistique a légitimement une certaine structure, quelle approches accorder face aux nouvelles méthodes d'analyse? Et comment interpréter les données d'analyse? D'un point de vue théorique, cet ensemble d'interrogations, nous mène à penser aux approches établis par les théoriciens en matières d'analyse du discours mais aussi à l'influence des corpus (médiatisés, numériques, informatisés, etc.) récemment étudiés.

Plusieurs méthodes à travers les quelles le chercheur puise dans ce domaine, à l'instar des logiciels, des techniques statistiques et des nouvelles disciplines tel que la lexicométrie.

Par la lexicométrie, le mot est le noyau de l'analyse faisant de ce dernier une unité d'analyse dans son contexte donnant ainsi un prototype de listing de mots et des occurrences selon les points représentatifs assignés par le chercheurs qui sont par la suite comparables, hiérarchisés, spécifiés et analysables. Cette méthodologie représente une aide indispensable dans le traitement des discours et de la collecte des données, deux laborieuses démarches qui demandent de la persévérance et de la patience. À cet égard, nous proposons la problématique suivante; Dans quelle mesure une méthode lexicométrique peut-être utile dans l'analyse du discours numériques médiatisés?

La structure multimodale et pluridisciplinaire du discours numérique et ces mutations fréquentes, poussent à penser à l'apport des nouvelles méthodes d'analyse conformes à l'évolution du numérique d'où la lexicométrie comme méthode de traitement des discours.

Il est donc question d'exposer une réflexion méthodologique sur le traitement automatique des discours et de défendre une nouvelle démarche qu'est "la lexicométrie", cette notion que nous définissons dans les lignes qui suivent.

2. Méthodologie et présentation du champ d'étude

2.1 La lexicométrie

La procédure lexicométrique aide à comparer le lexique de différents émetteurs, de différents discours et de différents énoncés d'un même discours, d'analyser ce lexique à partir d'un corpus représentatif, homogène, dont les points sont comparables, au niveau du lexique, du genre et du contexte d'énonciation.

Pierre Lafon, Maurice Tournier et Robert-Léon Wagner, sont les précurseurs du courant lexicométrique vers les années 1970. En conséquence, quelques appellations sont à signaler dans ce domaine, telles que;

1. "textométrie" (André Salem),
2. "logométrie" (Damon Mayaffre),

3. "Analyse de discours assistée par ordinateur" (Pascal Marchand),
4. "Traitement automatisé du discours" (Pierre Fiala).

Il existe donc un « glissement de la lexicométrie originelle vers une logométrie pleine et entière, susceptible de renouveler la discipline » **Damon Mayaffre (2015)**.

En effet, les recherches menées dans ce domaine qui nous intéresse se fixent sur la notion du style en le considérant parmi les traits distinctifs lors d'une analyse. Également, la fréquence d'emploi des mots et leurs occurrences qui font la distinction entre les émetteurs d'où les indices d'étude suivants:

- la diversité du vocabulaire,
- la spécialisation,
- la richesse,
- la longueur et la structure des phrases,
- la densité des catégories grammaticales.

2.2 Typologie de sélection de corpus

Le corpus est donc défini selon le dictionnaire de didactique des langues comme un « ensemble fini d'énoncés pris pour objet d'analyse [...] réunis pour servir de base à la description et, éventuellement, à l'élaboration d'un modèle explicatif [...] il s'agit donc d'une collection de documents, soit oraux (enregistrés et/ou transcrit), soit écrits, soit oraux et écrits » (**Galisson et Coste 1976 : 131**).

Par ailleurs, du point de vue d'une phénoménologie du langage, l'accès vaste à de nouvelles données multimodales, souligne l'urgence de reconstituer une approche récente à base de corpus réel susceptible d'être analysé, traité et interprété, et s'opposant à des données ad hoc construites au préalable à base de théories ou d'études.

Toute émission linguistique a ses caractéristiques selon la situation de communication, et le fait de nier cette dernière, c'est faire comme s'il n'y avait pas d'altérité entre une émission culinaire et une autre sportive à la télévision.

À ce propos, se joint **M. BAKHTINE (1984, [1953]: 285)**; tout en étant convaincu par cette conception et l'utilité de marquer le point sur cette divergence qui existe et qui permet non seulement de donner un sens au discours (corpus) mais aussi une classification des genres de ce dernier.

« [...] Les formes de langue et les formes type d'énoncés, c'est-à-dire les genres du discours, s'introduisent dans notre expérience et dans notre conscience conjointement et sans que leur corrélation étroite soit rompue. Apprendre à parler, c'est apprendre à structurer des énoncés (parce que nous parlons par énoncés et non par propositions isolées et

encore moins, bien entendu, par mots isolés). Les genres du discours organisent notre parole de la même façon que l'organisent les formes grammaticales (syntaxiques). [...] autrement dit, dès le début, nous sommes sensibles au tout discursif qui, ensuite, dans le processus de la parole, dévidera ses différenciations. Si les genres du discours n'existaient pas et si nous n'en avions pas la maîtrise, et qu'il nous faille les créer pour la première fois dans le processus de la parole, qu'il nous faille construire chacun de nos énoncés, l'échange verbal serait quasiment impossible. »

Cette citation mène à penser à la situation dont laquelle le corpus est préconstruit (lorsque le locuteur émet son énoncé) et aux particules d'interaction, d'où l'apparition d'un autre phénomène celui de l'expérience du locuteur, de sa position et de sa prise en charge énonciative le moment de la production des discours (corpus). Cela en découle aisément, la notion de cadre référentiel du corpus qui elle aussi est au centre d'un environnement descriptif.

Précisons encore que la conceptualisation de tel ou tel corpus n'est présente que par son appartenance à un lieu qui le détermine et l'associe dans un cadre référentiel.

En s'inspirant des travaux de **B. PINCEMIN (1999)**, nous pouvons conclure que les genres de corpus sont multiples et leur classification paraît imprécise, vu l'hétérogénéité des discours par rapport à différent sphère de l'activité humaine et leurs caractères délimitables et en constante interaction, mais aussi par rapport aux critères de classement (signifiante, acceptabilité et exploitabilité).

Ce qui nous révèle vers l'existence de série de corpus, de façon à pouvoir satisfaire le goût du jour par l'adaptation des nouvelles formes de traitement et l'inclusion des différentes données, qui à leurs tours participent dans l'évolution des pratiques langagières et donnent lieu à de nouvelles études et découvertes. Tout en s'appuyant sur l'idée «*qu'un corpus est une collection de données langagières qui sont sélectionnées et organisées selon des critères linguistiques explicites pour servir d'échantillon du langage.*» (**Habert, Nazarenko, Salem 1997, p. 11**).

Les conditions ci-dessous représentent des repères importants dans les études de corpus tels que l'analyse conversationnelle, dont chacun s'articule aux deux autres par le biais de certains paramètres d'où la nature (*combinaison, distribution, composition, dilatation*), le lieu, la situation...etc.

➤ **Conditions de signifiante:** Un corpus est conçu en vue d'effectuer un travail de recherche construit à base d'idées antérieures (*pertinence*), ayant des questions de recherche, une problématique, des hypothèses, un objet d'étude déterminé et une méthodologie d'analyse (*cohérence*).

➤ **Conditions d'acceptabilité:** Le corpus collecté présente une acception authentique et fidèle (*représentativité*), évitant ainsi toute forme d'intervention externe (*régularité*), afin de respecter le principe de scientificité, allant du plus général au plus spécifique et arrivant à des réponses compatible avec nos questions de départ (*complétude*).

➤ **Conditions d'exploitabilité:** Les corpus d'étude choisis doivent être de nature identique (**homogénéité**). Un corpus doit comporter un nombre important d'éléments capable d'être analysés (**volume**).

2.3 Choix du corpus d'analyse

Partant de cette typologie, nous tenons à présenter à partir d'une classification de sous corpus, notre corpus constitué de conversations communicatives, observables et authentiques de type d'interaction radiophonique médiatisée de vulgarisation religieuse, tout en se basant sur un ancrage méthodologique allant du concept de «**corpus**» jusqu'à la constitution de différentes typologies.

Le concept de «données» en terme d'un travail de recherche renvoie à «*tout échantillon de discours-en-interaction supposé représentatif du/des phénomène(s) à étudier*». (C. **KERBRAT-ORECCHIONI, 2005: 24**). C'est pour cela que le chercheur en analyse du discours et en analyse conversationnelle se fixe comme premier objectif, la collecte des données pour l'élaboration de son corpus, souvent recueillie à l'oral, après avoir effectué son observation sur le terrain, sa sélection, son enregistrement et sa transcription.

Ces données peuvent être soit naturelles et authentiques (enregistrées en situation réelle) ou élicitées, construites par le chercheur (entretien, interview...etc.)

Quant aux données qui constituent notre champ, il est question d'une émission radiophonique de vulgarisation islamique, que nous avons sélectionné et organisé en contexte réel, sans aucune modification de notre part; c'est ce qu'en fait, **F. NEVEU** appelle **corpus défini**: «*Une collection de ressources langagières sélectionnées et organisées à partir de critères linguistiques, explicites et destinées à servir d'échantillons représentatifs*». (2004:86).

Le travail que nous estimons présenter, s'effectue sur un corpus constitué d'enregistrement à distance de données authentiques de l'émission "**Au coeur de l'islam**" réalisée par une animatrice qui, selon le thème de discussion, fait intervenir toutes personnes concernées dans le plateau d'une radio nationale (radio Alger chaîne 3). Il s'agit donc d'un enregistrement, à l'insu des participants ce qui donne un caractère d'authenticité à notre corpus et s'échappe de toute falsification possible. Or, les interactants essaient d'intervenir tout en gardant une certaine distance et intimité à partir du moment où ils savent qu'ils sont sur les ondes d'une radio, qu'un large public peut écouter; ils effectuent donc une conversation exceptionnelle. Quant à l'animatrice elle est plutôt dans le professionnalisme. Diffusée tous les vendredis de 6h à 7h et rediffusée de 14h à 15h, «**Au coeur de l'islam**» est une émission, interactive de type «*antenne ouverte*» ou «*phone in*» qui traite de tout ce qui se rapporte à la religion. Elle revient sur différents préceptes de l'islam face à une multitude d'avis religieux et sur des réalités identitaires auxquelles nous pouvons tous être confrontés.

Le concept de l'émission est pris de son nom «**Au coeur de l'islam**», dont le but est très sérieux puisqu'on vise non seulement de vulgariser la religion mais aussi et surtout de créer

un espace de réflexion autour des valeurs innées et de l'esprit Malékite. Ce thème va être discuté avec le co-énonciateur qui est l'animatrice de l'émission *Bouchra YAHIA*.

De par son triple caractère: *scientifique, utilitaire, vulgarisateur*, elle constitue une dimension institutionnelle médiatisée pour un large public écoutant.

3. Traitement numérique du discours

Des l'arrivée du numérique, le passage au traitement des discours via des logiciels devient indispensable pour une bonne méthodologie de recherche en matière des sciences humaines. Du besoin d'analyser des corpus volumineux est née la nécessité d'apporter l'outil informatique dans le traitement automatique des textes.

Ainsi, le logiciel présente un répertoire nettement performant, ordonné et satisfaisant par rapport à une étude manuelle. À cet égard, le logiciel, qualifié de rapide et intelligent, permet de modifier et de revenir facilement aux données d'analyse par un simple clic.

Par ce traitement numérique, l'analyse des textes et des discours connaît une rénovation, la vision n'est plus linéaire mais plutôt transversales par rapport aux lexiques étudiés. Ceci dit, qu'il est permis au chercheur d'avoir une lecture riche de ces résultats tout en ayant plusieurs interprétation du même texte par une situation de communication donnée. Par ailleurs, le traitement via le logiciel ouvre plusieurs voix aux analystes et s'échappe de toutes les techniques imposées et fermées.

Toutefois, les sciences sont indispensable d'où le rapport entre l'analyse du discours et les statistiques. Système, représentativité, exhaustivité et automatique sont les maîtres mots dans la lexicométrie.

La source de cette discipline est le calcul ou ce qu'on appelle l'inventaire de tout le lexique dans un corpus. Ces travaux quantitatifs ont pour objectifs une analyse du discours purement linguistiques à travers les marques énonciatives récurrentes.

En ce qui nous concerne, nous étudions les modalisateurs, faisant une étude des occurrences de celle-ci à travers l'approche lexicométrique. À travers son appellation, le but est de mesurer la qualité du lexique après avoir quantifier les mots répétés dans un même contexte. Le logiciel tel que *Trope*¹, procède par le découpage du corpus puis le calcul des mots dans les unités linguistiques, leurs occurrences, etc.

¹ Un logiciel d'analyse sémantique, apparut en 1994, avec les travaux de Pierre Molette et Agnès Landré, sur la base des travaux de Rodolphe Ghiglione. C'est un outil de fouille de textes qui permet de catégoriser la terminologie, le style, l'époque ou le temps... etc.

4. *Analyse statistique: Classification de la modalisation*

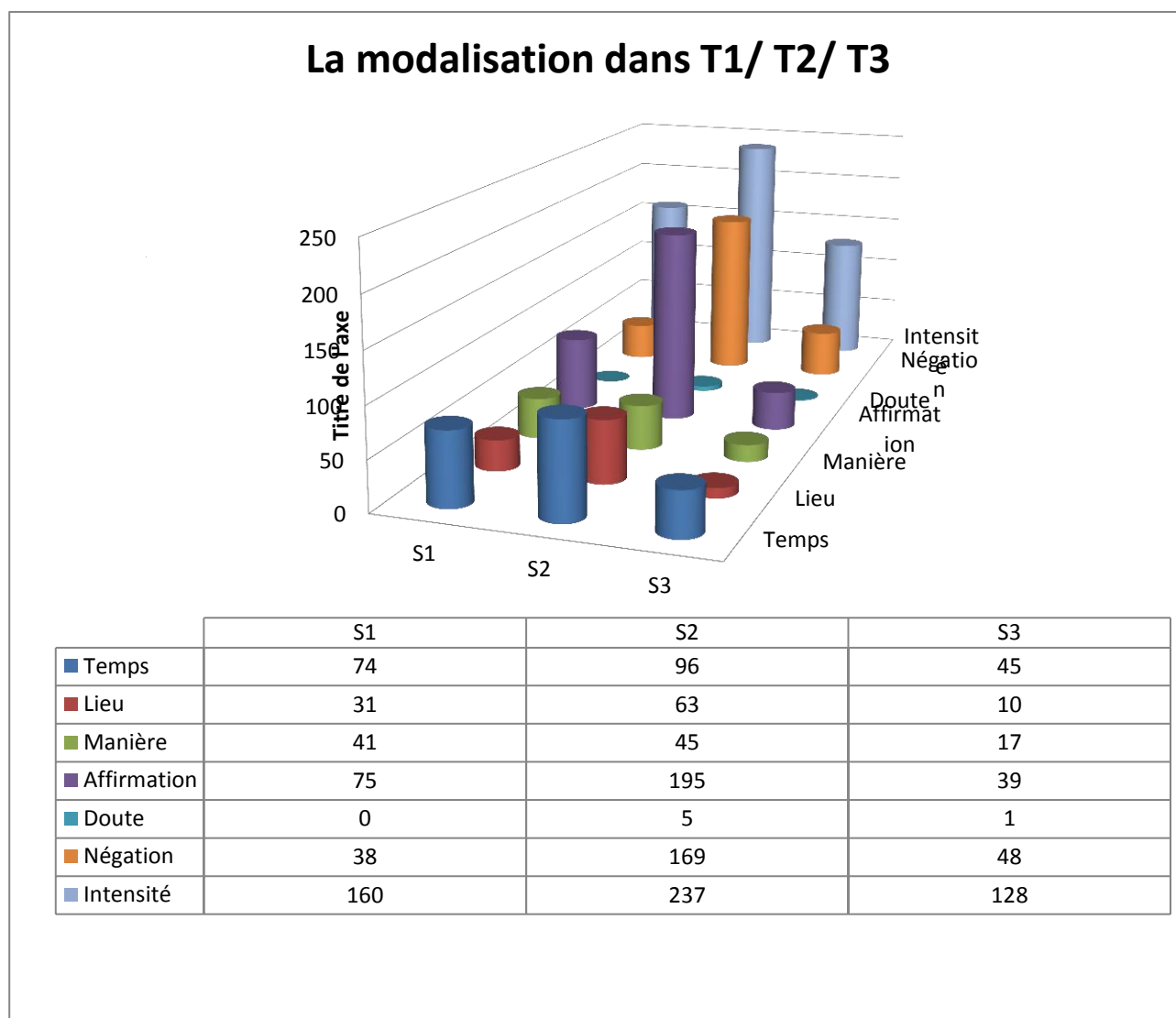
Un tel travail n'aurait pu être mené efficacement si nous nous ne sommes pas référés à une étude quantitative, qui vise à analyser des données quantifiables par le biais des outils statistiques reposant sur un système informatique. En effet, cette approche nous a permis de faire la relation entre notre conception méthodologique théorique (observation et méthode, concepts et notions) et les données statistiques relatives à l'étude qualitative. Ces données collectées sont issues d'un travail de terrain, accompli à travers une procédure automatique de traitement de texte qui s'ajustait avec nos besoins et permettait de fixer les priorités avec une large visibilité, ce dispositif est appelé «*Tropes* ».

À travers les résultats que nous a fournis le logiciel, illustrés dans le tableau ci-dessous (fig.01), nous remarquons que *la modalisation* traitée à une grande échelle dans l'émission "*Au cœur de l'islam*", est expliquée sous forme de variables statistiques en chiffres et en pourcentages, ces données recueillies, nous ont offert une vision descriptive analytique qui nous a permis de catégoriser les différentes modalisations existantes dans notre corpus, de les déterminer les unes vis-à-vis des autres et de repérer les points nécessaires à traiter.

Tableau N° 1. La modalisation par catégorie dans T1/T2/T3

	Contexte			Modalisation						
	animatrice	invités	Auditeurs (passifs)	Temps	Lieu	Manière	Affirmation	Doute	Négation	Intensité
1^{ère} thématique La foi				96	63	45	195	5	169	237
				11.9%	7.8%	5.6%	24.1%	0.6%	20.9	29.3%
Ouverture de l'émission	Bouchra YAHIA	Kamel CHEKKAT		44	20	15	68	0	23	76
				17.9%	8.1%	6.1%	27.6%	0.00%	9.3%	30.9%
Corps de l'interaction		/		2	1	1	3	0	2	11
		/		10.0%	5.0%	5.0%	15.0%	0.00%	10.0%	55.0%
		/		2	2	0	6	1	16	15
		/		4.8%	4.8%	0.00%	14.3%	2.4%	38.1%	35.7%
		/			0	1	1	2	1	0
		/			0.00%	14.3%	14.3%	28.6%	14.3%	0.00%
		/		1	1	0	0	0	1	8
		/		9.1%	9.1%	0.00%	0.00%	0.00%	9.1%	72.7%
		/			0	3	1	1	3	0
		/			0.00%	30.0%	10.0%	10.0%	30.0%	0.00%
Clôture de l'émission	Bouchra YAHIA	/		4	4	1	7	0	5	7
				14.3%	14.3%	3.6%	25.0%	0.00%	17.9%	25.0%
				0	0	0	5	0	0	2
2^{ème} thématique L'unicité de dieu	Bouchra YAHIA			41	35	24	103	3	119	118
				9.3%	7.9%	5.4%	23.3%	0.7%	26.9%	26.9%
3^{ème} thématique La fois en les anges	Bouchra YAHIA			74	31	41	75	0	38	160
				17.7%	7.4%	9.8%	17.9%	0.0%	9.1%	38.2%
				45	10	17	39	1	48	128
				15.6%	3.5%	5.9%	13.5%	0.3%	16.7%	44.4%

Fig. N° 1 : Schématisation de la modalisation dans T1/ T2/T3



Nous remarquons à travers cet échantillon que la modalisation est répartit en sept catégories;(temps, lieu, manière, affirmation, doute, négation, intensité), et que chaque catégorie est distinctes de l'autre différemment dans les trois thématiques (*T1/T2/ T3*), ce qui est explicité dans le graphique ci-dessus (*Fig.02*). Ainsi, la hauteur des barres informe sur le degré élevé d'une telle ou telle modalisation dans les trois thématiques T1/ T2/ T3.

La représentation graphique des données annonce le sommet pour la modalisation d'intensité dans *T1/ T2/ T3*, avec un intervalle approximatif mais qui s'affiche le plus dans **T2**, cela peut être expliqué par la nature du sujet «*l'unicité de dieu*» qui déclenche une réaction de l'opinion publique avec un haut degré d'intervention du spécialiste, étant touché par cette dernière. Par ailleurs, nous examinant la modalisation d'affirmation en second lieu, qui marque aussi une dominance dans T2.

C'est pour cela, notre étude sera penchée particulièrement vers ces deux catégories, marquant la variable la plus remarquable; la modalisation d'affirmation et la modalisation

d'intensité, manifestés sous différentes catégories (verbes, adjectifs, adverbes, locutions etc.). En ce qui nous concerne, nous focalisons notre analyse sur le repérage «*des adverbes*», ce point de langue qui permet de modifier le sens d'un énoncé, a attiré notre attention vu sa fréquence dans notre corpus.

Ce travail ne peut être mené qu'à travers une autre opération statistique, pour être en mesure de présenter une efficacité, justifiée par un recensement et une réalité fondée à partir d'hypothèses. C'est ce que nous pouvons démontrer dans une autre étude, à travers l'étude des occurrences de différents types d'adverbes.

5. Conclusion

En somme, le canevas structural que nous avons proposé a permis de présenter hiérarchiquement la structure d'une interaction spécifique à l'émission «*Au coeur de l'islam*», une émission numérique médiatisée de vulgarisation religieuse; ceci dit, que les résultats auxquels a abouti notre travail, ne sont valables que pour le corpus que nous avons adopté, sachant qu'un même corpus pourra faire l'objet d'énormes travaux.

De surcroît, il convient de souligner que dans le choix de l'objet de recherche, en général et dans la sélection du corpus d'étude en particulier, *les Conditions de signifiante, d'acceptabilité et d'exploitabilité* s'interpellent. La présence de l'un interroge l'autre ce qui se présente en flagrant dans le corpus proposé. Ceci, faisant appel à un ensemble de critères diversifiés et de méthodes d'analyse qui combinent entre le réel, le fidèle et l'exploitable. Tous cela, pour parvenir à proposer une méthodologie efficace de sélection de corpus et de méthodologie de traitement des discours.

En guise de conclusion, nous avouons que cette étude n'est qu'un commencement d'une recherche. Par ailleurs, elle peut constituer une ouverture à d'autre champ d'investigation abordant l'usage de tous types de corpus.

Liste bibliographique:

Livres

- BAKHTINE, Mickaël. 1984. *Esthétique de la création verbale*. Gallimard. Paris.
- CORBIN, Danielle. 2012. *Méthodologie de la théorisation enracinée: fondements, procédures et usages*. Presses de l'Université. Québec.
- GUMPERZ, John. 1982. *Engager la conversation Introduction à la sociolinguistique interactionnelle*. Les éditions de Minuit. Paris.
- GUIDÈRE, Mathieu. 2004. *Méthodologie de la recherche*. Ellipses.
- KERBRAT-ORECCHIONI, Catherine. 2005. *Le Discours-en-interaction*. Armand Colin. Paris.
- Lebart L., Salem A., 1994, *Statistique textuelle*, Paris, Dunod.
- Mayaffre D., 2015, *Textes augmentés et discours numériques*, Actes du colloque international du 14-19 septembre 2015 : Texte et discours en confrontation dans l'espace

européen pour un renouvellement épistémologique et heuristique, Université de Lorraine, Metz.

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

-Muller C., 1992, *Principe et méthodes de statistique lexicale*, Paris, Honoré Champion,

-POCHET, Bernard. 2003. *Méthodologie documentaire: comment accéder à la littérature scientifique à l'heure d'Internet?* De Boeck. Belgique.

Articles

-BOMMIER-PINCEMIN, Bénédicte (1999). «Définir un corpus». [En ligne], http://www.revue-texto.net/Corpus/Publications/pincemin_ad_1999.pdf [consulté le 24 septembre 2018].

-CHARAUDEAU, Patrick (2009). « Dis-moi quel est ton corpus, je te dirai quelle est ta problématique », *Corpus*. n°8, p. 37-66.

-HABERT, Benoît, NAZARENKO, Adeline, SALEM, André (1998). «Les linguistiques de corpus», *Persée*. [En ligne], https://www.persee.fr/doc/hism_0982-1783_1998_num_13_3_1758 [consulté le 15 novembre 2018].

-KERBRAT-ORECCHIONI, Catherine (2012). «Le contexte revisité», *Corela*. [En ligne], HS-11, (le 02 avril 2012). <http://journals.openedition.org/corela/2627> [consulté le 24 avril 2018].

Dictionnaire

-DUCROT, Oswald, et Jean-Marie, SCHAEFFER. 1995. *Nouveau Dictionnaire Encyclopédique des Sciences du Langage*. Éd. du Seuil. Paris.

-GALISSION, Robert. & COSTE, Daniel. 1976. *Dictionnaire de didactique des langues*. Paris : Hachette.

-NEVEU, Frank. 2004. *Dictionnaire des sciences du langage*. Paris: Armand Colin.

Site web

Damon Mayaffre, « De la lexicométrie à la logométrie », *L'Astrolabe*, 2005, p. 1-11, <https://halshs.archives-ouvertes.fr/hal-00551921/document>, site consulté le 5 octobre 2015.