

L'efficiency de l'outil informatique dans l'étude du texte littéraire

The efficiency of the computer tool in the study of literary text

Seghier Souad¹

¹ Université de Relizane, souad.seghier@univ-relizane.dz

Reçu: 24/ 05/ 2025 Accepté: 30/ 11/ 2025 publié 16/12/2025

Abstract:

In this study, we focused on content analysis that relies on lexicometric methods and tools. We will emphasize the theme of torture in Yasmina Khadra's novel "What Do Wolves Dream of?" using computerized procedures specific to content analysis. Our analysis will consist of two stages: the first (thematic study via frequencies) and the second (thematic study via concordance). The latter will allow us to highlight the thematic environment of the key word "torture". By the end of this research, we will demonstrate that occurrences reveal a theme, and the theme manifests itself in an occurrence.

Keywords:

Torture ; occurrence ; thème ; violence ; lexicometric.

L'efficacité de l'outil informatique dans l'étude du texte littéraire

Résumé:

Dans cette étude, nous nous sommes intéressés à l'analyse de contenu qui s'appuie sur des méthodes et outils lexicométriques. Nous allons mettre l'accent sur le thème de la torture dans le roman de Yasmina Khadra « A quoi rêvent les loups ? » en se servant des procédures informatisées propres à l'analyse de contenu. Notre analyse comportera deux étapes : la première (étude thématique via les fréquences) et la deuxième (étude thématique via la concordance). Cette dernière nous permettra de mettre en lumière l'environnement thématique d'un mot-pôle « torture ». Nous parviendrons au terme de cette recherche de montrer que les occurrences révèlent un thème et le thème se manifeste dans une occurrence.

Mots clés :

Torture ; occurrence ; thème ; violence ; lexicométrie.

1. Introduction:

Depuis l'apparition de la technologie numérique, la méthodologie de recherche en sciences humaines s'est beaucoup développée. L'usage des logiciels lexicométriques représente une nouveauté dans la manière de traiter les corpus textuels notamment les textes volumineux. Cela permet d'analyser, en un laps de temps relativement court, de grands ensembles de données en déterminant les motifs récurrents et en mesurant les variables linguistiques de façon performante.

Dans cette investigation, nous allons nous intéresser à l'analyse de contenu qui s'appuie sur des méthodes et outils lexicométriques. En effet, le logiciel assure un résultat plus rigoureux, plus performant et plus rapide qu'un travail manuel.

Cette analyse thématique de contenu est attachée à la recherche du sens d'une manière objective en utilisant des méthodes statistiques et quantitatives basées sur la comparaison des occurrences. Elle s'appuie sur le postulat suivant : les thèmes sont décelables à travers des expressions et des occurrences. Mais, inversement, ces manifestations concrètes et hétérogènes ne prennent sens qu'en référence aux thèmes qui les sous-tendent (Opler, 1945 ; Ryan & Bertrand, 2003). Ce qui veut dire qu'une occurrence révèle un thème, et le thème se manifeste dans une (ou des) occurrence(s). Dans cette contribution, nous allons mettre l'accent sur le thème de la torture dans le roman de Yasmina Khadra « A quoi rêvent les loups ? » en se servant des procédures informatisées propres à l'analyse de contenu. Notre problématique est la suivante : comment la distribution et la récurrence des unités lexicales permettent-elle de déceler la thématique de la torture ? Notre analyse comportera donc deux étapes : la première (étude thématique via les fréquences) et la deuxième (étude thématique via la concordance). Cette dernière nous permettra de mettre en lumière l'environnement thématique d'un mot-pôle « torture ».

2. Analyse automatique des textes

L'analyse automatique des textes est une approche issue de l'alliance entre la linguistique, l'informatique et la statistique. Elle vise à traiter et interpréter de manière systématique de grands ensembles de données textuelles à l'aide d'outils informatisés. Cette méthode, qui s'inscrit dans le champ de la linguistique de corpus, permet d'explorer les régularités lexicales, syntaxiques et sémantiques présentes dans un texte ou un ensemble de textes.

L'efficacité de l'outil informatique dans l'étude du texte littéraire

L'objectif principal de cette démarche est de décrire objectivement le fonctionnement du langage à travers l'observation des fréquences, cooccurrences et réseaux lexicaux. Grâce aux logiciels de traitement automatique, le chercheur peut repérer des tendances discursives, mettre au jour des structures récurrentes, et établir des corrélations significatives entre les éléments du discours.

Cette approche présente un double avantage :

1. Elle réduit la subjectivité de l'interprétation en s'appuyant sur des mesures quantitatives.
2. Elle accroît la portée de l'analyse en rendant possible l'examen de corpus volumineux, souvent inaccessibles à une lecture manuelle exhaustive.

L'analyse automatique ne se limite toutefois pas à la simple quantification des occurrences ; elle vise à dégager le sens à partir des régularités linguistiques. En combinant les dimensions quantitative et qualitative, elle contribue à une compréhension fine des discours, qu'ils soient littéraires, politiques, médiatiques ou sociaux.

En somme, l'analyse automatique des textes représente une méthode rigoureuse et interdisciplinaire, à la croisée de la linguistique computationnelle et des sciences humaines numériques, offrant au chercheur des outils puissants pour explorer la structure et le sens des productions textuelles contemporaines.

3. Hyperbase

C'est un logiciel d'analyse textuelle et lexicométrique développé à l'Université de Nice sous la direction d'Étienne Brunet. Il figure parmi les outils les plus anciens et les plus robustes dans le domaine de la linguistique de corpus. Conçu initialement pour l'étude des textes littéraires, il s'est progressivement imposé comme un instrument incontournable pour l'analyse automatique du discours dans les sciences humaines et sociales.

L'un des atouts majeurs d'Hyperbase réside dans sa capacité à gérer de vastes corpus textuels tout en offrant une analyse statistique approfondie. Il permet de calculer les fréquences lexicales, d'établir des concordances, d'identifier les cooccurrences significatives et de repérer les particularités lexicales d'un texte ou d'un auteur. Ces fonctionnalités facilitent la mise en évidence de tendances stylistiques, thématiques ou idéologiques au sein d'un corpus.

Hyperbase propose également des visualisations graphiques et des représentations statistiques telles que :

- les nuages de mots et cartes lexicales pour observer la distribution des unités dans le corpus ;
- les analyses factorielles de correspondances (AFC) pour comparer les profils lexicaux de différents textes ou auteurs ;
- les études de spécificités permettant d'identifier les mots caractéristiques d'un texte par rapport à un ensemble de référence.

Sur le plan méthodologique, l'utilisation d'Hyperbase permet d'allier rigueur quantitative et lecture interprétative. Le chercheur

L'efficience de l'outil informatique dans l'étude du texte littéraire

peut ainsi objectiver des hypothèses de lecture tout en conservant une approche qualitative, notamment à travers l'analyse des contextes d'apparition des termes.

Grâce à ces fonctionnalités, Hyperbase s'impose comme un outil scientifique de référence pour l'exploration lexicale, stylistique et thématique des textes, et demeure un pont entre la linguistique computationnelle et l'analyse littéraire traditionnelle.

4. L'approche théorique d'Opler (1945)

Selon Morris Opler (1945), l'analyse thématique repose sur le principe que les thèmes d'un texte se manifestent à travers des expressions linguistiques récurrentes — c'est-à-dire des occurrences de mots, de syntagmes ou de structures sémantiques. Autrement dit, les thèmes ne sont pas nécessairement déclarés de manière explicite, mais se déduisent à partir de la répétition et de la cooccurrence de certains éléments du discours.

Ainsi, dans une perspective lexicométrique, une occurrence révèle un thème, car la fréquence et la distribution d'un mot témoignent de l'importance conceptuelle qu'il revêt dans le texte. Inversement, le thème donne sens aux occurrences, puisqu'il constitue la catégorie interprétative qui permet de relier entre elles les unités linguistiques observées.

Ce principe est à la base de l'analyse de contenu thématique, qui vise à dégager les structures de sens latentes d'un corpus à travers la quantification et la contextualisation des occurrences. En d'autres termes, le travail du chercheur consiste à :

1. Repérer les récurrences lexicales et sémantiques,
2. Les regrouper en ensembles cohérents,

3. Les interpréter comme indices de thèmes dominants ou sous-jacents.

Ainsi, pour Opler (1945), « les thèmes sont décelables à travers des expressions et des occurrences », mais ces dernières ne prennent sens qu'en référence aux thèmes qui les sous-tendent — établissant une relation circulaire entre les signes linguistiques et les structures de sens qu'ils véhiculent.

4. Méthodologie de recherche

Dans le cadre de cette recherche à visée lexicométrique, nous adoptons la théorie d'Opler (1945), selon laquelle les thèmes d'un texte se révèlent à travers les expressions et les occurrences linguistiques. L'analyse thématique que nous proposons repose ainsi sur l'examen systématique des formes lexicales afin de dégager les structures de sens qui traversent le roman de Yasmina Khadra.

Pour la mise en œuvre de cette approche, nous recourrons au logiciel Hyperbase, un outil d'analyse textuelle qui permet de traiter de grands corpus de manière quantitative et qualitative. Ce logiciel fournit la liste exhaustive des formes présentes dans le texte, accompagnée de leur fréquence d'apparition, ce qui constitue la première étape de notre analyse.

Dans un premier temps, nous procéderons à la sélection des mots pleins à forte fréquence, car en lexicométrie, la distribution et la récurrence de ces unités lexicales constituent des indicateurs pertinents de la thématique dominante d'un texte. Dans un second temps, nous examinerons leur environnement

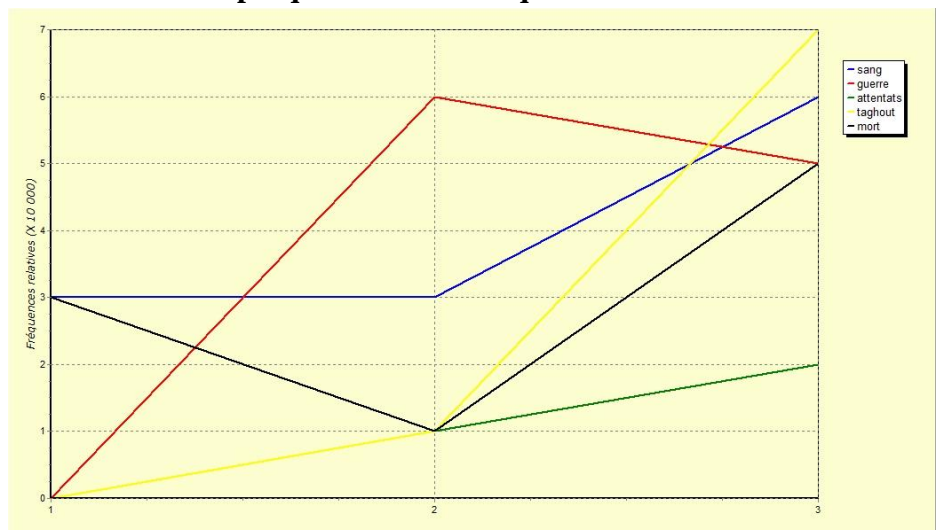
L'efficacité de l'outil informatique dans l'étude du texte littéraire

lexical à l'aide de la fonction de concordance, qui permet d'observer les mots apparaissant à proximité du terme étudié. Cette opération vise à identifier les réseaux sémantiques et à comprendre comment chaque mot-pôle participe à la construction thématique du roman.

6. Analyse du corpus

Selon les statistiques de la fonction dictionnaire du logiciel hyperbase, le roman de Yasmina Khadra comporte 12466 formes. Les occurrences de haute fréquence sont : guerre (31), mort (26) morts (12), taghout (27), sang (32), attentats (11). Le graphique de ventilation pour la partition (chapitre 1, 2 et 3) ci-dessous montrent la redondance de ces occurrences qui représentent la thématique de la guerre et de la violence.

Graphique N° 1. Les fréquences relatives



Le graphique montre que le mot *sang* apparait dans le premier chapitre avec une fréquence moyenne. Sa fréquence évolue, à partir du deuxième chapitre pour atteindre le maximum de fréquence au dernier chapitre. Le mot *guerre* est répété dès le premier chapitre, il atteint le plus haut degré dès le début du deuxième chapitre. Le mot *attentat*, quant à lui, il n'est cité qu'au deuxième et au troisième chapitre avec une fréquence moyenne. Le mot *taghout* apparait tout au long du roman avec une légère évolution du premier au second chapitre pour atteindre le maximum d'occurrences au dernier chapitre. Enfin, le mot *mort*, apparait d'une fréquence moyenne dès le début du roman ; une légère présence de ce mot est constatée dans le deuxième chapitre suivie d'une forte présence dans le dernier chapitre.

6.1. Environnement lexical du mot/morts

Concordance de : mort/morts

Les escadrons de la	mort	investissaient les douars
confirmaient la tragédie. La	mort	frappait partout. Tous les
islamique. Ton père est	mort	, lâchement assassiné. Il
. Si ce dernier est	mort	, mon frère doit l'

L'efficience de l'outil informatique dans l'étude du texte littéraire

sa chambre. Depuis la	mort	de son frère, elle
: – Omar Ziri est	mort	? – Et comment !
ai pas revue depuis la	mort	de mon père. –
Tapies dans un angle	mort	. Ignorant que leurs es
il est tombé, raide	mort	. Son cœur a lâché
Sidi Moussa. Il est	mort	dans mes bras. Tu
tandis que vous semez la	mort	et le feu, dis
sang ; il était la	mort	en marche. Cousin de
être destitué et condamné à	mort	par le Conseil national.
cour martiale et condamné à	mort	pour intelligence avec
Abdel Jalil lutte contre la	mort	avec l'énergie du désespoir
. Abdel Jalil n'est	mort	que depuis une semaine.

Nous avons choisi le mot “**mort**” comme mot pivot de notre étude, dans la mesure où il figure parmi les unités lexicales les plus fréquentes du corpus, avec un total de trente-huit occurrences. Cette fréquence élevée justifie son statut central dans la construction thématique du texte et témoigne de sa valeur symbolique dans l'univers narratif de Yasmina Khadra.

Afin d'effectuer une analyse approfondie et rigoureuse de l'item *mort*, nous procéderons, dans un premier temps, à une étude de **concordance** permettant d'identifier les mots cooccurents — autrement dit, les éléments lexicaux constituant son environnement linguistique immédiat (voisinage proche) et son environnement discursif élargi (voisinage lointain). Cette étape

visé à déterminer la nature et la régularité des associations lexicales autour du terme étudié.

Dans un second temps, nous analyserons les relations sémantiques et affectives entre le mot pivot *mort* et ses colocataires lexicaux, en les classant selon leur valeur connotative. Ce classement distinguera deux catégories principales : méliorative et péjorative.

6.1.1. Collocation des mots-outils

Nous observons que les cooccurents du mot “**mort**” sont, dans la majorité des cas, des mots outils qui le précèdent. Parmi eux, on relève principalement l’article défini “**la**”, employé à douze reprises. Il introduit le substantif “mort” dans des groupes nominaux du type “la mort de + complément”, exprimant la disparition d’une personne précise (*la mort de Hanane, la mort d’un individu, la mort de son frère, la mort de mon père*, etc.). Dans d’autres occurrences, le syntagme “la mort” est suivi d’un verbe par exemple *la mort frappait partout* où *la mort investissait les douars*. Ces deux énoncés partagent une valeur sémantique commune, celle de la douleur et de la généralisation du phénomène mortel.

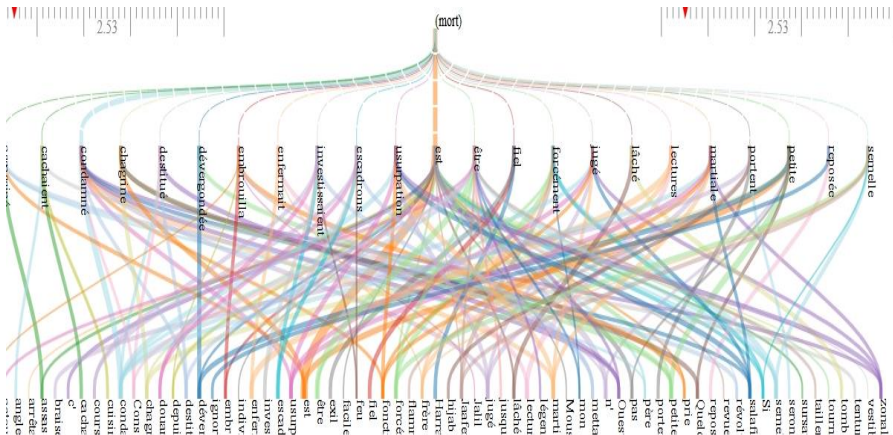
Les articles définis *le* et *les* renvoient, quant à eux, à des victimes identifiées, connues du locuteur, tandis que les articles indéfinis (*un, des, de*) désignent des victimes anonymes ou non spécifiées. Enfin, la présence d’adjectifs possessifs tels que *leur* ou *nos* précédant le mot “mort” renforce la dimension affective et

référentielle du discours : ils renvoient toujours à des victimes appartenant au champ personnel ou collectif du locuteur.

6.1.2. Collocation des mots pleins

Parmi les mots pleins, amis du mot *pôle mort*, il y a le verbe être qui apparaît 6 fois et utilisé pour annoncer la mort d'untel comme dans la phrase « Handala est mort ». Dans d'autres phrases, le mot mort est un substantif précédé de l'article défini *la* et suivi de la préposition *de* pour annoncer la mort d'untel qui est souvent proche comme le montrent les exemples (la mort de son frère, de mon père). Parfois, c'est le chiffre qui précède ce mot *pôle* tels que (vingt et un morts, douze morts, etc.) et servent à recenser le nombre de victimes. Les mots voisins proches ou lointains de ce mot *pôle* font également montre de la torture et de la violence (blessés (3 occurrences), l'énergie du désespoir, six disparus, enterre, chagrine, condamné, son cœur a lâché, le drame, l'allégeance, etc.)

Figure N° 2. Cooccurrent de *mort*



La relation entre le mot pivot **“mort”** et ses cooccurrents lexicaux tels que *condamné*, *assassiné*, *chagrine*, *fiel*, *feu*, *flammes*, *braises* constitue un réseau sémantique cohérent évoquant la mort sous ses diverses manifestations émotionnelles et symboliques.

Par ailleurs, l’association du même mot avec des termes comme *révolution*, *drame*, *désespoir*, *exil*, *usurpé*, *usurpation*, *choqué*, *destitué*, *forcément* met en évidence une extension sémantique vers le champ de la violence, de la perte et du bouleversement social.

Ainsi, le lexique de la mort se relie ici étroitement à celui de la souffrance collective et de la désintégration du monde humain, révélant une forte charge émotionnelle et idéologique dans le discours.

Indice de spécificité	Fréquence corpus	Fréquence partie	Taille corpus	Taille partie	Mot
--------------------------	---------------------	---------------------	------------------	------------------	-----

L'efficiencia de l'outil informatique dans l'étude du texte littéraire

4,13	636	14	88585	520	est
3,68	4	2	88585	520	condamné
2,61	25	2	88585	520	petite
2,53	2	1	88585	520	assassiné
2,53	2	1	88585	520	cachaient
2,53	2	1	88585	520	chagrine
2,53	2	1	88585	520	destitué
2,53	2	1	88585	520	dévergondée
2,53	2	1	88585	520	embrouilla
2,53	2	1	88585	520	enfermait
2,53	2	1	88585	520	escadrons
2,53	2	1	88585	520	fiel
2,53	2	1	88585	520	forcément
2,53	2	1	88585	520	investissaient
2,53	2	1	88585	520	jugé
2,53	2	1	88585	520	lectures
2,53	2	1	88585	520	lâché
2,53	2	1	88585	520	martiale
2,53	2	1	88585	520	portent
2,53	2	1	88585	520	reposée
2,53	2	1	88585	520	semelle
2,53	2	1	88585	520	usurpation
2,53	81	3	88585	520	être
2,47	31	2	88585	520	feu
2,39	3	1	88585	520	Harrachi
2,39	3	1	88585	520	Ouest
2,39	3	1	88585	520	braises
2,39	3	1	88585	520	choqué
2,39	3	1	88585	520	désespoir
2,39	3	1	88585	520	exil

2,39	3	1	88585	520	prie
2,39	3	1	88585	520	sursauta
2,39	3	1	88585	520	énergie
2,37	36	2	88585	520	Si
2,33	38	2	88585	520	depuis
2,33	38	2	88585	520	frère
2,3	1690	17	88585	520	la
2,29	4	1	88585	520	Conseil
2,29	4	1	88585	520	Jaafer
2,29	4	1	88585	520	Moussa
2,29	4	1	88585	520	Toutes
2,29	4	1	88585	520	douars
2,29	4	1	88585	520	facile
2,29	4	1	88585	520	mettaient
2,29	4	1	88585	520	revue
2,29	4	1	88585	520	tourné
2,26	42	2	88585	520	père
2,25	43	2	88585	520	Abdel
2,25	43	2	88585	520	Jalil
2,21	197	4	88585	520	mon
2,2	5	1	88585	520	angle
2,2	5	1	88585	520	arrêtait
2,2	5	1	88585	520	auquel
2,2	5	1	88585	520	commandait
2,2	46	2	88585	520	contre
2,2	5	1	88585	520	course
2,2	5	1	88585	520	cuisine
2,2	5	1	88585	520	hijab

L'efficiencia de l'outil informatique dans l'étude du texte littéraire

2,2	5	1	88585	520	ignorais
2,2	5	1	88585	520	individu
2,2	5	1	88585	520	légende
2,2	5	1	88585	520	salafites
2,2	5	1	88585	520	tailler
2,2	5	1	88585	520	tombé
2,14	122	3	88585	520	c'
2,13	6	1	88585	520	drame
2,13	6	1	88585	520	fonction
2,13	6	1	88585	520	jusque
2,13	6	1	88585	520	seront
2,13	6	1	88585	520	sirènes
2,11	531	7	88585	520	n'
2,07	7	1	88585	520	acteur
2,07	7	1	88585	520	frappait
2,07	7	1	88585	520	révolution
2,07	7	1	88585	520	tenture
2,07	7	1	88585	520	vestibule
2,07	7	1	88585	520	zonal
2,02	8	1	88585	520	Quelque
2,02	8	1	88585	520	flammes
2,02	8	1	88585	520	plafond
2	806	9	88585	520	pas

L'étude statistique des fréquences met clairement en évidence l'importance thématique du champ lexical de la torture. Celle-ci se manifeste non seulement à travers le nombre élevé d'occurrences du mot central, mais également par la récurrence de termes sémantiquement associés. Cette densité lexicale confirme la centralité du motif de la torture dans l'univers

narratif du roman, où il constitue un noyau symbolique majeur autour duquel s'articulent les notions de violence, de souffrance et de domination.

Ainsi, la fréquence et la cooccurrence de ces unités lexicales ne relèvent pas d'un simple effet de style : elles traduisent une insistance discursive révélatrice de la dimension thématique et idéologique que Yasmina Khadra accorde à la torture dans sa représentation du conflit et de la condition humaine.

6.2. Environnement lexical de « la guerre »

Concordance du mot pole « guerre »

<u>partie gauche</u>	<u>pivot</u>	<u>partie droite</u>
fiel) C'est la	guerre	que l'on réclame, en
en bas pas la	guerre	, la dignité. Sid Ali
véritables déclarations de	guerre	. Les appels à la mobilisation
machettes, d'arsenal de	guerre	et d'escadrons obscurs. [
regrettable chahut de gamins. La	guerre	est là. Nous sommes foutus
en contrefiche. Nous sommes en	guerre	, figure - toi. Tes
Puisqu' on est en	guerre	, est -ce que je peux
Je suis invalide de	guerre	, j'ai droit à des
tranquilles. Le spectre de la	guerre	de 54 revenait gâcher le
d'armes. L'armement de	guerre	que nous récupérons à l'issue
le problème des équipements de	guerre	. Nous avons des bases arrière
avons besoin d'auxiliaires. La	guerre	exige de mettre, derrière chaque
probablement des armes de	guerre	démontées. Cela ne le contrariait
membre actif de l'effort de	guerre	, certes dans les coulisses,
sobriquet et réclame un nom de	guerre	. Il s'éloigna vers la
lui. Il a fait la	guerre	d'Indochine, la révolution de
la révolution de 54 et la	guerre	des frontières contre les
méfier. En revanche, notre	guerre	est sainte. Ce n'est
gisaient à côté de fusils de	guerre	, de grenades artisanales, de

L'efficacité de l'outil informatique dans l'étude du texte littéraire

puisait dans le trésor de la	guerre	pour élever sa villa, à
vous change un homme, la	guerre	, pas vrai ? Cette voix
C'est mon nom de	guerre	. Je suis dans Katiba El
le serment du GIA : la	guerre	, rien que la guerre,
la guerre, rien que la	guerre	, jusqu'à l'extermination radicale
maître, nous humilient. La	guerre	fait rage, et nous tenons
les meilleurs équipements de	guerre	et étaient les seuls à disposer
murs nus, quelques trophées de	guerre	: deux sabres croisés, un
tôt ? – Nous sommes en	guerre	. Nul ne peut prévoir ce
mentaire. – Nous sommes en	guerre	. – Nous venons de la
la perdre, émir. Une	guerre	est perdue dès lors que des
Considérées comme butin de	guerre	, elles constituent le bordel de

Dans le TSC, le lexème “**guerre**” revient à trente et une occurrences, un chiffre qui témoigne de sa forte saillance textuelle et de sa valeur symbolique au sein du récit. Cette récurrence n’est pas fortuite : elle traduit la présence obsédante de la guerre dans la mémoire et la conscience des personnages, mais aussi dans la structure discursive de l’œuvre elle-même. La fréquence lexicale du terme devient ainsi un indice linguistique de son rôle thématique majeur.

Dans la première partie de notre analyse, nous nous attacherons à étudier la relation entre le mot-pôle “guerre” et son voisinage lexical à partir des concordances relevées. Cette approche permettra de situer le mot dans ses contextes d’emploi — qu’il soit précédé, suivi ou encadré par d’autres termes révélateurs — afin d’observer les associations sémantiques récurrentes. Ces associations constituent ce que l’on appelle le réseau lexical ou collocationnel, qui renseigne sur la manière dont la guerre est linguistiquement représentée dans le texte.

Dans un second temps, nous analyserons les liens sémantiques et contextuels unissant les cooccurrents du mot “guerre”. L’objectif est de repérer la formation de mini-thèmes ou de champs lexicaux secondaires gravitant autour de ce terme central. Ces réseaux périphériques — qu’ils relèvent de la violence, de la souffrance, de la perte, ou encore de la résilience — participent à la construction symbolique de la guerre comme expérience humaine et sociale dans le roman.

Ainsi, l’étude du mot “guerre” ne se limite pas à une simple mesure quantitative : elle devient un outil d’accès au sens profond du texte, révélant la manière dont Yasmina Khadra articule le lexique, la mémoire et l’émotion pour représenter l’empreinte durable du conflit sur les êtres et sur la société.

Tableau N°1. Renseignant sur l’environnement du mot *guerre*

Mot outil	Mot plein ou groupe nominal
En (4 occurrences)	Arme
La (12 occurrences)	Fusils
De (12 occurrences)	Grenades artisanales
	Deux sabres croisés
	Un casque de soldat
	Butin (de guerre)
	Trophées (de guerre)

6.2.1. La collocation des mots outils

Dans le corpus étudié, la préposition “**en**”, utilisée pour marquer un rapport temporel ou situationnel, apparaît quatre fois avant le

L'efficacité de l'outil informatique dans l'étude du texte littéraire

mot-pôle “*guerre*” — comme dans les expressions “*nous sommes en guerre*” ou “*on est en guerre*”. Ce type de construction met en évidence la dimension circonstancielle du terme, renvoyant à un état global et continu plutôt qu'à un événement ponctuel.

L'article défini “**la**” précède quant à lui le mot “*guerre*” à douze reprises. Son emploi a pour effet d'actualiser le substantif, en le dotant d'une valeur référentielle précise. Ainsi, les syntagmes qui suivent cet article — qu'il s'agisse d'un complément du nom (*la guerre de 54, la guerre des frontières*) ou d'un verbe attributif (*la guerre est à nous, la guerre est sainte*) — contribuent à spécifier le référent et à déterminer la signification contextuelle du mot “*guerre*”.

De même, la préposition “**de**”, relevée douze fois avant ce mot-pôle, assume une fonction de complémentation nominale. Elle introduit des groupes nominaux complexes tels que *armement de guerre, armes de guerre, trophées de guerre, fusils de guerre*, ou encore *trésor de guerre*. Dans ces exemples, *de* établit un lien sémantique de dépendance entre le mot “*guerre*” et le nom qu'il qualifie, participant ainsi à la formation d'un champ lexical structuré où la guerre devient à la fois catégorie de référence et principe organisateur du sens.

6.2.2. Le voisinage du mot pôle « guerre »

L'examen de l'assortiment des mots voisins, qu'ils soient proches ou éloignés du terme “*guerre*”, révèle une forte cohérence thématique. En effet, ces cooccurrents entretiennent une relation sémantique directe avec le champ de la guerre et en constituent les éléments lexicaux constitutifs.

D'une part, plusieurs de ces mots renvoient aux matériaux et instruments de guerre, tels que *arme, fusils, grenades artisanales, deux sabres croisés, casque de soldat, butin de guerre* ou encore *équipement de guerre*. Ces unités lexicales participent à la construction concrète et matérielle de la thématique, en évoquant les objets et symboles de la violence armée.

D'autre part, d'autres termes identifiés, tels que *soldat, combattant* ou *taghout*, désignent les acteurs humains impliqués dans le conflit. Ils introduisent ainsi une dimension anthropologique et sociale, centrée sur les participants et les protagonistes de la guerre.

En somme, la proximité lexicale du mot "*guerre*" avec ces cooccurrents souligne sa fonction structurante dans le corpus : il fédère autour de lui un réseau sémantique dense, articulant à la fois les objets, les acteurs et les symboles du champ guerrier.

6.3. L'environnement lexical du mot pôle « sang »

Concordance de « sang »

<u>partie gauche</u>	<u>pivot</u>	<u>partie droite</u>
gît dans une mare de	sang	, les yeux hagards et
fraction de seconde, son	sang	a fui son visage et
éclatement de chair et de	sang	, plaquant des grumeaux
revenir si vite, bon	sang	, par un temps aussi
le camp. – Bon	sang	! Je ne peux tout

L’efficience de l’outil informatique dans l’étude du texte littéraire

<u>partie gauche</u>	<u>pivot</u>	<u>partie droite</u>
se chamailler. Gardons notre	sang	- froid, et réfléchissons
contrarié que préoccupé. Son	sang	- froid raviva mon malaise
m’éclaboussant de giclées de	sang	et de fragments d’os
pour quelques gouttes de ton	sang	. Ne les écoute pas
, dans des mares de	sang	, mutilés, foudroyés,
es un cadre, bon	sang	! tu mérites de la
et des yeux injectés de	sang	; j’étais jaloux de
la veste déchirée, du	sang	sur la chemise et des
du fil, glaçait-le	sang	: « Tu vas crever
Il apprit à garder son	sang	- froid et à dresser
régulièrement pataugeant dans le	sang	d’un renégat. Parfois
à des pantins. Leur	sang	se mêla aux débris de
son immeuble lui glacèrent le	sang	. La rue était obstruée
fraction de seconde, son	sang	a fui son visage et
lui recommandait de garder son	sang	- froid et l’assurait
si une futaille pleine de	sang	frais se fracassait au sol
petit malin. – Bon	sang	, c’était le chanfre
Une puissante giclée de	sang	le gifla. Salah l’
Abou Tourab remarqua le	sang	coagulé qui maculait
Ses chaussettes étaient en	sang	. Il retroussa son pantalon

<u>partie gauche</u>	<u>pivot</u>	<u>partie droite</u>
seulement de chair et de	sang	; il était la mort
dissoutes ou transférées dans le	sang	. Le remaniement remit
passer la nuit glaçait le	sang	. On dormait recroque
matiquement réprimés dans le	sang	. La tyrannie aveugle des
toutes les rivières de	sang	et de larmes n'étanchèrent
giclaient plus haut que le	sang	. Les portes frêles des
patios, bientôt le	sang	rougit les flaques de pluie
main. J'avais du	sang	jusque dans les yeux.
leur langue frétille dans le	sang	frais de leur proie acc
que des terroristes assoiffés de	sang	. Ils ne savent rien
suffoquait. Un filament de	sang	gouttait de son cou

Le nom « sang » qui apparaît 12 fois accolé au mot-outil « de » sert à compléter le nom qui le précède et apporter plus de précision dans la description des faits de torture comme le montrent les exemples suivants : « une marre de sang ; un effroyable éclatement de chair et de sang, giclées de sang, quelques gouttes de sang, une futaille pleine de sang, un filament de sang, des yeux injectés de sang ». Quant aux mots pleins qui se trouvent au voisinage de cette forme, ils ont le rôle d'apporter plus de précision et de description à cette forme comme le montrent les exemples suivants : « du sang sur la chemise, le sang se mêla aux

L'efficacité de l'outil informatique dans l'étude du texte littéraire

débris de verres, le sang a fui son visage, sang frais, de sang le gifla, le sang coagule, le sang rougit, le sang gouttait de son cou ». Les mots voisins lointains de cette forme pole sont en parfaite adéquation avec le sens de la torture comme le montrent les mots suivants : « pataugent (dans le sang), fracasser, crever, éclatement de chair, des grumeaux de cervelle, larme, réprimé, ensanglante ».

Force est de constater, que pour renforcer le sens du mot sang, l'auteur a employé tout un réseau sémantique de ce mot qui représente la violence. Il s'agit notamment des relations de synonymie qui unissent ce mot aux autres occurrences comme le montre le tableau de formes répétées ci-dessous :

Tableau N °2. Forme répétée « sang »

Forme répétée	fréquence
Sang	32
Sanguinolent	22
Sanglant	1
Ensanglanter	1
sangloter	1

7. Conclusion

À travers l'étude du roman de Yasmina Khadra, *Ce que le jour doit à la nuit*, nous avons mis en évidence l'intérêt de la lexicométrie dans l'analyse des textes littéraires. L'exploitation du logiciel Hyperbase s'est révélée particulièrement utile, car elle a permis d'établir la liste des formes les plus fréquentes du corpus et d'identifier les unités lexicales significatives qui contribuent à la compréhension globale du texte.

L'analyse lexicométrique des mots les plus récurrents nous a ainsi conduits à dégager les thématiques dominantes de la torture et de

la violence, thèmes centraux dans un roman qui revisite les événements douloureux des années de braise (1900-2000). Cette approche quantitative, combinée à une lecture interprétative, a permis de confirmer que la fréquence lexicale peut être un indicateur pertinent des préoccupations thématiques et symboliques d'une œuvre littéraire.

Par ailleurs, la méthode statistique, initialement développée dans les sciences de la nature, a démontré sa pertinence et sa transférabilité aux sciences humaines et sociales. Elle illustre la capacité des outils numériques à renouveler les perspectives d'analyse textuelle, en offrant aux chercheurs des moyens de traitement à la fois objectifs, rapides et précis.

Enfin, cette recherche souligne la nécessité de favoriser une interdisciplinarité scientifique, en intégrant les méthodes issues de différents champs du savoir. Une telle démarche permettrait d'enrichir les études littéraires par des approches issues de la linguistique, de la statistique et de l'informatique, ouvrant ainsi la voie à une compréhension plus globale, critique et approfondie des textes. Enfin, l'évolution des outils numériques d'analyse linguistique et l'intégration de techniques d'intelligence artificielle pourraient offrir de nouvelles perspectives pour explorer la dimension sémantique et symbolique des textes littéraires contemporains. Cette étude pourrait être prolongée par une analyse comparative d'autres œuvres de Yasmina Khadra, afin de vérifier si les thématiques de la violence, de la torture et

L'efficiency de l'outil informatique dans l'étude du texte littéraire

du désespoir s'inscrivent comme des constantes stylistiques dans l'ensemble de son univers romanesque.

8. Bibliographie

Guillaume (J). 1986. L'historien du discours et la lexicométrie Histoire et mesure.

Marchand (P).1998.L'analyse du discours assistée par ordinateur. Paris : Armand Colin.

Martinez (W). 2003. ***Contribution à une méthodologie de l'analyse des cooccurrences lexicales multiples dans les corpus textuels***. Thèse de doctorat en sciences du langage. Université Paris 3 ; Sorbonne nouvelle, Paris.

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.

Mazière. (F). 2005. L'analyse du discours. Paris. PUF.

Muller. (C). 1992b. Principes et méthodes de statistique lexicale. Paris, champion.

Opler (Moris). 1945. Concept of thèmes.

Paillé, (P) et Mucchielli, (A). (2008) L'analyse qualitative en sciences humaines et sociales. Armand Colin, Paris.

Salem. (A).1986. « ***Segments répétés et analyse statistique des données textuelles*** ». Etude quantitative à propos du Père Duchesne de Hébert. Histoire & mesure.1 (2). 149-175.

Salem. (A). 1988. Approche du temps lexical. Statistique textuelle et séries chronologiques. Mots/Ordinateur/Textes/Sociétés, 17, 105,143 ;