



CLAUDIA CAGNINELLI

Corpus discursifs, hétérogénéité des données, hétérogénéité des méthodes

Une approche méthodologique des variations
intra- et intergénériques

CLAUDIA CAGNINELLI

**Corpus discursifs,
hétérogénéité des données,
hétérogénéité des méthodes**

Une approche méthodologique
des variations intra- et intergénériques

préface de Julien Longhi

UNIVERSITÀ

tab edizioni

© 2025 Gruppo editoriale Tab s.r.l.
viale Manzoni 24/c
00185 Roma
www.tabedizioni.it

Prima edizione giugno 2025
ISBN versione cartacea 979-12-5669-151-7
ISBN versione digitale open access
(licenza CC BY-NC-ND 4.0) 979-12-5669-152-4

Sommaire

- p. 7 Préface de Julien Longhi
- 11 Introduction

Cadrage théorique et fondements méthodologiques
Partie I

- 19 Chapitre 1
Corpus et recherches linguistiques. Conceptions, approches et perspectives
 - 1.1. Corpus et discours : perspectives en analyse du discours, 20
 - 1.2. Explorer le corpus : l'apport épistémologique de la textométrie, 24
 - 1.3. Corpus et discours numériques : textualités et genres de discours, 29
 - 1.4. Aborder en corpus les variations liées au genre de discours, 34

- 39 Chapitre 2
Conception du corpus. Entre réflexion et analyse
 - 2.1. Concevoir le corpus : conception du langage, cadre théorique de référence et objet d'étude, 40
 - 2.2. Élaborer le corpus : cohésion et hétérogénéité, 42

- 53 Chapitre 3
Structuration du corpus en fonction des spécificités génériques
 - 3.1. Spécificités d'un corpus d'articles journalistiques, 53
 - 3.2. Spécificités d'un corpus de tweets, 67
 - 3.3. Démarches d'exploration du corpus : une approche multidimensionnelle combinée, 88

Analyse du corpus

Partie II

- p. 95 Chapitre 4
Analyse du corpus du discours journalistique
- 4.1. Aborder le niveau global du corpus : hétérogénéité sous-générique et métaénonciatrice, 95
 - 4.2. Analyse déductive et contextualisée des niveaux locaux : hétérogénéité métaénonciatrice dans PN-INFO, 116
 - 4.3. Synthèse comparative et articulation du local au global, 133
- 149 Chapitre 5
Analyse du corpus de tweets « simples »
- 5.1. Aborder le niveau global du corpus : hétérogénéité pluritechnosémiotique et sous-technogénérique, 150
 - 5.2. Analyse déductive et contextualisée des niveaux locaux : hétérogénéité sous-technogénérique, 168
 - 5.3. Synthèse comparative et articulation du local au global, 185
- 195 Bilan final. Hétérogénéités intra- et intergénérique : rapprochements et divergences
- 201 En guise de conclusion
- 207 Références bibliographiques
- 217 Remerciements
- 219 Index des notions
- 223 Index des noms

Préface

L'analyse du discours, discipline en plein essor en France depuis la fin des années 1960, connaît en Italie un développement important, qui s'illustre par une maîtrise des cadres théoriques et des approches traditionnelles, mais aussi une grande élégance dans l'analyse de la matérialité des discours et des observables. Parmi le grand nombre de travaux remarquables publiés dans ce contexte (valorisés notamment avec l'association Do.Ri.F. qui donne à voir cette richesse), l'ouvrage de Claudia Cagninelli se distingue tout particulièrement : il s'appuie sur les principaux résultats de son travail de thèse de doctorat soutenue en 2022, et propose une articulation théorique et méthodologique adaptée de l'analyse des variations intra- et intergénériques dans la mise en discours de la fin de vie dans le débat public français.

Cet ouvrage, qui se caractérise à la fois par sa clarté, sa rigueur, et sa finesse, illustre deux tendances majeures de la recherche de Claudia Cagninelli :

- une contribution théorique forte et originale à propos d'une question scientifique centrale dans le domaine de l'analyse du discours ;
- une contribution pratique et applicative fondamentale, en proposant l'éclairage minutieux d'un débat de société qui recouvre des dimensions politiques, sociales et religieuses notamment.

La recherche de Claudia Cagninelli repose sur un postulat ambitieux et affirmé d'une relation entre les genres discursifs et la notion de positionnement discursif, qui permet de saisir les processus de stabilisation ou de négociation du sens dans les discours liés au débat public. En analysant les

spécificités et contraintes de deux genres discursifs distincts, le discours de presse et le discours numérique sur Twitter, l'auteur décrit de manière éclairante la manière dont les dynamiques génériques influencent les positionnements discursifs dominants autour d'une controverse sociale récente. La discursivité, et la généricité, sont donc placées au centre de la prise en compte du débat public.

Cet ouvrage s'inscrit dans les dynamiques actuelles de la recherche en analyse du discours et dans les enjeux de formation des analystes. Il est donc un outil essentiel, du point de vue pédagogique, car il livre une synthèse complète et exemplifiée, de l'étude des variations discursives d'ordre intra- et intergénérique au sein d'un corpus caractérisé par plusieurs niveaux d'hétérogénéité. Deux axes méthodologiques structurent l'ouvrage : la constitution du corpus et l'exploration des données par des méthodes variées. Ainsi, l'ouvrage ne se limite pas à une description théorique : il fournit les clés pour appréhender cette complexité, à travers le traitement complet et éclairant d'un vaste corpus. Celui-ci est constitué de données hétérogènes mais comparables, soumises à une analyse mixte combinant outils informatisés et analyses qualitatives. L'hétérogénéité est présente à deux niveaux : les données elles-mêmes (textualité, visées discursives, instances d'énonciation) et les méthodes d'analyse mobilisées (textométriques et qualitatives), illustrant la nécessité de dépasser certains clivages parfois construits en analyse du discours. Cette complémentarité permet d'appréhender les variations discursives intra- et intergénériques et de tester des concepts clés de l'analyse du discours dans les genres abordés. La perspective énonciative et phénoménologique est notamment renouvelée au regard des concepts de genre et de positionnement. L'ouvrage fournit en outre des pistes innovantes pour l'analyse du discours numérique, et l'articulation réelle de traitements qualitatifs et quantitatifs, à différents stades de la recherche.

Comme le lecteur l'observera, la structuration de l'ouvrage, claire et lisible, permet une exposition efficace du cadre théorique et méthodologique, en précisant la conception du corpus et la place de l'outil informatique dans l'analyse textométrique. Les spécificités du web 2.0 sont prises en compte, mais comme la démarche est mise en valeur de manière progressive, la reproductibilité de la recherche est manifeste : les chapitres détaillent la constitution du corpus, les niveaux d'hétérogénéité envisagés et les critères de comparabilité intergénérique, pour procéder ensuite à l'analyse du corpus, distinguant l'exploration des discours journalistiques et celle des discours issus de Twitter/X. Chaque sous-corpus est étudié selon une approche en trois étapes : exploration outillée globale, analyse qualitative approfondie et

mise en relation des résultats. Les variations intragénériques sont examinées avant de dresser un bilan des similitudes et divergences intergénériques. Aussi, loin de surfer sur la mode des « méthodes mixtes », très souvent insuffisamment articulées, ou des « humanités numériques », convoquées pour leur prestige, Claudia Cagninelli propose, par la preuve, la nécessité d'une combinaison processuelle des différentes approches, selon des temps de la recherche qui nécessitent une parfaite maîtrise de leurs enjeux.

Cet ouvrage apporte une contribution essentielle à l'analyse du discours et des données textuelles en mettant en lumière une méthodologie qui articule constitution et exploration du corpus à différents niveaux de textualité. Il enrichit la réflexion épistémologique et contribue au renouvellement des méthodes en analyse du discours : outil précieux pour le chercheur en analyse du discours, soucieux de se doter d'une méthodologie rigoureuse et transparente pour l'analyse de corpus, il est également bienvenu pour le chercheur en exploration de corpus qui souhaite conférer à sa recherche la dimension sociodiscursive nécessaire pour appréhender le débat public.

Dans cet ouvrage, Claudia Cagninelli relève le défi d'éclairer un sujet complexe et controversé. Sa rigueur scientifique et la solidité de son argumentation facilitent la compréhension du phénomène, tandis que la lisibilité des résultats vient valider la méthodologie.

À Cergy, le 30/03/2025

Julien Longhi

Introduction

Toute activité de recherche pose des questions relatives à la délimitation de son objet d'étude, aux conditions matérielles pour y avoir accès, au(x) cadre(s) théorique(s) sur le(s)quel(s) s'appuyer et aux méthodologies à adopter, tester voire adapter. L'objectif est d'analyser au mieux l'objet de son étude sur la base de la problématique et des questions de recherche à explorer. Il faut ainsi se situer du point de vue aussi bien théorique que méthodologique, en proposant un cadre et une démarche pertinents à l'objet et aux objectifs de la recherche.

Tiré de notre thèse de doctorat, cet ouvrage¹ se propose de présenter, en l'affinant, la réflexion méthodologique qui l'a caractérisée. Il illustre une approche discursive des variations intra- et intergénériques dans la construction d'un sujet, aussi délicat et controversé que celui de la fin de vie, dans le cadre du débat public français. Considérant le rôle du langage dans la mise en discours de l'expérience du réel, un même objet du discours peut en effet être appréhendé à partir de plusieurs perspectives. D'où l'intérêt de chercher à comprendre comment ces diverses représentations circulent à travers des contextes et des genres discursifs différents. Ces derniers participent tous aux processus de stabilisation ou bien de négociation des significations et des connotations des formes linguistiques et des procédés discursifs employés. Cela est encore plus significatif dans le cadre d'une controverse so-

1. Cet ouvrage représente une version remaniée d'une partie de notre thèse de doctorat (Cagninelli 2022). La réflexion théorico-méthodologique proposée dans la première partie a été affinée et mise au point eu égard à la nature méthodologique de l'ouvrage, qui se veut d'ordre à la fois théorique et appliqué. L'analyse illustrée dans la seconde partie a fait l'objet d'opérations de sélection et de clarification en fonction des objectifs de l'ouvrage. Des analyses supplémentaires ont également été intégrées à cette fin.

ciale, où les procédés de référenciation et de mise en discours des objets de discours principaux – c'est-à-dire les référents discursifs principaux autour desquels différentes positions s'affrontent et se polarisent – contribuent à donner corps à des positionnements qui se déploient dans la matérialité discursive.

Notre réflexion s'est ainsi concentrée sur l'influence que les spécificités et les contraintes propres à divers genres de discours exercent sur les positionnements discursifs dominants dans le débat public sur la fin de vie, abordé en relation avec un cas d'actualité qui l'a récemment relancé. La mise en place de conditions permettant l'exploration des variations de phénomènes discursifs et interdiscursifs s'avérait donc primordiale. Il était nécessaire de prendre en considération des données hétérogènes mais comparables, qui étaient aussi susceptibles d'être soumises à une analyse outillée par des méthodes informatisées. L'hétérogénéité des données entraînait en conséquence le recours à des méthodes d'analyse également diversifiées et adaptables en fonction des spécificités génériques des différents discours sélectionnés. Par genres de discours, nous entendons ici : l'article de journal et le tweet (post sur Twitter/X)², chacun contribuant à sa façon au débat.

Cet ouvrage se propose de creuser deux problématiques méthodologiques, qui sont étroitement liées : le rôle et la construction du corpus, d'une part ; la mise au point et l'application d'une démarche d'exploration combinant diverses méthodes et divers niveaux d'approche des données, de l'autre. Le choix du titre vise à mettre en évidence la centralité du corpus et la question discursive à l'origine de l'étude, qui s'intéresse aux traits distinctifs des genres de discours et à leur impact sur les procédés de construction discursive d'un sujet controversé. La problématique discursive est ainsi en relation avec deux formes d'hétérogénéité qui sont fondamentales pour l'analyse de l'objet d'étude envisagé. Dans ce travail, la significativité épistémologique de l'hétérogénéité est en effet sollicitée en ce qui concerne tant la nature des données analysées que les méthodes et les niveaux d'analyse mobilisés. Ces deux plans matérialisant les différentes formes d'hétérogénéité sont également mis en relief par le titre : (a) la nature hétérogène des données rassemblées en corpus, qui relèvent de genres de discours différents, tout comme de types de textualité, de finalités communicationnelles et d'instances d'énonciation également divers ; (b) l'hétérogénéité des méthodes d'analyse adoptées (analyses textométriques et analyses qualitatives)

2. Nous signalons que le réseau social n'avait pas encore changé de nom en X au moment de l'étude. Nous garderons ainsi les dénominations qui étaient en vigueur à cette époque-là.

et des niveaux d'approche sollicités (micro-, méso- et macrolinguistique). Il s'agit néanmoins de formes d'hétérogénéité comparables du point de vue des discours sélectionnés, mais aussi de complémentarité méthodologique des approches et des angles d'observation adoptés.

La nature hétérogène à la fois des données et des démarches d'analyse représente donc la condition nécessaire pour accéder aux variations de phénomènes discursifs à l'aune des influences exercées par les spécificités et les contraintes propres à des genres et sous-genres de discours différents. L'étude prendra en outre appui sur la notion de *positionnement discursif*, qui a été mise au point dans le cadre de notre thèse, en tant qu'exemple de phénomène discursif diffus dans la matérialité langagière et à travers divers paliers de la textualité. Nous proposons ainsi l'une des approches possibles de l'étude des variations discursives d'ordre intra- et intergénérique – à savoir des variations entre des sous-genres et des genres discursifs différents – au sein d'un corpus caractérisé par plusieurs niveaux d'hétérogénéité. Cette approche s'appuie sur une conception spécifique de l'activité langagière, du rôle et de la démarche d'exploration du corpus, qui s'ancre dans le sillage des travaux français d'analyse du discours et de textométrie et logométrie. L'objectif du présent ouvrage est de contribuer à la réflexion méthodologique dans ces domaines, en réinterrogeant les objets et les démarches d'analyse à la lumière des transformations et des défis posés par les nouvelles productions de l'environnement interconnecté du web. Cela va en outre de pair avec la nécessité de tester des notions fondamentales de l'analyse du discours comme celle de *genre de discours* à travers une approche comparative-différentielle qui se veut attentive aux déterminations discursives et interdiscursives dans les processus de construction de sens en discours.

Suivant une démarche multidimensionnelle qui combine des explorations outillées avec des analyses qualitatives, l'analyse du corpus montre l'influence de variables d'ordre macro-générique et sous-générique sur la mise en discours des objets principaux de la controverse en relation avec les traits lexico-sémantiques et pragmatico-énonciatifs des productions envisagées. Cette démarche permet également de mettre en évidence la polarisation des perspectives d'appréhension véhiculées par les constructions des objets de discours examinés, dont relèvent les positionnements dominants dans ce débat public. La démarche méthodologique proposée s'attache en outre à contourner la part de la subjectivité intrinsèque au processus interprétatif de construction du sens en discours. D'où la nécessité d'articuler des explorations complémentaires de nature textométrique, qui se fondent

sur des calculs statistiques et sur des analyses mesurables, avec un constant retour au texte à travers divers niveaux de contextualisation.

Cet ouvrage se compose de deux parties principales. La première partie développe le cadrage théorique de l'étude et les fondements méthodologiques sur lesquels se fonde la conception et la constitution du corpus. La deuxième partie présente une analyse du corpus, réalisée en fonction des spécificités génériques des textes rassemblés, des niveaux d'hétérogénéité envisagés ainsi que des diverses méthodes d'exploration adoptées.

La première partie de l'ouvrage – *Cadrage théorique et fondements méthodologiques* – situe le travail d'un point de vue théorique et méthodologique. Elle s'ouvre sur la notion de *corpus* telle qu'elle a été développée en sciences du langage, puis précise l'acception qu'elle prend dans les travaux en analyse du discours française (1.1). La mise en évidence des apports épistémologiques de l'outil informatique dans le cadre des recherches textométriques et logométriques (1.2) contribue ensuite à souligner la portée heuristique d'une exploration outillée du corpus, qui renouvelle l'approche des données, en fondant le processus d'interprétation sur des résultats mesurables et contextualisés. Compte tenu des objets et des objectifs de l'étude, les notions de *textualité* et de *genre de discours* sont abordées en relation avec les transformations engendrées par l'informatique dans la nature des discours eux-mêmes, comme c'est le cas des discours numériques du web 2.0 (1.3). La dernière section du premier chapitre expose et justifie la conception, la constitution et l'exploration du corpus faisant l'objet de la présente étude en fonction de la problématique discursive des variations liées aux genres et aux sous-genres de discours (1.4).

Le deuxième chapitre donne un exemple d'application concrète de la réflexion développée dans le chapitre initial à un cas spécifique, en se concentrant sur les enjeux théorico-méthodologiques entraînés par l'élaboration d'un corpus qui recueille des données hétérogènes du point de vue textuel et générique. La conceptualisation du corpus s'appuie sur les fondements théoriques du domaine disciplinaire de référence et sur la conception de l'activité langagière qui le caractérise (2.1). Les critères de délimitation du corpus et de sa structuration interne sont ensuite illustrés (2.2.1), de même que les différents niveaux d'hétérogénéité qu'il implique, tout en soulignant les éléments comparables au niveau intergénérique (2.2.2).

Les spécificités des deux genres de discours dont relèvent les productions rassemblées dans le corpus soumis à l'analyse sont traitées dans le troisième chapitre. La présentation de leurs caractéristiques distinctives respectives (3.1.1, 3.1.2, 3.2.1, 3.2.2), ainsi que l'identification des niveaux d'hétérogénéité

à retenir en tant que variables de l'analyse, représentent des étapes fondamentales déterminant les opérations de constitution matérielle du corpus (3.1.3, 3.2.3, 3.2.4). Des niveaux d'hétérogénéité sous-générique ont été également distingués au sein de chaque sous-corpus afin d'avoir accès aux variations intragénériques, qui s'avèrent nécessaires pour réaliser une analyse comparative intergénérique. L'illustration des démarches d'exploration du corpus dans une approche multidimensionnelle et combinée achève la première partie du volume (3.3), favorisant la transition vers la deuxième partie qui met en œuvre une application de cette démarche méthodologique sur le corpus présenté.

La deuxième partie du volume est centrée sur l'analyse du corpus. Les chapitres 4 et 5 se caractérisent par une progression similaire et portent respectivement sur l'exploration du corpus de discours journalistiques et sur celle du corpus de discours de la plateforme Twitter, renommée X en juillet 2023. Dans les deux cas, les spécificités d'ordre générique et sous-générique propres à chaque (sous-)corpus ont été prises en considération. L'analyse de chaque sous-corpus générique se déroule selon une démarche méthodologique en trois étapes, qui s'appuie sur la réflexion autour du cadre théorique développée dans la première partie du volume. Elle commence par des explorations outillées du niveau global du corpus, suivies par des analyses qualitatives d'éléments linguistiques et de phénomènes discursifs plus pointus. Les résultats des différentes perspectives et niveaux d'analyse ressortissant des deux premières phases de l'exploration du corpus sont enfin mis en relation dans une synthèse finale. Plusieurs types de variations intragénériques sont explorés dans les deux (sous-)corpus, avant de dresser un bilan final concernant les similitudes et les divergences intergénériques.

L'ouvrage retrace ainsi les étapes d'une approche méthodologique où la constitution du corpus et la démarche de son exploration représentent deux moments fondamentaux qui permettent à l'analyste d'avoir accès à l'étude de phénomènes discursifs à travers divers niveaux et paliers de la textualité, nourrissant en retour la réflexion épistémologique.

Cadrage théorique et fondements méthodologiques

partie I

Chapitre 1

Corpus et recherches linguistiques

Conceptions, approches et perspectives

Au tournant des années 2000, de nombreuses pratiques discursives se sont développées dans divers environnements du web, multipliant ainsi les données langagières au format numérique. La transformation des phénomènes discursifs à observer et de la nature des données à étudier elles-mêmes s'est accompagnée en outre de la mise au point d'outils d'analyse favorisant des explorations informatisées. Les sciences du langage et, plus largement, les sciences humaines et sociales se sont donc trouvées confrontées à des défis épistémologiques et méthodologiques inédits, posés aussi bien par les nouveaux objets d'étude que par les nouvelles méthodes disponibles pour les aborder. En particulier, les outils de traitement et d'exploration informatisés ont joué un rôle considérable, grâce aussi au développement de recherches interdisciplinaires notamment dans le cadre des humanités numériques.

Dans les études centrées sur le langage, ces transformations ont contribué à renouveler et à revitaliser un élément caractéristique de la méthodologie : le corpus. En effet, sa conceptualisation, ses formes de constitution et les démarches d'analyse auxquelles il donne lieu ont profondément changé à la suite d'« un nouveau rapport à l'empirique » (Rastier 2011 : 12) déterminé par le numérique. Cela a eu des répercussions significatives sur la nature des productions discursives, sur les dispositifs et les procédés d'appréhension des objets d'étude, ainsi que sur les approches et les perspectives méthodologiques à adopter en conséquence. Dans son acception traditionnelle, le corpus indique « un recueil large, et quelquefois exhaustif, de documents ou de données » (Beacco 2002 : 148). En revanche, dans les « linguistiques de corpus » (Habert *et al.* 1997 ; Condamines 2005), les critères linguistiques

qui déterminent la sélection des données ainsi que la représentativité de ces dernières deviennent des traits définitoires.

La mise en place de corpus plus volumineux en format numérique et l'élaboration d'outils d'analyse informatisés ont donné un nouvel élan à la recherche linguistique, favorisant l'émergence de nouveaux objets d'étude et de nouvelles pistes de réflexion, tout comme le renouvellement des méthodologies et des perspectives d'analyse. Dès leurs débuts, les travaux en linguistique de corpus et en analyse des données textuelles – aussi bien dans le domaine anglo-saxon (Sinclair 1991, 1996 ; Biber 1993, Biber *et al.* 1998 ; Tognini-Bonelli 2001 ; McEnery & Hardie 2012) que dans la recherche francophone (Lebart & Salem 1994 ; Habert *et al.* 1997 ; Williams 2006 ; Mayaffre 2010 ; Poudat & Landragin 2017 ; Lebart *et al.* 2019) – ont montré la centralité des corpus et la richesse de leurs explorations informatisées dans les études linguistiques. Comme notre travail s'inscrit dans le domaine de l'analyse du discours de tradition française (désormais ADF pour *analyse du discours française*), nous parcourons brièvement ci-dessous la conception du corpus et les approches qui lui sont propres.

1.1. Corpus et discours : perspectives en analyse du discours

1.1.1. *Le corpus : entre textes et ancrages sociaux*

La centralité du corpus dans la démarche méthodologique en tant qu'objet d'appréhension de l'activité langagière à partir de ses réalisations concrètes et attestées caractérise les travaux de la tradition française d'analyse du discours (cf. Charaudeau & Maingueneau 2002 ; Maingueneau 2014a ; Angermüller 2017 ; Mazière 2018 ; Moirand 2020). Comme l'affirme Mazière (2018 : 10), la « mise en corpus » de données authentiques est ce qui permet d'analyser le *discours*, à l'« articulation » entre la « langue structure » et la « langue empirique » (p. 15). Plus précisément,

La mise en corpus se définit donc contre le simple recueil de textes, et autonomise l'AD [Analyse du discours en France] [...]. Elle est la construction d'un dispositif d'observation propre à révéler, à faire appréhender l'objet discours qu'elle se donne pour tâche d'interpréter (Mazière 2018 : 11).

D'où le rôle du croisement interdisciplinaire caractérisant les recherches en ADF, comme en témoigne, entre autres, l'ouvrage de Guilhaumou *et al.* (1994) à l'intersection entre histoire et linguistique.

À travers l'analyse du corpus, « on cherche alors à expliquer les différences et les régularités par les extérieurs du discours », pour reprendre les mots de Moirand (2018 : 56). Parmi ces extérieurs, un rôle clé est joué par les genres de discours, qui influencent la sélection des ressources lexicales, les procédés de mise en discours et, plus généralement, les processus de construction du sens. D'ailleurs, comme le soutient Rastier (2011 : 56), l'analyse de corpus rassemblant des productions authentiques doit prendre en considération les genres auxquels elles se rattachent, dans la mesure où « le genre relie le texte à un discours », c'est-à-dire aux conditions sociodiscursives de production qui permettent d'avoir accès à « un autre statut d'intelligibilité ». Si l'on se réfère aux travaux de Maingueneau et de Charaudeau, le genre de discours relève en effet de « l'articulation des contraintes linguistiques et des contraintes situationnelles » (Maingueneau 2007 : 29) ou encore de l'« articulation entre le lieu des pratiques sociales et celui de la configuration textuelle » (Charaudeau 2001, en ligne).

Dans une conception discursive (Mazière 2018 ; Moirand 2018) et « *philologique-herméneutique* » (Rastier 2011), le corpus s'avère essentiel pour aborder la matérialité langagière à la lumière de ses déterminants sociodiscursifs ainsi que les « rapports de texte à texte, ce qui n'est possible qu'au sein d'un discours » (Rastier 2011 : 35). Le corpus discursif permet donc d'appréhender une réalité sociale par les pratiques discursives qui la caractérisent. Ainsi, suivant la proposition de Longhi (2018), nous l'envisageons comme un « "terrain" d'analyse » (p. 29), c'est-à-dire « un agencement de discours à situer dans leur écologie » (p. 20) « qui permet de saisir les objets discursifs, dont la description rend sémantiquement compte des mécanismes socio-discursifs et sémiotiques propres aux formations sociales envisagées » (p. 57). Longhi met en outre en relation la conception du corpus comme *terrain* avec celle du discours comme *champ*, c'est-à-dire « un observatoire qui permet de considérer les formes linguistiques (et sémiotiques) qui s'y déploient » (Longhi 2020a : 40). Par conséquent, le corpus se configure comme un terrain d'enquête sociodiscursive qui donne accès à différentes pratiques, formes et stratégies de discursivisation, véhiculant des appréhensions du réel diversifiées et socialement influencées, qui se manifestent en discours. En effet, comme le souligne Longhi,

le monde social ne peut se décrire d'une manière qu'on jugerait objective ou objectivée : il peut être saisi et décrit à travers les processus discursifs des sujets parlants, qui laissent les traces de l'appréhension qu'ils en ont dans la mise en mots (Longhi 2018 : 26).

Si l'étude discursive du monde social ne peut donc pas prétendre à l'objectivité, l'outillage de la démarche d'analyse vise pour autant à objectiver l'approche des données (cf. Pêcheux 1969), dont l'interprétation est en tout état de cause marquée par la subjectivité interprétante du chercheur (cf. 1.2). Le format numérique du corpus est ce qui rend possible son exploration par des outils informatiques ; d'où l'importance non seulement de la conceptualisation du corpus mais aussi du processus de sa constitution matérielle. Le corpus est ainsi au cœur de la démarche méthodologique de l'étude. Rastier (2011 : 13) voit dans la recherche linguistique appliquée sur corpus par des outils informatisés la possibilité d'« élaborer des instruments et définir une méthode expérimentale propre : elle ouvre aussi des champs d'application nouveaux et engage un mode spécifique d'articulation entre théorie et pratique ». L'exploration informatisée de corpus discursifs peut donc donner accès à une pluralité de terrains et de niveaux d'analyse, contribuant à enrichir également la réflexion théorique.

1.1.2. *Appréhender le sens en corpus : délimitation du « terrain » et approche discursive-interprétative*

D'après Poudat et Landragin (2017 : 11-13), le corpus peut être défini comme un ensemble de productions langagières attestées, recueillies en fonction de critères établis par le chercheur, dans un format électronique permettant des traitements informatiques. Suivant Pincemin (1999), la constitution du corpus devrait respecter des *conditions de signifiante*, afin de garantir sa *pertinence* et sa *cohérence*, mais aussi des conditions de *représentativité*, d'*acceptabilité* et d'*exploitabilité*. Le respect du critère d'*interprétabilité* (Pincemin 2012a) s'avère d'ailleurs essentiel pour déterminer la portée pragmatique et herméneutique du corpus. Comme l'affirme Pincemin (2012a : 15), il faut « fai[re] passer la connaissance du corpus au premier plan » pour « en faire un usage juste et approprié par la bonne connaissance de son contenu et de ses limites ».

La phase de constitution du corpus représente donc un moment fondamental de réflexion mais aussi de connaissance du corpus (cf. Pincemin 2012a), avec pour objectif de délimiter le terrain d'analyse de manière à ré-

pondre efficacement aux objectifs de l'étude. L'analyste devrait ainsi « se doter de critères de recueil et de mise en forme de ces données afin de garantir la représentativité de ce terrain vis-à-vis de ce qu'il est censé représenter » (Longhi 2018 : 30). Par conséquent, l'idée de représentativité se détache de celle incarnée par les premiers grands corpus de référence élaborés au cours de la deuxième moitié du siècle dernier (cf. par exemple Biber 1993 et Sinclair 1996). Dans les recherches discursives, les données rassemblées en corpus doivent rendre compte du terrain d'analyse délimité en fonction des objectifs de la recherche. Le corpus se veut donc représentatif d'un terrain spécifique parmi d'autres possibles.

Constitué en tenant compte de la problématique abordée et des objectifs de recherche (cf. Charaudeau 2009), le corpus n'est pas donné. Il s'agit au contraire d'un objet construit (cf. aussi Mellet 2002 ; Dalbera 2002 ; Mayaffre 2002a, 2005), qui résulte d'« une médiation consciente entre le chercheur et le fait linguistique » (Mellet 2002, en ligne : §5). Si le corpus relève de la conception théorique du chercheur et de son approche des questions de recherche (cf. Charaudeau 2009), l'empreinte subjective qui préside à sa constitution n'affecte pourtant pas sa portée heuristique. Elle participe au contraire aux processus de construction du sens en corpus. Nous rejoignons en effet la position de Mayaffre qui affirme que le corpus

est une construction arbitraire, une composition relative qui n'a de sens, de valeur et de pertinence qu'au regard des questions qu'on va lui poser, des réponses que l'on cherche, des résultats que l'on va trouver. Ce n'est pas un donné disciplinaire mais un objet heuristique. [...] C'est l'intention du chercheur qui est importante et lui donne son sens (Mayaffre 2002a, en ligne : §16-17).

Une dynamique interactive s'établit ainsi entre l'analyste, d'une part, et le corpus et les démarches de son exploration, de l'autre. Les objectifs de l'étude et le cadre théorico-méthodologique de l'analyste donnent corps au corpus, en déterminant à la fois sa structuration et les méthodes adéquates pour son exploration. Le corpus étudié grâce à une approche outillée favorise à son tour l'objectivation de l'accès aux données de la part de l'analyste, en permettant différentes visualisations et explorations de ces dernières à travers des démarches et niveaux d'analyse divers. En bref, si la construction du corpus et la méthodologie adoptée pour son exploration relèvent des choix théorico-méthodologiques de l'analyste, les deux nourrissent en retour sa réflexion et son analyse interprétative.

1.2. Explorer le corpus : l'apport épistémologique de la textométrie

Les conditions d'accès à l'objet d'étude dépendent à la fois de la mise en corpus de la matérialité langagière et de la démarche d'analyse à adopter. Dès ses débuts, l'ADF considère le recours à l'informatique comme partie intégrante d'un processus d'objectivation de la lecture des données (Pêcheux 1969). Ces dernières peuvent en effet être appréhendées depuis plusieurs perspectives différentes grâce aux traitements informatiques. L'outil d'analyse informatisée délinéarise le texte, favorisant l'identification d'éléments récurrents, d'associations et de relations constantes entre les mots, ou encore de sous-emplois de ressources qui seraient plus difficilement saisissables à la lecture linéaire traditionnelle. L'introduction d'outils d'exploration informatique assure ainsi une nouvelle force heuristique à l'analyse linguistique du corpus au service de l'activité herméneutique (Mayaffre 2002b, 2010), en décentrant le sujet interprétant (Mayaffre 2010) et en désobjectivant sa lecture du texte (Pêcheux 1969).

En France, la lexicométrie¹ (Lebart & Salem 1994) a été propulsée initialement par les travaux de Maurice Tournier et du centre de recherche de lexicologie politique de l'ENS de Saint-Cloud (Tournier 1969 ; Demonet *et al.* 1975). Tout comme ses évolutions ultérieures telles que la textométrie (Brunet 2016 ; Mayaffre *et al.* 2019 ; Pincemin 2020, 2022) et la logométrie (Mayaffre 2010), ces recherches ont permis de mettre au point une méthodologie d'analyse qui dépasse l'opposition entre quantitatif et qualitatif, en assurant « des outils pour une lecture renouvelée et une exploration systématique » (Pincemin 2020 : 26). Par ailleurs, comme le souligne Pincemin (2020), cette évolution terminologique met en évidence l'élargissement progressif des objets et des niveaux d'observation, abordant des niveaux de contextualisation de plus en plus larges.

Mayaffre (2002b, 2010) voit ainsi dans l'exploration outillée des corpus les principes de base d'une « herméneutique numérique ». Les opérations de déconstruction et reconstruction des textes en corpus contribuent de fait à délimiter la dimension subjective de l'interprétation, qui peut s'appuyer sur des mesures plus objectives, bien que son rôle reste primordial tout au long de la recherche. En effet, comme le met en évidence Mayaffre,

1. Pour une synthèse de l'évolution de la lexicométrie à la textométrie, voir Pincemin (2022). On signale également l'article de Guilhaumou (2002, en ligne) qui retrace les évolutions de la conception du corpus et des méthodes d'exploration en analyse du discours, en mettant en évidence les apports des travaux discursifs « du côté de l'histoire ».

Le chercheur ne se précipite plus sur le texte sans autre médiation que ses préjugés. Ce sont les informations [...] du texte indexé, trié, mesuré qui viennent l'interpeller dans ses compétences interprétatives. Si l'interprétation pourra être l'objet de discussions, les bases descriptives et interrogatives seront elles irréfutables : la subjectivité des chercheurs s'en trouve ainsi incontestablement repoussée pour être cantonnée au niveau de l'interprétation, alors qu'elle sévissait déjà au niveau descripto-interrogatif (Mayaffre 2002b : 159).

L'apport épistémologique de l'informatique et du numérique concerne aussi bien la représentation que l'exploration des productions langagières. Cela a marqué une étape fondamentale pour le processus d'objectivation de l'accès aux données, tout comme pour la vérification et la reproductibilité des calculs quantifiés sur lesquels se fondent les interprétations. Néanmoins, l'analyse outillée d'un corpus ne peut pas être réduite à une série de calculs statistiques et computationnels. Ces derniers constituent plutôt un point de départ – mais aussi une étape intermédiaire – d'une démarche qui combine divers angles d'approche de la matérialité langagière. Les opérations d'extraction, de segmentation et d'atomisation des données sont mises en relation avec des explorations qui assurent leur co(n)textualisation par un nécessaire retour au texte.

Ces différentes perspectives d'approche des données complexifient ainsi l'appréhension de l'objet d'étude. En s'appuyant aussi sur les travaux de Jean-Guy Meunier, Longhi (2019) soutient en effet la possibilité de « [d]épasser la contradiction entre humanités et numérique » (p. 132), et plus précisément, de « sortir des considérations clivantes entre qualité et quantité, pour redistribuer les perspectives des deux domaines selon les différents types de contextualisation » (p. 134). Il conclut ainsi que « les humanités garderont leur capacité à interpréter, et l'informatique à traiter de manière puissante, mais l'un et l'autre seront en mesure d'appréhender leur objet comme un tout cohérent » (Longhi 2019 : 134).

Comme nous le montrerons plus loin (cf. 3.3), la démarche d'analyse qui caractérise notre étude s'attache à dépasser l'opposition qualitatif-quantitatif, envisagée comme « une opposition épistémologique irréductible entre une position herméneutique interprétative et une position empirique sinon même positiviste » (Meunier 2020 : 194). Cette opposition implique d'ailleurs le contraste non seulement entre des méthodes d'analyse différentes, mais encore entre des objets d'étude de nature et de taille différente. Au contraire, l'approche multidimensionnelle combinée que nous proposons

s'attache à mettre en relation des méthodes d'analyse fondées sur des principes statistiques et computationnels avec d'autres privilégiant plutôt la dimension contextuelle des données. Cela donne lieu à une démarche diversifiée à l'articulation entre plusieurs niveaux et perspectives d'analyse différents qui s'enrichissent réciproquement, favorisant une appréhension plus complexe de l'objet.

À ces fins, la constitution du corpus devrait pour sa part favoriser l'appréhension de deux ordres de contextualisation. D'une part, la contextualisation assurée par les déterminants socio-historiques présuppose d'envisager la matérialité langagière en relation avec le cadre situationnel de production. D'autre part, il s'agit de (re)construire les formes de contextualisation des déterminants textuels, discursifs et interdiscursifs en corpus. Comme le sens n'est jamais donné, la constitution du corpus devrait permettre d'accéder à la matérialité sur laquelle s'ancrent les parcours interprétatifs de la construction intersubjective du sens et à la circulation interdiscursive des discours. D'où la proposition de Mayaffre (2002a) de construire des corpus *réflexifs* dont la structuration comporte leur contexte immédiat. Les constituants de ce type de corpus participent à la création d'« un réseau sémantique performant dans un tout (le corpus) cohérent et auto-suffisant » (Mayaffre 2002a, en ligne : §35), qui résulte des relations et des renvois qu'ils matérialisent les uns avec les autres au sein du corpus. Le corpus donne ainsi accès aux relations entre les textes et aux dimensions intertextuelle et interdiscursive des parcours de sens à travers les textes rassemblés en corpus.

Quant aux discours numériques, on peut envisager d'intégrer en corpus un niveau de contextualisation supplémentaire pour chercher à les saisir dans leur complexité. La prise en compte de leurs divers constituants technodiscursifs sous forme de métadonnées s'avère en effet nécessaire pour aborder leur sémiose de nature pluritechnosémiotique (cf. Cagninelli 2024a). Un tel corpus peut donc être qualifié de « corpus réfléchi » selon le concept proposé par Longhi (2021 : 111), dans la mesure où sa constitution permet différents « retour[s] sur lui-même ». Ce type de corpus crée ainsi les conditions pour explorer les données à travers plusieurs formes de contextualisation matérialisées par la structure du corpus.

Étant donné l'importance accordée aux formes de contextualisation ainsi qu'à une approche différentielle-contrastive interne au corpus, différents travaux mettent d'ailleurs en évidence plusieurs convergences entre les principes fondamentaux de la textométrie et ceux de la sémantique interprétative de Rastier. En premier lieu, le principe avancé par Rastier (2011) que le global détermine le local trouve son application dans le « traitement

macroscopique d'ordre quantitatif », qui assure la « décontextualisation et appréhension paradigmatique » du corpus, comme le précise Mayaffre (2010 : 35).

En deuxième lieu, la prise en compte de différentes formes de contextualisation, indispensable pour l'interprétation du sens, se réalise grâce à l'articulation des traitements quantitatifs avec des traitements qualitatifs, qui offrent une « lecture locale » c'est-à-dire une « recontextualisation et appréhension syntagmatique » (Mayaffre 2010 : 35). Rappelons à ce propos le concept de « passage » élaboré par Rastier (2011) pour indiquer le « lieu privilégié de la sémosis locale » (Rastier 2011 : 62), qui « n'a pas de bornes fixes » (p. 63) mais assure l'appréhension contextualisée du sens au niveau local.

Enfin, en tant que forme de contextualisation hiérarchiquement supérieure aux différents niveaux et unités d'analyse, le corpus présuppose une exploration endogène (Bonnafous & Tournier 1995 ; Rastier 2011 ; Mayaffre *et al.* 2019 ; Pincemin 2020), caractérisée également par des analyses contrastives internes au corpus. Ainsi, « [l]e sens étant fait de différences, le détour méthodologique par l'instrumentation permet de construire des différences » (Rastier 2011 : 51). Plusieurs méthodes textométriques se fondent en effet sur un principe d'analyse contrastif qui favorise l'émergence des similitudes et des écarts entre les sous-parties du corpus établies par l'analyste, à l'instar de l'analyse factorielle des correspondances (Benzécri 1973) ou du calcul des spécificités (Lafon 1980).

La textométrie – ou logométrie (Mayaffre 2010) – permet donc de matérialiser, au niveau méthodologique, une conception épistémologique bien précise des processus de production de sens en discours appréhendés en corpus et par le corpus. Cette démarche outillée, à l'articulation entre global et local, entre quantitatif et qualitatif, entre processus d'objectivation et activités de la subjectivité interprétante intègre l'analyse conduite par le chercheur de façon heuristique (cf. Mayaffre 2002b, 2010). L'exploration outillée du corpus alimente l'analyse interprétative de l'analyste qui évolue, se modifie et s'ouvre sur de nouvelles pistes d'exploration. L'articulation de plusieurs méthodes d'exploration à travers différents niveaux d'analyse du corpus peut en effet contribuer à « faire apparaître de *nouveaux observables* » (Rastier 2011 : 19) sur la base desquels le chercheur peut formuler de nouvelles hypothèses de recherche.

Bien qu'il existe des divergences épistémologiques et méthodologiques entre la sémantique des textes et l'ADE, il nous semble que la dimension contextuelle et contrastive de certaines analyses textométriques – qui in-

carnent des principes de la sémantique textuelle – peut se révéler fructueuse dans l'étude de phénomènes discursifs et de leurs variations à travers différents paliers textuels et niveaux d'analyse. Les analyses textométriques peuvent en effet jouer un rôle significatif dans la mise en évidence d'appareillages et de contrastes dans les dynamiques discursives attestées en corpus, en attirant l'attention de l'analyste sur des aspects à approfondir ensuite en contexte(s). En l'occurrence, nous entendons à la fois le contexte textuel, discursif et interdiscursif matérialisé par le corpus et le contexte socio-discursif de production pris en considération par l'analyste dans son activité d'interprétation.

De fait, comme le souligne aussi Pincemin (2012b, 2020), le rôle de l'analyste reste central dans toutes les phases de la recherche : non seulement dans l'étude du corpus, mais aussi dans la sélection et l'utilisation des méthodes d'exploration, sans oublier son rôle dans les phases de conception et d'élaboration du corpus. En effet, c'est à l'analyste de définir les objectifs de la recherche, de sélectionner et mettre en forme les données en corpus, d'organiser l'analyse et de la moduler en fonction des objets d'étude et des résultats obtenus au fur et à mesure de l'exploration. Cette démarche permet à l'analyste d'examiner les parcours de sens des discours en corpus et de saisir les ressources et les procédés qui en sont spécifiques à la lumière des déterminants extérieurs matérialisés par la structure du corpus². Plusieurs travaux offrent par ailleurs une riche illustration théorique et appliquée de la pluralité d'outils et de méthodes mis au point au cours des dernières décennies, comme c'est le cas, entre autres, de l'ouvrage de Poudat et Landragin (2017), du volume collectif dirigé par Née (2017) et encore de celui de Lebart *et al.* (2019).

Pour conclure, le corpus est à la fois l'objet et le support de l'étude, participant aussi de la réflexion théorico-méthodologique et alimentant les démarches d'exploration. Même s'il est construit pour répondre à des questions de recherche déterminées, son exploration outillée ouvrira également de nouvelles pistes de réflexion, stimulant l'intuition du chercheur. À cet égard, nous estimons la portée heuristique de l'analyse outillée du corpus comme l'un des aspects essentiels de la démarche qui se réalise grâce à l'interaction entre l'analyste, le corpus et les outils sélectionnés et paramétrés.

2. Au moins une partie des déterminants extradiscursifs peut en effet être intégrée au corpus sous forme de métadonnées, comme on peut le voir dans la structuration des deux (sous-)corpus présentés ici (cf. 3.1.3 et 3.2.4).

1.3. Corpus et discours numériques : textualités et genres de discours

L'informatique et le numérique sont désormais devenus partie intégrante non seulement des méthodes d'analyse linguistico-discursive, mais aussi de la nature même des matériaux langagiers à examiner. C'est le cas pour les discours du web 2.0, qui ne sont pas seulement produits par le biais d'outils et de supports technologiques, mais qui sont eux-mêmes constitués d'éléments technologiques. Il est ainsi nécessaire de réinterroger les concepts théoriques et les démarches méthodologiques à la lumière des nouveautés et des défis qu'ils entraînent. Comme le souligne Longhi dans le cadre de sa réflexion sur les humanités numériques,

Le numérique modifie le point de vue que les humanités ont de leur domaine, agissant non comme une dimension ajoutée, une « plus-value » ou un « biais » selon la manière dont on peut qualifier l'apport, mais comme un prisme qui remodèle, récursivement, le champ des humanités (Longhi 2019 : 126).

Compte tenu des objectifs de notre étude, il est nécessaire d'aborder les concepts de *textualité* et de *genre de discours* en relation avec les spécificités des discours numériques qui font l'objet de l'analyse.

1.3.1. *Discours numériques : complexification et technologisation des textualités*

La notion de texte peut se caractériser par des traits définitoires différents selon le domaine de recherche. L'étymologie latine du terme en révèle néanmoins deux caractéristiques distinctives et largement reconnues : la nature unitaire et cohérente. Dans la lignée des travaux de Adam, nous adoptons la conception du texte comme « un tout signifiant » (Adam 2020a : 17), qui résulte de plusieurs opérations à la fois de segmentation et de liage (p. 55-56) agissant à divers niveaux (p. 53). Ainsi, le texte n'existe pas en soi, mais il est perçu en tant que tel par une subjectivité interprétante. Adam (2020a ; 2020b) conçoit en outre la textualité selon une approche modulaire qui envisage un continuum de formes textuelles résultant de la combinaison de plusieurs « régimes sémio-médiologiques ». La textualité est ainsi définie en fonction de sa complexité multiforme, qui a été encore accrue par le numérique et notamment par l'environnement interconnecté du web relationnel.

Les études de Paveau (2013a, 2015, 2017) ont été parmi les premières à aborder les textualités numériques en prenant en considération les spécificités déterminées par l'environnement technologique où elles sont produites et dont elles relèvent. En 2015, Paveau établit une classification tripartite des « ordres linguistiques du numérique » différenciant *numérisé*, *numérique* et *numérique*. Elle distingue tout d'abord la textualité *numérisée*, considérée comme le résultat de l'importation d'un texte imprimé en contexte numérique, de la textualité *numérique*, dotée en revanche d'une dimension composite technolangagière. Cette dernière est en effet créée dans un environnement numérique tout en n'étant pas en ligne. Par rapport aux textualités non numériques, les textes *numérisés* et *numériques* se caractérisent donc par un format différent qui modifie notamment les parcours possibles de lecture intra- et intertextuels.

La dernière catégorie envisagée par la classification de Paveau (2015) correspond aux textes *numériques*, qui sont des productions natives de l'environnement interconnecté du web. Par la suite, Paveau (2017) les appelle « technodiscours », car ils sont co-constitués d'éléments à la fois langagiers et technologiques. Elle les définit en outre sur la base de plusieurs traits distinctifs, tels que la *composition* technodiscursive et plurisémiotique, la *délinéarisation*, l'*augmentation* énonciative, la *relationalité*, l'*investigabilité* ainsi que l'*imprévisibilité* (Paveau 2017). Finalement, ces traits augmentent le degré de complexité technodiscursive des textes *numériques* par rapport non seulement aux textualités traditionnelles hors ligne, mais aussi aux textes *numérisés* et *numériques*.

La nature complexe et technologique des formes des textualités numériques natives du web 2.0 détermine en outre une riche prolifération des parcours de lecture possibles. Les technodiscours procèdent en effet d'une activité d'« écritecture » (Paveau 2017), qui résulte d'une transformation adaptative de la pratique de lecture-interprétation aux spécificités de l'environnement numérique. Les activités interprétatives d'écritecture des technodiscours sont ainsi fortement instables. Le récepteur-interprète accède à une forme temporairement stabilisée d'un texte, produit d'un ensemble de variations antérieures qui reste pourtant ouvert à des variations ultérieures, non seulement matérielles mais aussi sémiotiques.

La complexité et les spécificités des technodiscours demandent donc de réinterroger les cadres théoriques et méthodologiques existants. À cet égard, les travaux de Paveau (2013a, 2015, 2017) engagent un changement paradigmatique, en se donnant pour objectif de dépasser l'approche dualiste traditionnelle et l'opposition entre linguistique et extralinguistique.

Selon Paveau, il faut que l'analyse de ces textualités prenne en compte la relation entre leurs composantes langagière et technologique, qui détermine leur nature et leurs fonctionnements. Nous suivons ainsi la proposition de Paveau (2015, 2017) d'intégrer la perspective logocentrique du discours au sein d'une approche « écologique » qui aborde les discours numériques à la lumière de leurs constituants pluritechnosémiotiques et en relation avec l'environnement technologique qui les co-constitue. Dans ce cadre, il nous semble néanmoins pertinent d'envisager les technodiscours dans une continuité entre textualités traditionnelles et textualités numériques, comme le suggère Adam (2020b). Les technodiscours relèveraient ainsi d'une intensité différente des caractéristiques de la textualité. Ils impliqueraient une sorte d'augmentation, d'intensification de certains traits de la textualité du fait de leur nature hybride et pluritechnosémiotique, sans pour autant se différencier nettement des formes non numériques.

1.3.2. *Des genres aux « technogenres » de discours*

Les formes technodiscursives du web relationnel entraînent une redéfinition non seulement de la textualité mais aussi de la généricité, comme le soulignent plusieurs travaux en ADF (Maingueneau, 2014a, 2014b ; Paveau, 2015, 2017). Le genre de discours constitue une question complexe qui ne fait pas pour autant l'objet d'un consensus dans les recherches linguistiques et qui peut être abordée selon des perspectives différentes. Dans le cas des études en ADF, Maingueneau précise que les critères situationnels représentent un élément définitoire des genres de discours, qu'il envisage comme

des dispositifs de communication socio-historiquement définis [...]. Par nature, les genres évoluent sans cesse avec les sociétés dont ils sont partie prenante [...]. Mais cela ne signifie pas que le genre fonctionne de manière homogène pour toute activité de parole (Maingueneau 2007 : 29-30).

Le genre de discours participe ainsi à l'appréhension de la matérialité langagière, à la lumière des déterminants discursifs qui y sont impliqués. Ces derniers dépendent à leur tour du contexte situationnel et socio-historique de production ainsi que des relations interdiscursives avec d'autres productions associées de manière explicite ou implicite, ce qui relève de

leur dimension dialogique³. Rastier aussi met en évidence l'interrelation profonde entre texte, genre et discours, et en conséquence avec la dimension sociale de l'activité langagière :

Le texte isolé n'a pas plus d'existence que le mot ou la phrase isolés : pour être produit et compris, il doit être rapporté à un genre et à un discours, puisque tout texte relève d'un genre qui le rattache à un discours et par là à un type de pratique sociale (Rastier 2011 : 23).

Les transformations des pratiques communicationnelles et des formes (techno)discursives engendrées par le numérique et par l'avènement du web social ont de fait contribué au renouvellement de certains genres de discours traditionnels, mais aussi à l'émergence de nouveaux (cf. aussi Paveau 2017). D'une part, nous constatons, comme le fait Maingueneau (2014b, en ligne : §13), l'existence d'une « hypergénéricité » des discours du web, qui implique un « lissage des différences génériques », déterminé par un affaiblissement des contraintes d'ordre générique par souci d'homogénéisation. D'autre part, nous remarquons toutefois que cette hypergénéricité ne concerne que certaines formes de textualités numériques.

Pour illustrer l'hypergénéricité des discours numériques, Maingueneau (2014b) analyse à titre d'exemple le cas des sites d'information, en mettant en évidence les différences profondes qui marquent à la fois la textualité et la généralité de ces productions. Selon l'auteur, le site d'information ne peut pas être envisagé comme « un texte » mais plutôt comme « une mosaïque de modules hétérogènes, ce qui interdit de mettre en correspondance simple un texte et une scène d'énonciation » (Maingueneau 2014b, en ligne : §17).

L'hétérogénéité des scènes d'énonciation – pour reprendre la notion de Maingueneau – confluent dans ce qui est perçu comme une « unité » relève d'une appréhension différente de la production langagière propre au contexte numérique. Elle peut donc s'éloigner de la conception traditionnelle de la textualité, développée à partir du texte écrit, tout en gardant des éléments de continuité qui peuvent également expliquer l'émergence de nouveaux déterminants génériques. Il existe en effet d'autres formes de textualités numériques, notamment celles du web participatif, qui présentent

3. À ce propos, nous nous permettons de renvoyer le lecteur au concept de *dialogisme* introduit par Bakhtine (1970, 1984) et prolongé, entre autres, par Bres *et al.* (2019), ainsi qu'au concept d'*interdiscours* proposé par Pécheux (1975) et ensuite retravaillé en ADF (voir par exemple Garric & Longhi 2013). Pour une définition synthétique des deux concepts, voir également les articles respectifs dans le *Dictionnaire d'analyse du discours* (Charaudeau & Maingueneau (dirs.) 2002).

des traits distinctifs spécifiques, déterminés par les fonctionnalités de l'environnement technologique de production. C'est le cas par exemple pour un tweet qui présente plusieurs différences par rapport à un post sur un autre réseau social. Tout en pouvant partager certains des aspects distinctifs des textualités du web (cf. *supra* 1.3.1), ces technodiscours se caractérisent par des traits (techno)génériques diversifiés et spécifiques de l'environnement numérique de production. Les formes de textualités natives des différents réseaux socio-numériques (RSN) sont déterminées par les caractéristiques des environnements et des fonctions diversifiés, différenciant les pratiques technodiscursives qui s'y produisent. Nous pensons ainsi que des technodiscours de ce type gagnent à être abordés en termes de *technogenres* de discours (Paveau 2017). De fait, puisqu'il est « marqué par ou issu de la dimension technologique du discours », le technogène de discours « implique un fonctionnement et des propriétés particuliers » (Paveau 2017 : 300).

Le composant technologique modèle les pratiques sociodiscursives des espaces numériques, constituant un nouveau type de contrainte générique qui influence la production technodiscursive. Il faut ainsi aborder ces technodiscours à l'aune des fonctions et des contraintes spécifiques de l'environnement numérique de production. L'analyse des technodiscours ne peut se dissocier d'un effort de caractérisation du technogène dont ils relèvent, afin de mettre en relief les contraintes technodiscursives qui en influencent la textualisation.

Dans le cadre de ce travail, notre appréhension de la généralité se fonde sur deux paramètres principaux : le contexte discursif de production et la matrice formelle de la configuration discursive. Le contexte discursif de production correspond à une macro-pratique sociodiscursive et à un type de discours (au sens de Maingueneau 2014a), c'est-à-dire le discours journalistique et le discours d'info-commentaire sur Twitter. La matrice formelle de la configuration discursive résulte en revanche des caractéristiques structurelles du format de base de la textualité envisagée (l'article de journal et le tweet), ainsi que du support et du dispositif communicationnel. Nous distinguons ainsi un niveau *macro-générique* (appelé également *générique* au fil du texte par souci d'économie) et un niveau *sous-générique*. Le niveau macro-générique permet de classer les productions sélectionnées en fonction du type de discours (discours journalistique/discours d'info-commentaire sur Twitter), du format textuel de base (l'article de journal/le tweet comme technodiscours), des instances énonciatives et du dispositif de communication (l'instance journalistique/pluralité d'instances s'exprimant sur la plateforme Twitter où l'on peut supposer une présence quantitativement

significative de locuteurs « ordinaires »), ainsi que des finalités communicationnelles principales identifiées sous forme de macro-actes discursifs (informer et commenter l'actualité). Au niveau macro-générique, nous opposons donc l'article de journal en ligne et le tweet, dont la comparaison permet d'identifier les variations de phénomènes discursifs d'une perspective intergénérique.

Le niveau sous-générique relève en revanche de formes de spécialisation de la pratique sociodiscursive identifiée au niveau macro-générique, pouvant concerner des variantes structurelles-formelles et/ou pragma-énonciatives, toutes dérivant du niveau macro-générique. En d'autres termes, le niveau sous-générique permet de distinguer différents modes discursifs associés au macro-genre. Ces modes matérialisent donc plusieurs configurations structurelles et pragma-énonciatives rattachées au macro-genre. Au niveau sous-générique, nous opposons ainsi les articles d'information aux articles d'opinion pour ce qui est du macro-genre de l'article journalistique. Concernant le macro-genre du tweet comme technodiscours, nous distinguons, au niveau sous-générique, les différentes pratiques technodiscursives de la plateforme Twitter (tweets « originaux », tweets-réponses, threads, retweets, etc.). Ces dernières donnent en effet lieu à des variantes formelles, caractérisées également par des finalités communicationnelles différentes. C'est à partir de ces sous-catégorisations génériques internes à l'article de journal en ligne et au tweet que nous analyserons les variations intragénériques.

1.4. Aborder en corpus les variations liées au genre de discours

Suivant les principes fondamentaux synthétisés dans les sections précédentes, la conception du corpus et sa constitution représentent une phase fondamentale de la réflexion théorique-méthodologique. Elles permettent en effet de créer les conditions nécessaires pour accéder à l'objet d'étude visé et à la diversité de ses manifestations discursives, en tenant compte également des spécificités génériques des discours envisagés. D'un point de vue méthodologique, les contraintes génériques et sous-génériques déterminent la constitution du corpus, ainsi que les démarches d'analyse adoptées. Comme le souligne Rastier (2011), la mise en corpus de données authentiques relevant de genres différents favorise l'accès aux variations influencées par les spécificités génériques de chacun. À ce propos, Rastier affirme :

L'étude des corpus en situation montre en effet que le lexique, la morpho-syntaxe, la manière dont se posent les problèmes sémantiques de l'ambiguïté et de l'implicite, tout cela varie avec les genres. [...] Pour parvenir à des traitements automatiques efficaces, il convient de spécifier les fonctionnements propres aux différents genres. [...] La linguistique de corpus doit permettre de refonder ou d'affiner la distinction intuitive et empirique des genres (Rastier 2011 : 72).

D'après les trois grands niveaux de corpus identifiés par Mayaffre (2005 : 11), notre étude s'appuie sur un corpus de type textuel « qui ne peu[t] aspirer ni à l'exhaustivité ni même à la représentativité » par rapport à la langue en tant que système. Néanmoins, il rassemble « toujours des données attestées puisqu'on ne saurait fabriquer artificiellement un texte pour prétendre en appréhender le sens » (*ibidem*). Le corpus représente ainsi ce qui permet d'avoir accès à des productions authentiques et de les analyser en tenant compte des relations avec les contextes socio(techno)discursifs de production et notamment de l'influence des contraintes génériques. Il constitue également un moyen pour tester ensuite les hypothèses et les intuitions (cf. aussi Longhi 2018) qui émergent à partir des résultats de l'analyse outillée.

La démarche d'analyse outillée textométrique dont nous donnons un exemple d'application ici prend appui sur des calculs de nature quantitative, mais elle « intègre une démarche qualitative englobante » (Pincemin 2012b : 6). L'activité d'interprétation de l'analyste reste donc essentielle, bien que sa subjectivité soit délimitée par la confrontation avec des résultats mesurables. Notre démarche vise à montrer le rôle fondamental de la structuration du corpus en fonction du cadre théorico-méthodologique adopté et des caractéristiques de l'objet d'étude envisagé. Cela permet de créer les conditions pour explorer le corpus à travers une démarche d'analyse qui s'attache à objectiver l'approche des données afin d'avoir accès aux variations intra- et intergénériques. L'analyse se réalise en effet par une multiplication des angles d'approche des données grâce au recours à plusieurs méthodes complémentaires. Des calculs statistiques, qui restituent des résultats *quantifiés*, sont articulés avec des retours constants au texte afin de les interpréter de manière contextualisée (cf. Longhi 2020b, 2021), à travers divers niveaux d'appréhension du corpus.

Dans la lignée de Longhi (2020b), nous employons ici le qualificatif *quantifié* au lieu de *quantitatif* pour insister sur le caractère mesurable des analyses textométriques effectuées et de leurs résultats. Ce choix terminologique s'éloigne de certaines acceptions propres à l'opposition *quantitatif-qualita-*

tif et des implications épistémologiques et méthodologiques qu'elle peut entraîner (cf. Longhi 2019 ; Meunier 2020), comme nous l'avons souligné plus haut (1.2). De fait, dans certains cas, le terme *quantitatif* est susceptible d'avoir une acception plus large renvoyant à des analyses non seulement statistiques et mesurables, mais aussi réalisées sur des corpus très volumineux. De même, l'opposition entre *quantitatif* et *qualitatif* peut parfois impliquer le contraste non seulement entre des méthodes d'analyse différentes mais encore entre des objets d'étude de taille différente. Cela dit, nous soulignons que la réalisation d'analyses quantifiées va toutefois de pair avec l'importance de la représentativité du corpus en fonction des objectifs de l'étude, ce qui ne peut pas pour autant se réduire à une question de volume des données. Dans un souci de clarté, nous désignons la complémentarité des méthodes adoptées dans cette étude en termes d'*inductives-quantifiées* et de *déductives-contextualisées*, en nous attachant ainsi à dépasser l'opposition entre quantitatif et qualitatif, comme le suggère Longhi (2019).

Dans cette démarche, les processus d'interprétation et d'analyse du corpus résultent d'un va-et-vient entre induction et déduction, d'une approche qui va du global au local et vice-versa, comme le soulignent aussi Mayaffre (2010) ainsi que Poudat et Landragin (2017). La tâche de caractériser le corpus est donc de première importance : il faut faire émerger ses caractéristiques distinctives au niveau global pour ensuite explorer ses diverses unités hiérarchiques et enfin revenir à des observations de portée générale. L'analyse s'ancre ainsi sur la variation en tant que voie d'accès à l'étude des dynamiques discursives attestées en corpus grâce à la matérialisation – dans la structuration de ce dernier – de plusieurs niveaux d'hétérogénéité (cf. aussi Garric 2012).

Le corpus qui fait l'objet de cette étude représente un observatoire délimité, qui se compose d'une sélection de textes relevant de genres de discours différents. Il représente le moyen pour avoir accès à l'influence des contraintes génériques sur les choix linguistiques, textuels et discursifs participant à la construction du sens au sein d'un cadre sociodiscursif déterminé. Comme le corpus rassemble également des textualités numériques d'ordre technodiscursif, nous proposons de le concevoir comme la capture d'un moment de ces textualités, qui sont susceptibles de variations et d'augmentations continues du fait de leur nature même. Le corpus saisit donc un état précis des discours sélectionnés ; il en donne un instantané en fonction des critères et des objectifs de la recherche, en ne restituant que l'une des configurations qu'ils peuvent assumer. Cela dépend en effet des spécificités des textualités numériques, qui influencent par conséquent les caractéris-

tiques des corpus les rassemblant. Il ne s'agit toutefois pas d'une limite, mais plutôt d'une richesse des textualités numériques. L'analyse de ces textualités en corpus doit donc rendre compte des phénomènes, des stratégies et des pratiques dominantes à un certain moment.

Le corpus matérialise ainsi un objet d'observation déterminé et stable en fonction des critères de sélection et de constitution, mais il sert également de support à l'observation, ce qui n'empêche pas des transformations ultérieures. Le corpus doit en effet être envisagé « comme [une] construction, potentiellement hétérogène et donc évolutive, destinée à analyser certains fonctionnements linguistiques non identifiables en son absence parce qu'ils relèvent de la variété des usages » (Garric & Longhi 2012 : 6). Compte tenu de la centralité des questions théorico-méthodologiques relevées par la création du corpus, nous consacrons les deux prochains chapitres à la description des phases de conceptualisation (chapitre 2) et de constitution (chapitre 3) des données en corpus, avant de procéder à l'analyse (partie II). L'étude du corpus présentée dans la deuxième partie de l'ouvrage constituera une application de la réflexion théorico-méthodologique illustrée dans les chapitres de cette première partie.

Chapitre 2

Conception du corpus

Entre réflexion et analyse

Les phases de conception et de constitution du corpus représentent deux moments clés de la démarche méthodologique de la recherche. La réflexion qui préside à l'élaboration du corpus vise à matérialiser, dans sa structuration même, les fondements théoriques et méthodologiques sur lesquels s'appuie l'étude. Le corpus est en effet appelé à actualiser les conditions nécessaires pour répondre aux questions de recherche posées. Ainsi, la constitution et la structuration du corpus représentent en même temps une première étape de l'analyse, dans la mesure où elles influencent la délimitation du terrain – l'une parmi d'autres possibles – ainsi que les angles d'approche et les démarches d'exploration des données envisagées. Il est donc essentiel de réfléchir en amont à ce qui peut guider aussi bien la conception du corpus que sa constitution, en motivant les choix à effectuer et en illustrant les défis méthodologiques à relever.

Précisons dès maintenant que le corpus sur lequel s'appuie ce travail est issu d'un autre plus large, qui a été conçu et analysé dans le cadre de notre thèse de doctorat (Cagninelli 2022). Il nous semble toutefois pertinent de présenter le processus de conception de sa forme originale (2.1 et 2.2), avant d'approfondir les traits caractéristiques du corpus restreint qui sera analysé par la suite (chapitre 3). Par souci de clarté, on désignera le corpus de thèse comme le « corpus de référence premier » (CR1) et le corpus restreint qui fait l'objet de l'analyse présentée dans le présent ouvrage comme le « corpus de référence second » (CR2), en nous appuyant sur la notion de « corpus de référence » proposée par Rastier (2011 : 36) pour indiquer l'« ensemble de textes sur lequel on va contraster les corpus d'étude ».

2.1. Concevoir le corpus : conception du langage, cadre théorique de référence et objet d'étude

Dans la phase d'élaboration d'un corpus, trois aspects gagnent à être pris en compte : (a) la conception du langage et le cadre théorico-méthodologique sur lesquels se fonde l'étude, (b) la conception de corpus adoptée et son rôle dans l'étude, ainsi que (c) les questions de recherche auxquelles il devrait permettre de répondre. En d'autres termes, il faut réfléchir à la façon dont les positionnements épistémologique (cadre théorique de référence) et méthodologique (conception du corpus et démarches d'exploration) de l'analyste déterminent la structuration des données en corpus afin d'atteindre les objectifs posés.

Dans notre cas, la recherche s'inscrit dans le cadre de l'analyse du discours de tradition française (cf. 1.1). D'un point de vue épistémologique, elle présuppose une approche pragma-discursive de l'énonciation, celle-ci étant conçue comme une activité intersubjective qui participe à l'appréhension (phénoménologique) et à l'élaboration (linguistico-discursive) du réel, et dont le sens se construit de manière dynamique et située (cf. Cagninelli 2022). La matérialité langagière est donc analysée en relation avec les déterminants extérieurs d'ordre situationnel (déterminants socio-historiques), discursif (déterminants génériques¹) et interdiscursif (circulation interdiscursive des productions et relations dialogiques) qui la façonnent, à travers une mise en corpus contextualisée.

Il s'ensuit que l'organisation interne du corpus est fondamentale pour donner corps à ce qui est considéré comme à la fois un dispositif et un objet d'analyse permettant d'aborder les variations intra- et intergénériques. Le genre de discours est ainsi traité en tant que paramètre influençant « les corrélations entre plans du contenu et de l'expression » (Rastier 2011 : 42) ou encore entre les « visées » et les « formes discursives » (Longhi 2011 : 12). Les spécificités et contraintes génériques représentent en effet l'un des niveaux des « variables globales » (Rastier 2011 : 35) qui participent à la textualisation des productions langagières. Comme le soutient Longhi (2012 : 1137), il existe « un lien étroit entre discursivité et textualité » et « la forme, dont découle la textualisation, est consubstantielle au dire ». Étant donné que le genre de discours se pose à « l'articulation entre des

1. La question du genre de discours est centrale dans notre étude. Soulignons néanmoins que les déterminants de type discursif peuvent concerner plusieurs niveaux des typologies et des classifications textuelles et discursives. Cf. par exemple les classifications proposées par Rastier (2011) et par Maingueneau (2014a).

lieux sociaux et des modes d'organisation textuelle » (Maingueneau 2014b, en ligne : §2) et qu'il peut être « appréhend[é] sur différents modes et selon ses lieux d'apparition » (Longhi 2011 : 46), les objectifs de notre étude demandent de prendre en compte « des manifestations discursives extrêmement diverses, non toujours explicites, voire diffuses et même floues » (Charaudeau 2009 : 51).

Afin d'avoir accès à la circulation des discours à travers différents genres et sous-genres qui alimentent le débat public, le corpus doit rendre compte de productions langagières attestées dans une pluralité de contextes discursifs différents, appartenant toutefois à un même réseau interdiscursif. L'hétérogénéité des données du corpus et l'hétérogénéité des méthodes, des perspectives et des niveaux d'analyse² adoptés constituent à la fois une nécessité et une opportunité méthodologique, comme le souligne aussi Garric (2012). L'auteure met en effet en relief la valeur heuristique de l'hétérogénéité tout en soulignant les défis que ce principe de constitution et d'analyse pose à la démarche méthodologique. Nous soutenons avec Garric qu'il s'agit d'une « difficulté méthodologique [...] pourtant requise pour l'activité interprétative » (Garric 2012 : 76).

La constitution d'un corpus pluri-générique vise à fournir les conditions nécessaires pour étudier l'impact d'un niveau macrostructurel tel que le genre de discours sur les formes, les contenus et la portée pragmatique de la mise en discours d'un sujet controversé. Parmi les diverses approches possibles, cette étude prend appui sur le concept de « positionnement discursif » (Cagninelli 2022)³. L'analyse portera ainsi sur quelques traits caractéristiques des positionnements discursifs qui s'avèrent dominants dans les discours rassemblés en corpus à partir du débat public sur la fin de vie. Par *positionnements discursifs dominants*, nous nous référons aux positionnements principaux qui résultent des constructions discursives les plus fréquemment employées sous la forme de choix de référencement et de procédés de mise en discours. D'une part, cela permet de faire ressortir les variations intra- et intergénériques des principaux aspects lexi-co-sémantiques et des traits énonciatifs-pragmatiques des productions du corpus. D'autre part, cette approche rend possible l'identification des dif-

2. Rappelons à ce propos que c'est sur un principe différentiel que se fonde également la textométrie (cf. 1.2).

3. Le concept de *positionnement discursif* a été développé dans le cadre de notre thèse de doctorat (Cagninelli 2022). Un résumé de ses caractéristiques principales ainsi qu'une mise en œuvre appliquée du concept de nature illustrative sont présentés dans Cagninelli 2024b.

férentes constructions discursives de quelques objets de discours fondamentaux du débat sélectionné et analysé plus largement dans notre thèse (Cagninelli 2022).

En s'appuyant sur une analyse multidimensionnelle des niveaux macro-méso- et microtextuel du corpus, notre étude s'attache ici à saisir un phénomène discursif et interdiscursif qui se construit par le biais d'une pluralité de ressources lexicales et de configurations discursives. L'analyse de ce phénomène est donc réalisée à travers divers paliers de textes qui relèvent de genres et de sous-genres différents, tout en s'inscrivant notamment dans l'interdiscours sur la fin de vie dont ils s'alimentent. Dans notre étude, nous envisageons l'interdiscours⁴ comme l'ensemble des discours produits sur un même sujet, influençant donc les déterminations sémantiques et les implications pragma-discursives des discours futurs qui y sont liés de manière directe et explicite ou indirecte et implicite. Cette vision s'articule également avec le principe dialogique constitutif du discours (cf. Bres *et al.* 2019), d'après lequel tout discours entre en relation avec les discours antérieurs et futurs au sein d'un réseau interdiscursif conflictuel responsable des dynamiques de construction et de (re)négociation du sens en discours, comme le mettent en évidence aussi Garric et Longhi (2013). L'objectif de notre analyse est ainsi double : (1) mettre en évidence la polarisation des perspectives d'appréhension véhiculées par les constructions des objets de discours examinés, dont relèvent les positionnements dominants dans le débat public ; (2) dégager l'influence de variables d'ordre macro-générique et sous-générique sur la mise en discours de tels objets en relation avec les traits lexico-sémantiques et pragmatico-énonciatifs des productions envisagées, à travers une perspective interdiscursive d'ordre dialogique.

2.2. Élaborer le corpus : cohésion et hétérogénéité

L'élaboration de notre corpus (aussi bien CR₁ que CR₂) procède des considérations théoriques et méthodologiques présentées dans le premier chapitre. Elle implique tout d'abord l'identification et la délimitation du « terrain » (Longhi 2018) d'analyse en fonction des objectifs de l'étude. Ce sont ensuite les spécificités discursives et génériques des productions sélection-

4. L'interdiscours constitue une notion fondamentale en ADF, introduite par Pêcheux (1975) et ensuite retravaillée par plusieurs chercheurs, comme l'illustre la synthèse proposée par Paveau (2010).

nées qui président à la structuration et au partitionnement du corpus en relation avec les paramètres d'analyse envisagés. Par conséquent, l'objet de recherche ne préexiste pas au corpus, qui joue au contraire un rôle fondamental pour sa saisie.

Dans le cadre de notre étude, le corpus comporte plusieurs niveaux d'hétérogénéité comme condition nécessaire pour explorer la variation (cf. Rastier 2011 ; Garric 2012 ; Pincemin 2012a) à la lumière des contraintes génériques. Comme le soutient Garric :

La mise en œuvre du principe de variation relève des corpus définis et des méthodes d'analyse utilisées, chacun instaurant différents réseaux intertextuels pour autant de parcours interprétatifs. Le principe de variation consiste à multiplier ces réseaux afin d'instaurer des dynamiques intertextuelles plurielles qui informent et guident l'interprétation de l'analyste (Tognini-Bonelli 2001) (Garric 2012 : 80).

Nous distinguons deux types de variation principaux : les variations intragénériques et les variations intergénériques. Nous envisageons les variations intragénériques comme les variations internes ou endogènes aux productions qui relèvent d'un même genre discursif sur la base de sous-catégorisations internes. Ces dernières s'attestent à des niveaux d'analyse hiérarchiquement inférieurs tout en dépendant de la macro-catégorie du genre de discours (cf. chapitre 3). Nous considérons en revanche les variations intergénériques comme les variations externes ou exogènes aux productions qui appartiennent à un même genre de discours. Elles ressortent de la comparaison contrastive de traits et phénomènes linguistico-discursifs entre des genres de discours différents, permettant ainsi d'avoir accès à leur dimension interdiscursive.

Afin d'aborder en corpus les variations qui peuvent caractériser – en interne et en externe – les productions issues d'un même genre de discours, il est pertinent de réfléchir avant tout aux critères susceptibles d'assurer la cohérence et la cohésion des données hétérogènes à envisager.

2.2.1. *Délimitation du terrain : cohésion et cohérence du corpus*

L'unité thématique et les liens interdiscursifs entre les données à sélectionner constituent les deux critères visant la cohérence et la cohésion de notre corpus.

2.2.1.1. Unité thématique

Le choix d'un sujet partagé par les différents discours sélectionnés représente le fil rouge qui garantit l'unité thématique du corpus, en rendant possible des analyses comparatives au-delà de son hétérogénéité interne, comme le suggèrent à la fois Garric (2012) et Pincemin (2012a).

Parmi la multiplicité de sujets envisageables, le corpus initial (CR₁) dont est issu le corpus sur lequel repose le présent travail (CR₂) rassemble une sélection de productions langagières qui ont récemment alimenté le débat français sur la fin de vie. La fin de vie constitue un sujet sensible qui soulève des questions éthiques, sociales et politiques complexes, et fait l'objet d'un débat de longue date en France. Depuis la fin du dernier millénaire, plusieurs lois ont progressivement marqué des étapes dans la reconnaissance des droits des malades et dans l'accès aux soins palliatifs, avant de dresser un cadre législatif spécifique à la fin de vie. Les lois Kouchner de 1999 et de 2002 ont été suivies de la loi Leonetti de 2005 et de la loi Claeys-Leonetti de 2016, la plus récente sur ce sujet, bien qu'un nouveau projet de loi ait relancé la discussion parlementaire en 2024. Parallèlement au débat institutionnel, la médiatisation de plusieurs cas d'actualité – on peut citer, à titre d'exemple, le cas de Vincent Humbert au début des années 2000 et, dans les deux décennies suivantes, le cas de Vincent Lambert ou encore celui d'Alain Cocq – ont contribué à relancer à plusieurs reprises le débat public sur la fin de vie, en faisant émerger des points de vue divergents au sein de l'opinion publique.

Compte tenu de la complexité du sujet débattu et des nombreux discours qui l'abordent, il est fondamental de délimiter le terrain d'analyse en construisant un corpus cohérent et significatif en fonction des objectifs de recherche. Dans le but de circonscrire notre étude aux débats les plus récents, nous avons porté notre attention, d'une part, sur les discours institutionnels de référence sur ce sujet et, d'autre part, sur le débat public lié à la médiatisation pluriannuelle d'un cas d'actualité, l'« affaire Vincent Lambert ». Cette expression a en effet surgi dans le discours de presse traitant le cas de Vincent Lambert, un infirmier victime d'un accident de la route qui l'a plongé dans un « état végétatif »⁵. Ce type d'expression est d'ailleurs courant dans ce genre discursif pour thématiser et identifier des événements liés à un fait ou à un individu

5. La controverse a porté également sur la désignation la plus appropriée pour établir l'état de santé de Vincent Lambert. Loin de vouloir prendre position à ce sujet, l'expression ici retenue relève de la décision du Conseil d'État du 24 juin 2014. <https://www.conseil-etat.fr/decisions-de-justice/jurisprudence/les-grandes-decisions-depuis-1873/ce-ass.-24-juin-2014-mme-lambert-n-s-375081-375090-et-37509> (dernière consultation le 24/03/2023).

déterminés (cf. aussi Maingueneau 2014a). En particulier, ce cas d'actualité a longtemps retenu l'attention de l'opinion publique et des médias en raison du conflit familial qui l'a accompagné. D'une part, les deux positions qui se sont opposées sur cette question ont reflété la polarisation plus générale du débat sur la fin de vie. D'autre part, ce cas s'est produit concomitamment à plusieurs initiatives en matière de fin de vie qui ont vu le jour en France pendant les onze années où il est revenu à plusieurs reprises dans l'actualité. Le cas de Vincent Lambert est ainsi devenu emblématique du débat français sur la fin de vie, prolongeant le débat institutionnel et politique dans les sphères médiatique et sociale de l'espace public.

Dans le cadre de l'analyse du discours, l'affaire Vincent Lambert peut être considéré comme un « moment discursif » (Moirand 2007 : 4) dans la mesure où « il donne lieu à une abondante production médiatique et qu'il en reste également quelques traces à plus ou moins long terme dans les discours produits ultérieurement à propos d'autres événements ». Suivant Maingueneau (2014a : 87), la constitution d'un corpus thématique – en l'occurrence, à partir d'un moment discursif spécifique – peut en outre donner accès à une « formation discursive thématique » du débat public sur la fin de vie. La délimitation du terrain d'étude et, en conséquence, du corpus ne vise pourtant pas à « étudier le référent correspondant », mais « les énoncés qui en parlent » et cela « à l'intérieur de certaines limites spatiales et temporelles qu'il convient de définir » (Maingueneau 2014a : 87). Dans notre cas, la question thématique est donc circonscrite par un événement spécifique, « l'affaire Vincent Lambert », pendant un moment déterminé de sa médiatisation longue de plusieurs années, qui est étroitement lié au cadre législatif en vigueur et notamment à la loi Claeys-Leonetti, auxquels on ne peut pas s'abstenir de faire référence.

2.2.1.2. Liens interdiscursifs

Une fois les contours thématiques esquissés, il reste à identifier des contextes et des genres discursifs en interrelation dans le débat public. Comme le précisent Badouard *et al.* (2016), la nature plurielle et fragmentée de l'espace public contemporain va de pair avec une multiplicité de pratiques sociales et discursives, chacune caractérisée par ses propres spécificités à prendre en considération. Chaque genre de discours contribue donc de manière différente au débat tout en étant étroitement lié aux autres. Les relations interdiscursives qui s'établissent entre ces discours représentent ainsi le deuxième critère à retenir pour saisir la dimension dialogique des variations.

Le corpus constituerait donc une voie d'accès – nécessairement délimitée et incomplète – au débat public en question. En l'occurrence, la condition était de veiller à l'hétérogénéité de contextes, de genres de discours et de sources énonciatives où puiser ensuite sélectivement des éléments significatifs. Par ailleurs, sans viser ni l'exhaustivité, ni la clôture, ce type de corpus reste toujours ouvert à de possibles intégrations et modifications, à la fois sur le plan des textes et sur celui du débat. Il pourrait en effet être susceptible d'intégrer de nouvelles sous-parties ainsi que de nouveaux textes en fonction d'autres questions de recherche. De fait, le corpus ne représente qu'une portion du débat en question, qui est pourtant significative compte tenu des objectifs de l'étude. Ainsi, dans notre travail, le corpus constitue

un point de départ indispensable pour une analyse du discours, les textes devant être confrontés en permanence à d'autres textes et d'autres corpus. Car les significations sociales se répartissent dans une large combinatoire des mots en rapport avec leurs différents contextes. C'est également pourquoi aucun corpus n'est jamais définitivement fermé, sa clôture ne pouvant être que le fait d'une décision provisoire à des fins opératoires (Charaudeau 2009 : 56).

En ce qui concerne la structure générale, notre corpus de thèse (CR₁) se caractérisait par une tripartition interne en fonction des contextes socio-discursifs et des genres de discours sélectionnés dans le but de donner corps à une expansion progressive du contexte interdiscursif reliant diverses « arènes » (au sens de Badouard *et al.* 2016) de l'espace public. L'objectif était d'observer la circulation interdiscursive d'une partie des discours alimentant le débat sur la fin de vie à travers un parcours de désinstitutionnalisation progressive. Ce parcours débutait par les productions officielles du discours politico-institutionnel et se terminait par le discours social, en passant par la médiation assurée par le discours journalistique.

Trois genres discursifs ont été ainsi retenus : le débat parlementaire, l'article de journal et le tweet. Le débat parlementaire relatif à la dernière loi promulguée en matière de fin de vie – la loi Claeys-Leonetti de 2016 – a été pris en considération en tant qu'exemple de discours institutionnel « au sens strict » (Oger & Ollivier-Yaniv 2003 : 127). Le discours journalistique, sous la forme d'articles de journaux en ligne, et les échanges sur Twitter par le biais des tweets ont été en revanche sélectionnés afin d'analyser des prolongements de la controverse autour de ce sujet dans la sphère médiatique et dans la sphère sociale de l'info-commentaire de l'espace public. En l'occurrence, à la suite de la délimitation du terrain, la sélection a porté sur les

productions relatives à la médiatisation du cas de Vincent Lambert – qui a contribué à relancer le débat public sur la fin de vie – au cours de sa phase conclusive (entre avril et juillet 2019), celle-ci étant peut-être la plus « polémique » au sens de Amossy (2014). Alors que le discours journalistique joue un rôle de médiateur d'événements et de discours d'intérêt public, le réseau socionumérique Twitter peut être considéré comme un espace privilégié de diffusion de l'information et de commentaire (Kwak *et al.* 2010 ; Bruns 2018), mais aussi comme un espace politique et polémique (Mercier 2015).

La tripartition interne du corpus en fonction des genres sélectionnés est illustrée par la figure 1. La flèche au bas du schéma signale l'ouverture progressive du contexte interdiscursif : du discours institutionnel aux interactions sociales du web participatif, via la médiation du discours journalistique.

La mise en discours de la fin de vie est ainsi abordée dans sa forme institutionnelle par le débat parlementaire, en tant que lieu d'origine du débat politique, qui contribue à établir la référence normative de la question, et, en conséquence, du débat public. Elle est ensuite observée dans ses formes médiatiques et sociales en relation avec le cas de Vincent Lambert, comme exemples de discours d'information, d'opinion et de commentaire d'un cas d'actualité à la lumière du cadre législatif en vigueur.

La conception du corpus originel (CR₁) aspirait en outre à donner lieu à un corpus « réflexif » (Mayaffre 2002a), dont les liens interdiscursifs étroits entre les sous-corpus assurent une forme de contextualisation réciproque. En d'autres termes, un tel corpus permet d'avoir accès à la circulation des discours entre des contextes distincts mais interreliés.

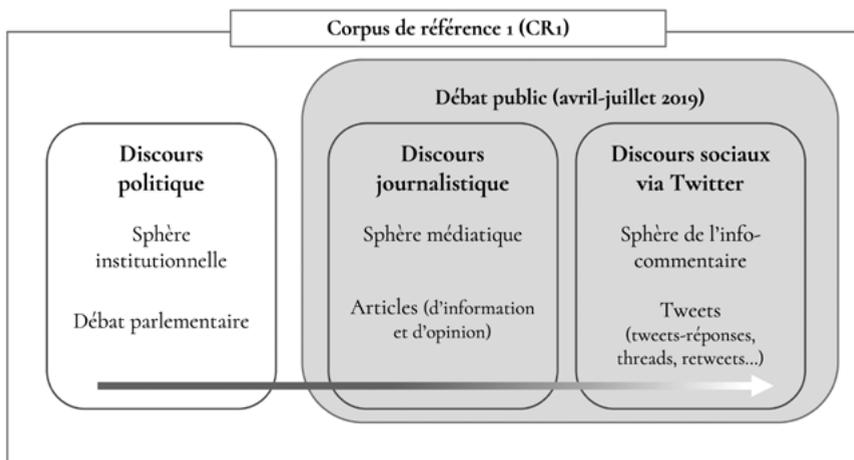


Figure 1. Relations interdiscursives entre les trois sous-parties du corpus CR₁.

2.2.2. Niveaux d'hétérogénéité et hétérogénéité comparable

Le corpus ainsi conçu comporte plusieurs niveaux et degrés d'hétérogénéité interne pour favoriser une approche comparative-contrastive qui mette en évidence les spécificités génériques et leurs influences sur les phénomènes discursifs. À ce propos, Charaudeau affirme :

Suivant l'hypothèse saussurienne qui dit que le *sens naît de la différence*, le sens de discours ne peut être saisi que dans ce qui fait contraste, du fait de la mise en regard de diverses séquences discursives. Ainsi les corpus doivent-ils être construits selon certaines variables permettant de les comparer, des variables *externes* ou *internes* (Charaudeau 2009 : 53).

L'hétérogénéité des données et l'hétérogénéité des procédés d'exploration s'imposent donc comme une « nécessité épistémologique » (Garric & Longhi 2012) pour avoir accès à la circulation interdiscursive des formes et du sens et, en conséquence, aux variations intra- et intergénériques. En effet, « introduire une hétérogénéité maîtrisée des données peut permettre d'actualiser des contextes pluriels larges qui seuls sont susceptibles de délimiter les conditions d'une construction intertextuelle » (Garric & Longhi 2012 : 4).

Le premier niveau d'hétérogénéité de CR₁ concerne les contextes discursifs sélectionnés : le contexte politique institutionnel, le contexte médiatique journalistique et le contexte que l'on pourrait définir « hybride » de la plateforme Twitter, en raison des différents types de discours sociaux qu'il peut réunir. Ce premier niveau peut, d'une part, donner accès aux trois formes impliquées dans l'élaboration du sens commun selon la progression *canon-vulgate-doxa*, développée par Sarfati (2008). De l'autre, il permet de constituer un exemple matérialisant le déploiement concentrique de l'institution et de ses discours, de même que leur ouverture vers l'extérieur, en suivant la théorisation de Oger et Ollivier-Yaniv (2003).

Ensuite, un deuxième niveau d'hétérogénéité intéresse les types et les genres de discours sélectionnés. Le débat parlementaire et l'article de journal correspondent à deux genres de discours distincts et bien définis, relevant respectivement du discours politico-institutionnel et du discours journalistique. Au contraire, les tweets peuvent être difficilement rattachés à un seul type de discours (au sens de Maingueneau 2014a), en raison de la variété des pratiques et des finalités sociodiscursives qui les caractérisent, de même que des divers statuts des twitteurs. D'où notre proposition de les

qualifier en termes d'« interactions sociales » ou encore « socio-politiques », pour renvoyer, d'une part, au caractère relationnel et interactionnel des discours sur Twitter et, de l'autre, aux enjeux sociaux ou sociopolitiques qu'ils peuvent entraîner. Nous considérons pourtant le tweet comme un technogène de discours spécifique sur la base de ses spécificités textuelles et technodiscursives (cf. 1.3).

Enfin, les formes de textualité rassemblées dans chaque sous-corpus générique représentent un dernier niveau d'hétérogénéité. D'un point de vue textuel, le sous-corpus du débat parlementaire se compose de textes écrits qui résultent de la transcription d'une production orale. Comme ils sont produits dans un environnement numérique et ensuite mis en ligne, ils relèvent d'une textualité « numérique », selon la classification proposée par Paveau (2015 ; cf. aussi 1.3.1). Il en va de même pour l'article de journal qui, dans le cadre de notre étude⁶, gagne à être envisagé en tant que textualité « numérique » (Paveau 2015), à savoir comme un texte produit en ligne et enrichi par des fonctionnalités propres à l'environnement numérique. Les tweets constituent en revanche des textualités « numérisées » (Paveau 2015), puisqu'il s'agit de productions natives du web social, qui présentent des traits composites et amplifiés par rapport aux textualités aussi bien pré-numériques que numérisées et numériques.

Au-delà des niveaux d'hétérogénéité propres aux trois genres, une divergence supplémentaire ressort de l'écart temporel entre le discours institutionnel-politique – les débats parlementaires ayant eu lieu entre mars 2015 et janvier 2016 – et le débat public de l'espace social relatif à la médiatisation du cas de Vincent Lambert. Celui-ci se déploie en effet au long de plusieurs années et s'intensifie notamment dans sa phase conclusive entre avril et juillet 2019, période à laquelle appartiennent les discours sélectionnés. Néanmoins, des relations interdiscursives fortes s'établissent entre les deux. D'une part, la loi Claeys-Leonetti, en association avec celle de 2005, constitue la référence législative principale et la plus récente, qui encadre et règle de tels cas. De l'autre, le cas de Vincent Lambert est évoqué à plusieurs reprises dans le débat institutionnel et politique. Ce décalage temporel n'affecte donc pas l'unité thématique de fond de la tripartition interne au corpus.

Pour résumer, le corpus ainsi conçu relève de plusieurs niveaux d'hétérogénéité, nécessaires pour aborder la variation à la lumière des spécifici-

6. La question de la textualité des productions rassemblées dans le corpus d'étude du discours journalistique sera abordée dans le détail plus loin (cf. 3.1.1).

tés génériques en interne et en externe. Cependant, il représente en même temps un ensemble cohérent et cohésif, grâce à son unité thématique et aux liens interdiscursifs qui relient les trois genres sélectionnés. Le corpus implique par ailleurs une hétérogénéité comparable dans la mesure où la structuration interne des différents sous-corpus repose sur des critères assimilables dans le respect des spécificités génériques de chacun. De plus, l'intégration d'informations contextuelles et d'annotations sous forme de métadonnées (cf. chapitre 3) enrichit les données langagières, favorisant des analyses plus complexes et plus fines.

Dans le cadre du présent ouvrage, nous délimitons toutefois l'illustration de quelques variations intra- et interdiscursives aux (sous-)corpus « presse en ligne » (PN) et « commentaire Twitter » (TW), qui constituent notre « corpus de référence second » (CR₂). Le corpus CR₂ implique les trois niveaux d'hétérogénéité propres au corpus CR₁, même si les deux (sous-)corpus PN et TW comportent également des niveaux d'hétérogénéité « endogène » supplémentaires en raison de leurs spécificités respectives. C'est le cas, d'une part, des différentes « unités discursives » (Moirand 2007 : 10-13) du discours journalistique⁷ (cf. 3.1.2), et, de l'autre, des diverses pratiques technodiscursives propres au dispositif de Twitter, qui se traduisent en autant de « types » de tweets (tweets « originaux », réponses, threads...) (cf. 3.2.2). Le sous-corpus TW fera l'objet d'une délimitation supplémentaire (TWS) afin de garantir la comparabilité avec le sous-corpus PN.

Cette délimitation du corpus CR₂ aux sous-corpus PN et TW s'appuie par ailleurs sur la notion de « valence générique externe » proposée par Maingueneau (2014a), qui permet de rendre compte des relations interdiscursives étroites entre les deux sous-corpus. D'après Maingueneau (2014a : 73), la valence générique externe concerne en effet un « processus d'irradiation d'un genre de discours », correspondant au « pouvoir qu'a un genre de faire parler de lui dans d'autres genres, et au-delà de nourrir les conversations de fractions plus ou moins vastes de la population ». Étant donné que le réseau Twitter est utilisé aussi pour diffuser de l'information et qu'il constitue par conséquent un moyen pour s'informer, la presse en ligne peut être considérée comme l'une des sources de l'information circulant sur Twitter. De plus, les discours sur Twitter peuvent à leur tour être repris dans le discours journalistique. Des relations étroites s'établissent

7. À ce propos, on renvoie aux travaux de Moirand (2007) et de Charaudeau (2011) qui abordent plus largement et dans le détail l'hétérogénéité interne respectivement du discours de presse et du discours médiatique.

ainsi entre les deux genres de discours retenus, malgré leurs fonctionnements discursifs distincts.

Compte tenu, d'une part, de la nécessaire hétérogénéité générique des sous-corpus et, de l'autre, des différences constitutives liées à leur structuration textuelle et discursive, ainsi qu'à leur volume, l'analyse du corpus est structurée en deux temps. Ce cheminement favorise l'étude des traits distinctifs au niveau intragénérique tout en maintenant la comparabilité au niveau intergénérique. Dans un premier temps, les deux (sous-)corpus d'étude (PN et TWS) seront ainsi analysés de manière séparée (chapitre 4 et chapitre 5) dans le but de faire ressortir les caractéristiques spécifiques à chacun en fonction des déterminations génériques respectives. Dans un second temps, la comparaison des résultats obtenus pour chaque sous-corpus permettra de mettre en évidence, dans une perspective intergénérique, des similitudes et des divergences interdiscursives quant aux ressources langagières et aux procédés exploités pour la construction discursive d'un même objet de discours.

Avant de passer à l'analyse, la présentation des spécificités textuelles, discursives et génériques de chaque corpus d'étude (discours journalistique et discours d'interaction sociale sur Twitter) s'avère indispensable, tout d'abord, pour justifier les choix qui ont présidé à la constitution du corpus, et ensuite pour contextualiser et situer les démarches de son exploration.

Chapitre 3

Structuration du corpus en fonction des spécificités génériques

La planification de la constitution du corpus et des démarches d'exploration à adopter doit tenir compte des déterminants extérieurs, y compris les contraintes génériques, qui participent activement au processus de construction du sens en discours. Il est donc essentiel de réfléchir aux traits spécifiques des genres de discours dont relèvent les textes rassemblés en corpus pour analyser leur influence sur les choix linguistiques, les procédés discursifs ainsi que les enjeux pragmatiques qu'ils sous-tendent.

3.1. Spécificités d'un corpus d'articles journalistiques

Le discours journalistique se manifeste par une grande pluralité de formes. Comme nous l'avons précisé plus haut (cf. 1.4), nous envisageons l'article de journal en ligne comme l'expression d'un macro-genre de discours en fonction de paramètres sociodiscursifs, textuels, énonciatifs et pragmatiques. La caractérisation macro-générique qui suit se concentre ainsi sur les traits généraux les plus significatifs du discours journalistique à la lumière de sa fonction de médiation d'événements saillants dans l'espace public.

3.1.1. *L'article de journal en ligne : quelle forme de textualité ?*

La question du genre du discours est étroitement liée à celle de la textualité (cf. 1.3). Il faut ainsi s'interroger sur la forme de textualité caractérisant les textes journalistiques à retenir en corpus. En fonction des objectifs

de l'étude, il nous semble que les articles de journal en ligne peuvent être abordés à partir d'au moins deux perspectives différentes, selon que l'on s'intéresse à l'article comme genre de discours journalistique, ou bien à sa textualisation comme « technogène » (Paveau 2017).

Dans le premier cas, une étude se penchant sur l'article de journal en ligne à la lumière des contraintes génériques déterminées par son appartenance au discours journalistique gagnerait à l'aborder en tant qu'exemple de textualité « numérique » selon la typologie de Paveau (2015). Ce type d'approche vise le déploiement textuel de l'article en ligne, à savoir son développement discursif, sa progression thématique, ainsi que ses liens de cohésion et cohérence. L'analyse ne prend toutefois pas en considération les éléments d'augmentation énonciative et discursive, à l'instar des textes seconds qui y sont associés tels que les commentaires, ou encore d'autres éléments qui peuvent entrecouper le texte principal de l'article. En l'occurrence, la forme textuelle retenue se rapproche de son correspondant imprimé. L'article en ligne se distingue ainsi de l'article au format papier par son support, qui ne modifie pas substantiellement le déploiement du texte primaire de l'article, dont la source énonciative principale est le journaliste. La nature numérique de l'article, qui comporte des éléments plurisémiotiques ou technodiscursifs, transforme plutôt l'activité de lecture, comme le montrent Achard-Bayle & Pešek (2022). En particulier, les fonctions technodiscursives intégrées enrichissent les parcours de lecture (et de diffusion) possibles, sans pour autant surdéterminer la construction textuelle et discursive de l'article en soi. D'ailleurs, il arrive que l'article en ligne soit structuré de manière similaire à sa version correspondante sur papier. On peut alors envisager l'article en ligne comme l'un des modules d'une unité du web, celle-ci désignée « webtexte » par Maingueneau (2022). Dans cette perspective, le webtexte relève d'un « hypergenre » (Maingueneau 2014b). Si l'article de journal en ligne a des spécificités propres à la « scénographie numérique » (Maingueneau 2014b), il présente néanmoins des affinités textuelles avec sa version papier. Par ailleurs, comme le rappellent Achard-Bayle et Pešek (2022 : 86) à partir d'autres travaux, « la dimension hypertextuelle des textes apparaît quasiment dès l'origine de l'écrit ».

Dans le second cas, le discours journalistique est en revanche étudié en tant qu'exemple de textualité « numérisée » (Paveau 2015) ou de « technodiscours » (Paveau 2017). Cette perspective implique de prendre en compte les différents espaces technodiscursifs du site de presse, qui inclut aussi bien la production première – l'article – que des productions se-

condes associées telles que les commentaires, mais aussi les partages et les commentaires dans d'autres environnements numériques. Comme l'observe Paveau, « les commentaires, textes seconds, sont coextensifs au texte premier et l'ensemble forment un seul entier discursif augmenté. La fonction commentaire modifie donc le statut du texte produit nativement en ligne » (Paveau 2017 : 44).

Bien que le commentaire constitue un prolongement de l'article, pouvant enrichir son sens, le texte premier – l'article – a un degré d'autonomie sémantique plus élevé que le texte second – le commentaire – lui étant hiérarchisé. En effet, si « les commentaires produisent un effet rétrospectif sur les unités premières et modifient donc leurs significations » (Paveau 2017 : 32), ils n'en modifient pas pour autant le fond sémantique principal. De plus, il n'est ni certain ni nécessaire que l'internaute prenne en compte les commentaires, pouvant de fait limiter sa lecture à l'article. La seule lecture de l'article garantit en effet sa compréhension-interprétation de manière complète, alors que la compréhension-interprétation des commentaires repose souvent sur la lecture préalable de l'article et, éventuellement, des commentaires précédents.

On peut ainsi conclure que la prise en compte des textes et des éléments technodiscursifs augmentant le texte premier de l'article de journal en ligne s'avère incontournable si l'on vise à sa caractérisation en tant que *technodiscours* et *technogénre* (Paveau 2017). Au contraire, si l'objectif est d'étudier l'article en ligne en tant que forme du genre journalistique, nous estimons que l'analyse gagne à être centrée sur le texte premier. Même si l'on envisage l'existence éventuelle¹ de textes seconds, aussi bien sous la forme de « commentaires de lecteurs » qu'à la suite d'actes de partage et de commentaire de l'article via les réseaux sociaux, une analyse portant sur le discours journalistique ne doit pas forcément en tenir compte.

Dans le cadre de ce travail relatif à la question du genre de discours, notre étude aborde l'article de journal en ligne en tant que textualité « numérique » au sens de Paveau (2015), c'est-à-dire comme un texte produit en ligne et enrichi par des fonctions propres à l'environnement numérique, et non pas comme un « technodiscours » (Paveau 2017). S'il intègre des fonctions technodiscursives augmentant les parcours de lecture possibles ainsi que sa diffusion, ces composantes hybrides ne surdéterminent pas la structuration textuelle et discursive de l'article en soi.

1. De plus, les commentaires peuvent ou non être présents selon les articles, de même que la fonction de commentaire peut être ou non accordée selon les journaux.

3.1.2. *Traits macro-génériques principaux de l'article de journal en ligne*

L'un des principaux objectifs généralement reconnus du discours journalistique est d'informer ses lecteurs des événements qui font l'actualité. Cependant, ce discours d'information ne peut pas se réaliser de manière neutre, contrairement à ce que ses caractéristiques formelles pourraient laisser supposer. Les recherches linguistiques – entre autres, Moirand (2007), Charaudeau (2011), Rabatel (2017) – ont en effet dévoilé les formes et les traces de la subjectivité qui traversent également le discours journalistique d'information. Cela n'est d'ailleurs qu'une conséquence de la nature intrinsèquement partielle de toute énonciation. Comme le soutient Rabatel (2017 : 18), « l'information correspond moins à des faits qu'à un discours établi à partir d'autres discours ». Une pluralité de constructions discursives partielles du réel peuvent ainsi converger dans le discours journalistique, offrant différentes perspectives d'appréhension d'une même question. L'instance d'énonciation journalistique joue néanmoins un rôle clé dans leur mise en discours. Moirand (2007 : 5) affirme en effet que « ce sont les médias qui construisent l'événement, ce sont les discours qui [...] "font" l'événement, les médias intervenant plutôt dans la mise en scène qu'ils fabriquent ». Si l'on peut s'interroger sur la typologie des faits qui sont construits en événements, l'analyse discursive s'intéresse notamment aux ressources linguistiques et aux procédés discursifs employés pour construire l'événement par le discours.

3.1.2.1. Construction de l'événement médiatique : effets d'objectivité et perspectivations partielles

La nature intrinsèquement subjective de l'énonciation implique une mise en perspective et une appréhension déterminée du réel re-construit par le discours. La mise en discours du réel se réalise nécessairement d'après un point de vue spécifique qui est susceptible d'en orienter la réception et l'interprétation. Il en va de même pour le discours journalistique qui construit l'événement à travers la mise en discours d'un « événement brut » (Charaudeau 2011 : 94), correspondant au fait tel qu'il s'est produit dans le monde réel extralinguistique. En premier lieu, la sélection des événements contribue déjà à mettre en évidence des sujets plutôt que d'autres, participant ainsi à la construction de ce que Charaudeau appelle l'« événement médiatique ». S'il existe des critères partagés pour ce processus de sélection

(cf. Charaudeau 2011), il faut également souligner que chaque journal est libre de choisir les faits dont il rend compte, selon sa ligne éditoriale, le type de lecteur ciblé et ainsi de suite. Cela relève, d'après Charaudeau (2011 : 115), de la « fonction de “filtrage” des médias, comme une tentative de maîtrise de l'événementialisation ».

Par ailleurs, un même fait peut également être mis en discours différemment selon les journaux, à travers divers choix lexicaux et divers procédés discursifs, mais aussi en soulignant certains aspects plutôt que d'autres. Les éléments linguistiques employés pour référer aux objets du discours principaux, ainsi que la façon dont ils sont mis en discours, participent en effet à la perception et à la catégorisation subjective du réel de la part de l'instance énonciative. Ces procédés linguistiques et discursifs contribuent à véhiculer le mode d'appréhension de l'objet nommé, qui implique une prise de position de l'énonciateur face à celui-ci. Nous avons ainsi proposé de désigner cette position « référentielle-idéologique » (Cagninelli 2022). Le journaliste met donc en récit l'actualité d'une perspective particulière par les choix de nomination (Moirand 2007) des référents, y compris la nomination des événements (Veniard 2013). Cette perspective transparaît dans l'objectivité apparente du discours journalistique, qui oriente en retour la perception du lecteur. Loin d'être objectif, le discours journalistique recherche plutôt un effet d'objectivité (cf. entre autres Moirand 2007 ; Charaudeau 2011 ; Rabatel 2017 ; Maingueneau 2021) par le biais de ressources et de procédés visant à effacer les traces explicites de la subjectivité énonciative.

Un même événement peut ainsi être construit en discours à travers plusieurs *perspectivations*, intrinsèquement subjectives et partielles. Comme nous l'expliquons dans Cagninelli (2024b), nous employons ce terme avec un sens spécifique pour désigner les opérations de mise en perspective des objets de discours principaux d'un sujet controversé. Par *perspectivation*, nous faisons en effet référence aux représentations orientées des objets de discours principaux, qui résultent des choix de référenciation² et des procédés principalement adoptés pour leur mise en discours, déterminant des polarisations dans les usages. Cette notion pourrait par ailleurs se rapprocher du terme anglais « *perspectivation* », employé dans la recherche linguistique à partir du concept élaboré au sein d'études en domaine littéraire (cf. Graumann & Kallmeyer 2002).

2. L'interrelation entre énonciation et référenciation a été traitée en détail dans notre thèse (Cagninelli 2022). Nous traiterons de ce choix terminologique plus loin, lors de l'exploration du corpus (cf. 4.2.1).

D'autres opérations contribuent également à véhiculer le point de vue du journaliste et/ou du journal. C'est le cas de l'agencement des unités du discours, de la sélection et des modes de représentation des propos des sources citées, ainsi que des choix lexicaux pour leur identification. Ces procédés influencent à leur tour la perception et l'interprétation de l'événement médiatique de la part du récepteur. En particulier, dans le discours journalistique, le recours à des procédés de « représentation de discours autres » (Authier-Revuz 2020) constitue un moyen « pour “laisser parler les faits” (arrangés par E1 [l'énonciateur premier enchâssant, soit le journaliste]) dans un sens qui agréé à E1, lui permettant de dire sans dire », comme le soutient Rabatel (2017 : 430). Il faut donc s'interroger sur les implications pragmatiques que ces formes et ces modes de mise en discours peuvent avoir sur le processus interprétatif :

La mise en scène des discours rapportés est lourde d'enjeux interprétatifs : en effaçant au maximum les traces du locuteur citant au profit des locuteurs cités, il se produit un effet d'objectivité, de transparence immédiate de la parole des acteurs [...] : la mise en scène donne l'illusion du vrai par l'effacement de son metteur en scène (Rabatel 2017 : 430).

Cette hétérogénéité énonciative qui est propre au discours journalistique ne concerne toutefois pas seulement la mise en scène de plusieurs discours pour construire l'événement, mais elle intéresse aussi le régime d'auctorialité de ce type de discours (Maingueneau 2021). Maingueneau (2021 : 125) observe en effet que le discours journalistique fait l'objet d'une élaboration collective, ce qui l'amène à considérer le journal comme « le *métaénonciateur* de l'ensemble de ses articles, qui sont complémentaires pour former un tout ». Il faut donc prendre en considération les relations qui s'établissent entre l'énonciateur premier de l'article (le journaliste), les éventuels énonciateurs seconds (les propos rapportés) et le métaénonciateur (le journal) afin de rendre compte des différentes perspectives et positionnements qui ressortent aux niveaux discursif et interdiscursif.

3.1.2.2. Polarisations pragma-énonciatives : discours d'information et discours d'opinion

L'énonciation journalistique, y compris celle de l'article de journal, présente des traits différents en fonction de ses expressions sous-génériques, ce qui représente un critère significatif pour sa caractérisation. D'un

point de vue pragma-énonciatif, il est en effet possible de distinguer, d'après Adam (1997 : 10-11), les diverses formes de discours journalistique en fonction de « deux positions énonciatives polaires », correspondant au *pôle distance-information* et au *pôle implication-commentaire*. L'auteur soutient en outre que, entre ces deux pôles, les productions journalistiques se distribuent selon un continuum sur la base de critères sémantique, énonciatif, pragmatique, compositionnel et stylistique. Cette bipartition majeure au sein du discours journalistique rejoint également la distinction avancée par Moirand (2007 : 12) entre les « genres qui vont du côté d'une énonciation plutôt *subjectivée* » et ceux « qui vont du côté d'une énonciation plutôt *objectivée* ».

Des paramètres spécifiques ont été proposés par Charaudeau (2011 : 175-176) pour dresser une typologie des divers sous-genres des « textes d'information médiatique », ce qui dépasse donc notre appréhension de l'article de journal en tant que macro-genre. Sa classification prend en considération, d'une part, les « *modes discursifs* du traitement de l'information » et, de l'autre, les « principaux types d'*instance énonciative* » ainsi que leur « *degré d'engagement* ». Par cette typologie, Charaudeau (2011) distingue trois modes discursifs pour construire l'événement médiatique, qui peut être « rapporté », « commenté » ou bien « provoqué ». Il identifie en outre deux types d'instances énonciatives en fonction de leur origine, interne ou externe au journal, chacune pouvant se caractériser par un degré d'engagement plus ou moins élevé et marqué.

Afin de catégoriser au niveau sous-générique les articles de journal à rassembler en corpus, nous avons mis en relation les paramètres de la typologie élaborée par Charaudeau (2011) avec la polarisation relevée par Adam (1997), qui se rapproche aussi de celle avancée par Moirand (2007). Il en résulte deux macro-catégories de sous-genres journalistiques : les articles d'information et les articles d'opinion. Les articles d'information comportent une énonciation qui se voudrait plus « objectivée » (Moirand 2007), sans pour autant que cela donne lieu à un discours objectif, mais plutôt à un discours produisant une impression d'objectivité. Dans ces articles, le journaliste est en effet supposé effacer les traces énonciatives de sa subjectivité. Il en résulte une énonciation déclarative-factuelle visant à rendre compte des faits de manière neutre, ce qui n'est d'ailleurs pas possible en raison de la nature même du langage. Puisqu'elle est intrinsèquement subjective, l'activité langagière ne peut faire abstraction de la subjectivité qui en est à l'origine. Même lorsque le discours – journalistique en l'occurrence – se voudrait objectif, il relève nécessairement

« du rapport de perception-construction que l'être humain entretient avec le réel » (Charaudeau 2011 : 34-35), c'est-à-dire de sa perception (inter)subjective de la réalité qu'il exprime par le langage (cf. Cagninelli 2022)³.

Les articles d'opinion se caractérisent au contraire par une énonciation plus « subjectivée » (Moirand 2007) qui matérialise de manière plus explicite la présence énonciative du locuteur-énonciateur principal (L1/E1)⁴ – à savoir la source énonciative de l'article de journal – ou bien celle de l'énonciateur second (e2) interpellé par le journaliste. Dans ces articles, l'énonciation du journaliste exprime ouvertement son point de vue sur l'événement, ou dans le cas, par exemple, d'un entretien, elle peut faire ressortir le point de vue de la personne interviewée à travers les propos rapportés dans le corps de l'article.

Par conséquent, on peut considérer les articles d'opinion comme des productions qui présentent des points de vue spécifiques, ceux-ci étant attribuables à un ou plusieurs individu(s) bien précis, qu'il s'agisse du journaliste (éditorialiste, chroniqueur, etc.) ou d'une personne sélectionnée par le journal pour donner son avis (tribune, entretien). Ils rendent donc compte de positionnements individuels sur l'événement, que le quotidien choisit de mettre en évidence ou en contraste. En revanche, les articles à visée informative sont censés proposer une mise en discours apparemment neutre de l'événement sélectionné, qui peut néanmoins laisser transparaître une ou plusieurs perspectives d'appréhension de l'événement. Quoique leur énonciation soit moins subjectivée, les articles d'information se font également révélateurs de la ligne éditoriale du journal face à la question.

Dans les deux cas, chaque journal choisit dans quel ordre et dans quelle mesure donner la parole aux partisans d'une position plutôt qu'à ceux d'une autre. On pourrait ainsi interpréter la distribution équilibrée des points de vue sur la question comme un effort de prise en compte neutre de l'affrontement polarisé. Au contraire, le fait de donner plus d'espace à un positionnement plutôt qu'à un autre pourrait être considéré comme une forme d'alignement du journal sur l'un des deux positionnements en opposition. D'ailleurs, comme le soulignent Garric et Longhi (2014 : 70),

3. Le concept de *positionnement discursif* que nous avons proposé prend appui sur cette appréhension de l'énonciation et en est la conséquence.

4. Les notions de locuteur et d'énonciateur sont précisées plus loin dans cette section. Soulignons toutefois dès maintenant que les notations L1/E1 et e2 sont empruntées aux travaux de Rabatel (2012, 2021).

« [l]e souci du journaliste est finalement d'inscrire à partir d'un événement advenu une orientation particulière, quitte à éluder complètement parfois la référence événementielle ».

Au-delà des stratégies adoptées par le journaliste pour effacer sa présence énonciative, il est possible de retracer sa perspective subjective, qui est susceptible d'orienter la réception-interprétation de l'événement, sur la base de plusieurs éléments, tels que la sélection des sujets à aborder, les choix de référencement, les sources énonciatives mises en scène dans son discours, la façon dont elles sont introduites et dont leurs propos sont représentés. En effet, dans le cas aussi des articles d'information, l'énonciation du locuteur-énonciateur premier (L_1/E_1), auteur de l'article, peut inclure la « représentation de discours autres » (Authier-Revuz 2020), donnant ainsi lieu à la co-présence de plusieurs actes d'énonciation et énonciateurs au sein d'un même discours (cf. 3.1.2.3).

Dans le sillage des travaux de Ducrot, de Bres et de Rabatel, nous considérons le *locuteur* comme l'instance qui « actualise l'énoncé dans sa dimension de *dire* – dimension corporelle de la voix (comme de l'écriture) » et l'*énonciateur* comme « l'instance à partir de laquelle l'énoncé est actualisé dans ses dimensions lexico-sémantique, déictique, syntaxique et modale » (Bres *et al.* 2019 : 24). Nous empruntons en outre le sigle L_1/E_1 aux travaux de Rabatel (2012, 2021) pour indiquer la coïncidence du locuteur premier avec l'énonciateur premier. Néanmoins, nous ne désignons l'énonciateur second que par *e2*, dans le but d'insister sur la rupture énonciative entre l'acte d'énonciation premier – attribuable à L_1/E_1 , en l'occurrence, l'auteur de l'article – et l'acte d'énonciation second, qu'il soit ou non effectivement attribué à un locuteur second. L'acte d'énonciation second est matériellement représenté ou évoqué – ou pour le dire avec les mots de Bres *et al.* (2019) « enchâssé » – dans l'acte d'énonciation de L_1/E_1 , mais il relève d'un autre énonciateur (*e2*).

La prise en compte de la distinction sous-générique entre articles d'information et articles d'opinion a finalement des répercussions dans les phases aussi bien de constitution que d'exploration du corpus. Il est en effet nécessaire de « traduire » ce paramètre sous forme d'annotations et de métadonnées pour pouvoir effectuer des analyses outillées comparatives. Il s'agit donc d'un niveau d'hétérogénéité supplémentaire interne à notre corpus qui favorise la caractérisation contrastive intra-générique.

3.1.2.3. Hétérogénéité énonciative et dialogisme : le cas des faits autonomiques dans le discours journalistique

La polarisation et, plus généralement, les caractéristiques énonciatives des diverses formes du discours journalistique influencent leur dimension pragmatique et notamment leur « dimension argumentative » (Amossy 2016). Cependant, au-delà des procédés adoptés par le journaliste pour construire discursivement l'événement, il faut également prendre en considération les formes d'« hétérogénéité énonciative » (Authier-Revuz 1984, 2020) qui habitent son discours. Cette hétérogénéité peut être « montrée », par l'intégration et la « représentation de discours autres » (Authier-Revuz 2020), ou elle peut procéder d'autres formes de dialogisme (Bakhtine 1970, 1984 ; Bres *et al.* 2019) qui traversent le discours du journaliste de manière moins évidente. Mais elle peut être tout aussi intrinsèque à la matérialité langagière comme dans l'épaisseur de la nomination (cf. Siblot 2001 ; Moirand 2007 ; Veniard 2013). Moirand affirme ainsi que « le texte journalistique devient alors une mosaïque de voix, constituée d'une pluralité de *files intertextuels*, et le fil horizontal du discours apparaît, dans sa matérialité même, fracturé par des marques de cette *hétérogénéité* » (Moirand 2007 : 85).

Le discours journalistique s'appuie en effet sur des discours qui lui sont extérieurs, c'est-à-dire des discours dont le journaliste n'est pas la source énonciative, aussi bien pour construire que pour commenter les événements dont il rend compte. Cette hétérogénéité de discours et d'énonciateurs dont se compose le discours journaliste change en outre selon les finalités que le journaliste et le journal poursuivent, ainsi qu'en fonction de ses sous-genres.

Le journaliste peut utiliser plusieurs formes de représentation de discours autres (Authier-Revuz 2020) tout comme d'autres procédés dialogiques (Bres *et al.* 2019) pour faire entendre, au sein de son discours, d'autres instances énonciatives, ou encore pour les évoquer. Cela peut se réaliser aussi bien par des procédés de « reformulation paraphrastique », tels que le « discours indirect » et la « modalisation en assertion seconde », que par des « modes autonomiques », par lesquels les propos représentés sont non seulement « en usage » mais aussi « mentionné[s] », selon les catégories proposées par Authier-Revuz (2020 : 311). Dans ce deuxième cas, on devrait avoir accès à la représentation aussi bien du contenu du discours autre que de ses aspects formels, sans toutefois que la reproduction fidèle en soit garantie (Authier-Revuz 2020 ; Maingueneau 2021). En prenant appui sur le travail de Authier-Revuz (2020 : 336sq), on emploie ici l'expression *modes autonomiques* au sens large pour renvoyer aux trois des cinq modes de

représentation du discours autre qui se caractérisent par le statut autonymique de la représentation de l'énoncé second au sein de l'énoncé premier : le *discours direct*, la *modalisation autonymique par emprunt*, ainsi que le *bivocal/discours indirect libre*.

Tout en étant bien consciente des implications énonciatives, pragmatiques et rhétoriques-argumentatives différentes de ces modes ainsi que de leurs diverses réalisations discursives, il nous semble intéressant de prêter attention à toute forme autonymique de représentation d'une altérité énonciative signalée par le journaliste à travers l'emploi des guillemets. D'une part, la présence des guillemets met graphiquement en évidence ces segments et, en conséquence, la rupture ou l'écart d'ordre énonciatif au sein du dire du journaliste. Bien que les guillemets ne constituent pas un élément indispensable pour l'identification de ces trois modes autonymiques, leur emploi permet de signaler clairement le statut autonymique des segments textuels qui en sont à l'intérieur, comportant à la fois la « monstration des mots » et le « blocage de la synonymie » (Authier-Revuz 2020 : 337). D'autre part, la prise en compte des segments marqués par les guillemets représente un critère empirique efficace pour procéder à leur traitement outillé dans les phases de partitionnement et d'analyse du corpus (cf. 3.1.3.2). Toute forme de représentation d'un discours autre délimitée par la présence des guillemets sera donc envisagée ici comme une *forme autonymique de représentation* d'un discours autre, voire comme un *discours autonymique*.

Le format des formes autonymiques de représentation de discours autres marquées par les guillemets est toutefois variable. Ces segments peuvent correspondre à un mot, à une suite de mots ainsi qu'à une ou plusieurs propositions. En tout cas, la représentation supposée littérale des mots d'autrui a une fonction citationnelle, qui peut également s'accompagner d'une fonction de modalisation (à l'instar du cas que Authier-Revuz (2020) appelle de « modalisation autonymique d'emprunt »). En l'occurrence, l'emploi des guillemets signale une prise de distance de la part de L₁/E₁ par rapport aux mots cités.

Ces marques typographiques permettent de signaler la présence d'une altérité énonciative au sein du discours du journaliste, favorisant un effet de réalisme. Cependant, ces formes autonymiques de représentation ont une fonction non seulement dramatique, mais aussi argumentative. Elles peuvent en effet participer activement à la perspectivisation de l'événement en relation avec le procédé de présentation de la source du discours cité (e₂), le verbe introducteur du discours citant, et la sélection des mots cités.

3.1.3. *Constituer un corpus journalistique selon un principe de variation interne et externe*

Les traits génériques distinctifs du discours journalistique retenus dans le cadre de cette étude correspondent à divers de ses niveaux d'hétérogénéité interne. Cependant, pour la création matérielle du corpus, il faut tout d'abord revenir au niveau global, c'est-à-dire au niveau textuel identifiant l'unité hiérarchiquement supérieure qui, en articulation avec le niveau macrostructurel du genre discursif, favorise l'étude de la variation intergénérique. La prise en considération du « rapport du local au global, médiatisé par plusieurs paliers d'organisation de complexité croissante » (Rastier 2011 : 39) est en effet essentielle dans une perspective à la fois intra- et intergénérique.

3.1.3.1. Hétérogénéité métaénonciatrice et critères thématique et temporel

Pour la constitution du corpus d'étude relatif au discours journalistique (dorénavant PN – presse numérique), un premier critère à établir, selon une progression hiérarchiquement descendante, concerne les sources énonciatives à retenir. Les journaux à sélectionner représentent ainsi le niveau d'hétérogénéité que nous proposons de nommer « métaénonciatrice » dans un souci de clarté. Cette dénomination entend en effet éviter toute superposition et confusion avec les notions d'« hétérogénéité énonciative » et de « méta-énonciation » (Authier-Revuz 2020), alors qu'elle prend plutôt appui sur la notion de « métaénonciateur » proposée par Maingueneau (2021). Le niveau d'hétérogénéité métaénonciatrice concerne ainsi les journaux sélectionnés.

Deux paramètres principaux ont guidé le choix des journaux en fonction des objectifs de notre étude : le degré de diffusion et la ligne éditoriale. La sélection a porté sur des quotidiens nationaux à diffusion élevée, qui incarnent également des positionnements politiques différents, afin d'avoir accès à une bonne variété de perspectives d'approche de la controverse lors de discours destinés à des publics variés. Parmi les quotidiens les plus diffusés au niveau national selon le classement de *l'Alliance pour les chiffres de la presse et des médias*⁵, quatre ont été choisis en tant que représentatifs d'autant

5. <https://www.acpm.fr/Les-chiffres/Diffusion-Presses/Presse-Payante/Presse-Quotidienne-Nationale> (dernière consultation le 02/04/2023).

de lignes éditoriales distinctes : « Le Monde », « Le Figaro », « La Croix » et « Libération ».

Pour chaque journal, les articles ont été ensuite sélectionnés en fonction de critères thématique et temporel afin de garantir l'unité et la cohésion du corpus dans une perspective aussi bien intragénérique – lors de la comparaison entre les différents journaux, ainsi qu'entre les deux sous-genres d'articles pris en considération – qu'intergénérique – lors de la comparaison entre les articles de journal et les tweets. Le critère temporel permet de délimiter la phase conclusive du moment discursif (Moirand 2007) lié à la médiatisation du cas de Vincent Lambert, du 24 avril 2019 – date à laquelle le Conseil d'État a validé la décision collégiale d'arrêt des traitements – au 12 juillet 2019 – jour suivant la mort de Vincent Lambert. Le critère thématique s'est en revanche traduit par la présence soit de l'expression exacte « Vincent Lambert » ou « Affaire Lambert » dans le titre, soit de l'expression « Vincent Lambert » dans l'introduction.

La sélection des données a été effectuée sur la base de ces trois critères – sources énonciatives, délimitation temporelle et mots-clés – en ne retenant que les articles centrés sur la médiatisation de ce cas d'actualité. En outre, d'éventuels hyperliens entrecoupant l'article ont été supprimés. Si ce choix permet de mieux rattacher le format de l'article au deuxième ordre linguistique du numérique identifié par Paveau (2015) (cf. 1.3), cela n'empêche pourtant pas de recréer – au moins partiellement – le réseau interdiscursif interne du journal matérialisé par les hyperliens. Étant donné que ceux-ci sont en relation thématique avec le paragraphe qui les précèdent (Achard-Bayle & Pešek 2022), on peut supposer que la plupart des articles identifiés par les hyperliens supprimés sont quand même inclus dans la sélection d'articles retenus pour chaque journal. Le réseau interdiscursif reste ainsi accessible grâce à la mise en corpus d'articles interreliés d'un point de vue thématique. La partition de chaque journal à l'intérieur du corpus PN fonctionne en conséquence comme un « architexte » (Rastier 2001) de l'article, et le corpus PN peut donc être envisagé à son tour comme « réflexif » (Mayaffre 2002a).

3.1.3.2. Hétérogénéité sous-générique et énonciative-pragmatique

La structuration interne du corpus d'étude PN a été réalisée grâce à l'ajout de métadonnées, c'est-à-dire d'informations supplémentaires liées à chaque texte sous la forme d'indications contextuelles (source énonciative, date, rubrique, etc.). Des annotations de nature discursive, concernant aussi bien

la catégorisation générique et sous-générique des textes que les discours autonomiques, c'est-à-dire les segments de texte entre guillemets, ont également été intégrées en tant que métadonnées. Les métadonnées s'avèrent en effet incontournables pour appréhender les productions langagières rassemblées en corpus dans leur écologie, tout comme pour mettre en œuvre des explorations outillées plus fines⁶. De plus, les métadonnées du corpus PN ont été conçues avec l'objectif de les rendre comparables dans une perspective intra- et intergénérique.

Les métadonnées principales associées à chaque article sont le genre discursif, la source énonciative (journal), la date, la rubrique et l'instance énonciative (journaliste). L'ajout d'une annotation sous-générique a permis ensuite de classer les articles sur la base de leur polarisation pragmatique-énonciative⁷, celle-ci étant susceptible d'influencer la variation intragénérique. En fonction des critères présentés plus haut (cf. 3.1.2.2), deux sous-corpus d'étude ont été créés. Le sous-corpus PN-INFO se compose d'articles d'information à visée informative-explicative, à savoir les articles qui « rapportent » l'événement (Charaudeau 2011) et qui se rapprochent en conséquence du « pôle distance-information » (Adam 1997) en raison d'une énonciation « objectivée » (Moirand 2007). Le sous-corpus PN-OP rassemble en revanche les articles visant au commentaire ou à l'expression de l'opinion. Il s'agit des articles qui « commentent » ou « provoquent » l'événement (Charaudeau 2011) et qui se rapprochent du « pôle implication-commentaire » (Adam 1997) en vertu d'une énonciation « subjectivée » (Moirand 2007).

Pour chacun des quatre journaux dont sont issus les textes retenus, la structure du corpus PN implique donc un partitionnement interne faisant la distinction entre une sous-partie qui relève de l'information-explication et l'autre qui regroupe des articles d'opinion, de commentaire ou d'analyse. Plus précisément, PN-INFO rassemble les articles d'actualité, les enquêtes, les articles d'approfondissement-explication et les reportages ; PN-OP comprend en revanche les éditoriaux, les analyses, les entretiens, les chroniques, les billets ou encore des articles d'opinion rubriqués comme « blogues ».

Après ces opérations de filtrage et de nettoyage des données, le sous-corpus PN se répartit comme l'illustre la figure 2.

6. Voir Cagninelli (2024a) pour une réflexion plus approfondie sur le rôle des métadonnées dans la structuration de corpus numériques. Une typologie de ces dernières y est également proposée.

7. Rappelons que la nature graduelle de la polarisation énonciative et pragmatique des sous-genres du discours journalistique peut être plus ou moins marquée selon les cas.

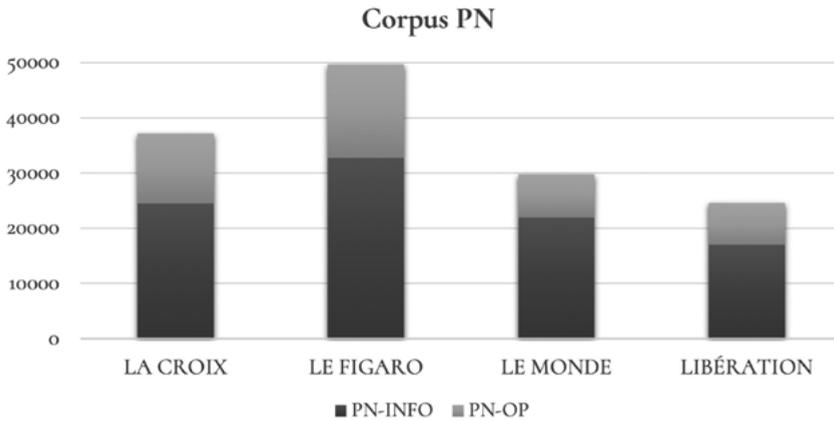


Figure 2. Structuration interne du corpus PN en fonction du nombre de mots.

Un dernier niveau d'annotation concerne la distinction entre « discours autonomiques » et « discours non autonomiques » (cf. 3.1.2.3). Comme on l'a anticipé, l'identification et l'annotation de ces segments de texte sur la base de la présence des guillemets résultent d'un filtrage informatisé, qui a permis de créer deux « sous-corpus de travail » (Rastier 2011). Ce niveau d'annotation est le seul qui ne concerne pas l'unité textuelle de l'article, mais les plans micro- et mésotextuel.

3.2. Spécificités d'un corpus de tweets

Les discours sur Twitter constituent un exemple représentatif des textualités numériques qui sont « natives » de l'environnement technologique du web, et qui peuvent être appelées aussi « technodiscours » suivant la terminologie de Paveau (2017) (cf. 1.3). Comme nous l'avons précisé dans l'introduction, le réseau socio-numérique Twitter ainsi que certaines de ses caractéristiques et fonctionnalités ont subi des modifications au fil du temps par rapport au moment où notre étude a été conduite. Si la nature même des dispositifs et des produits numériques rend possible, voire entraîne, des évolutions et des transformations, dans le cas présent divers changements ont été apportés depuis juillet 2023, quand la plateforme a été rebaptisée X. Les traits distinctifs des technodiscours et leurs fonctionnements technodiscursifs illustrés ici relèvent toutefois des caractéristiques du réseau socio-numérique lorsqu'il s'appelait encore Twitter.

3.2.1. *Le tweet : contexte de production, spécificités technodiscursives et textuelles*

D'un point de vue discursif, il faut mettre en évidence deux aspects distinctifs de l'espace de microblogage où les tweets sont produits, puisque le contexte de production influence les spécificités de la macro-pratique socio(techno)discursive qui s'y déploie. En premier lieu, le réseau Twitter est utilisé lui aussi pour diffuser et commenter l'actualité (cf. par exemple Mercier & Pignard-Cheynel 2018). En second lieu, il compte une grande variété d'utilisateurs⁸, à la différence d'autres RSN auxquels peuvent avoir recours des publics plus homogènes. La présence de discours d'information et d'actualité, relevant aussi d'instances médiatiques, ainsi que l'hétérogénéité des instances d'énonciation représentent donc deux éléments contextuels à prendre en considération aussi bien pour la constitution du corpus que pour son analyse. Les productions circulant dans cet espace peuvent ainsi contribuer à donner un aperçu des discours publics autour de sujets d'intérêt général tels que la fin de vie.

Par ailleurs, à l'époque où l'étude a été réalisée, contrairement à d'autres RSN, les discours sur Twitter étaient accessibles non seulement aux abonnés, mais aussi à tout internaute se rendant sur la plateforme⁹. De plus, les profils des twitteurs et les tweets étaient, par défaut, publics, bien que chaque utilisateur puisse choisir de rendre son compte – et en conséquence ses tweets – privé(s). Comme on pouvait le lire dans la *Politique de confidentialité* de Twitter,

Twitter est public et les Tweets peuvent immédiatement être vus et faire l'objet de recherches par qui que ce soit dans le monde entier. Nous mettons également à votre disposition des moyens non publics de communiquer sur Twitter par le biais des Tweets protégés et des Messages Privés. Vous pouvez également utiliser Twitter sous un pseudonyme si vous préférez ne pas utiliser votre propre nom¹⁰.

En tant qu'espace public, ce réseau socio-numérique invitait en outre ses utilisateurs à prêter attention aux contenus qu'ils publient, dans la mesure où

8. Pour quelques données statistiques à ce propos, voir par exemple : <https://www.blogdumoderateur.com/chiffres-twitter/> ; <https://blog.digimind.com/fr/tendances/twitter-chiffres-essentiels-france-monde-2020> (dernière consultation le 28/03/2023).

9. Des limites à la visualisation des tweets ont été introduites début juillet 2023.

10. Twitter, *Politique de confidentialité de Twitter*, https://twitter.com/fr/privacy/previous/version_16, (dernière consultation le 08/06/2023).

les tweets étaient en effet susceptibles d'être diffusés « aussi largement que possible », comme on pouvait le lire toujours dans la *Politique de confidentialité* de Twitter. La nature publique et l'intérêt public des tweets sont d'ailleurs reconnus, dans les faits, par le discours journalistique ou encore par les programmes télévisés. Ceux-ci ont de plus en plus tendance à faire référence à des tweets publiés, non seulement par des personnes célèbres mais aussi par des comptes « ordinaires », afin d'illustrer des exemples des conversations autour de divers sujets ayant retenu l'attention de l'opinion publique.

Selon une perspective discursive, un tweet peut être défini comme « un énoncé plurisémiotique complexe, [...] fortement contextualisé et non modifiable, produit nativement en ligne sur la plateforme de microblogging Twitter » (Paveau 2017 : 339). À l'instar de tout technodiscours, un tweet comporte non seulement plusieurs spécificités distinctives d'ordre technodiscursif (cf. 1.3.1), mais il présente aussi des spécificités particulières qui relèvent du dispositif technologique participant à sa création et assurant son existence même. Chaque dispositif comporte en outre des *affordances*, à savoir des « possibilités d'action » (Gibson 1979) offertes aux usagers. Les affordances favorisent l'interaction de l'utilisateur avec l'environnement numérique, puisqu'elles lui permettent d'établir des relations avec d'autres contenus et donc avec d'autres utilisateurs, tout comme d'élaborer parallèlement son propre technodiscours. C'est le cas par exemple des « technosignes » (Paveau 2015, 2017) de la plateforme, c'est-à-dire chaque « élément iconique non verbal cliquable doté d'une fonction sémantico-discursive et relationnelle, du type bouton de partage ou like » (Paveau 2015, en ligne, note 4 relative au §17). Ce type d'affordance permet ainsi d'accomplir « une opération discursive explicite [...] ou implicite » (Paveau 2017 : 319). Un autre aspect caractéristique de ces technodiscours concerne la présence possible d'éléments technolangagiers, à savoir des segments cliquables de nature langagière. Ces éléments assurent la relationalité du technodiscours ainsi que sa délinéarisation, composition et augmentation, à l'instar des « technomots » (Paveau 2013a, 2017) tels que les hashtags¹¹ et les mentions (cf. 3.2.2.1).

Le tweet constitue un format de base des technodiscours qui peuvent être produits au sein de l'environnement numérique de Twitter. En effet, il est possible de distinguer différents types de tweets en fonction des pra-

11. Dans la littérature francophone, certaines études préfèrent le terme *mot-dièse* pour désigner le hashtag. Nous privilégions en revanche la forme anglaise *hashtag* en raison de son usage plus fréquent, comme le remarque également Paveau (2017).

tiques technodiscursives qu'ils réalisent (cf. 3.2.2.2), bien qu'ils partagent tous un certain nombre de traits structurels constants. D'une part, figurent des éléments de contextualisation technodiscursive (nom et description du profil du compte ; date, heure, dispositif de création ; éléments statistiques relatifs aux taux de visualisation, diffusion et engagement du tweet, etc.). De l'autre, tout tweet se constitue de ce que l'on pourrait nommer une « fenêtre technodiscursive », réservée à la création du message qui ne peut pas dépasser 280 caractères, exceptés les liens URL. Cette fenêtre peut se composer d'éléments verbaux (ressources langagières) et technolangagiers (technomots et liens URL), tout comme de ressources iconiques, telles que les émojis. La fenêtre technodiscursive peut également être prolongée par ce que l'on peut considérer comme un espace médiatique, qui comporte des ressources visuelles (images, GIFs...) ou plurisémiotiques (vidéos). Le message du tweet peut ainsi ne correspondre qu'à la fenêtre technodiscursive ou, au contraire, intégrer d'autres objets plurisémiotiques (images, vidéos, GIFs, etc.), affichés dans l'espace au-dessous. Il va de soi qu'un tweet peut également n'inclure que des éléments sémiotiques non langagiers, sans aucun message technodiscursif.

Pour résumer et par souci de clarté, nous désignons par « partie textuelle » ou par « texte » la fenêtre technodiscursive du tweet comportant le message principalement linguistique limité à 280 caractères, y compris les technomots, tels que les hashtags et les mentions, ainsi que les émojis et émoticônes. La partie textuelle ne comprend pas en revanche les contenus multimédias complémentaires au texte, tels que les liens URL, les images, les GIFs et les vidéos.

3.2.2. *Traits macro-génériques principaux du tweet*

Après avoir traité des formes et des fonctions des hashtags et des mentions, nous aborderons les deux principaux critères distinctifs que nous avons établis dans le but de proposer une caractérisation technodiscursive des différents types de tweets : leur nature pluritechnosémiotique et les diverses pratiques sociotechnodiscursives qu'ils matérialisent.

3.2.2.1. Des éléments technolangagiers spécifiques : hashtags et mentions

L'interdépendance du langagier et du technologique qui est propre aux tech-

nodiscours se manifeste également par la présence d'éléments spécifiques, que Paveau (2017) nomme « technolangagiers », afin de mettre en évidence leur nature composite. Dans les tweets, deux catégories d'éléments technolangagiers jouent un rôle de premier plan : les hashtags et les mentions.

Le hashtag représente l'un des éléments les plus emblématiques des discours sur Twitter, même si d'autres réseaux socio-numériques l'ont ensuite intégré à leur tour. Ce segment technolangagier précédé du signe dièse # se charge en effet d'une pluralité de fonctions, ainsi que de nombreux travaux l'ont souligné (parmi d'autres, Zappavigna 2011 ; Paveau 2013a, 2017 ; Mercier 2013, 2018 ; Jackiewicz & Vidak 2014 ; Julliard 2016). Le hashtag remplit en premier lieu une fonction d'indexation qui inscrit le tweet dans un réseau interdiscursif spécifique. Tous les discours incluant un même hashtag sont ainsi liés matériellement entre eux. Il peut néanmoins jouer un rôle significatif aussi au niveau sémantique, pragmatique et argumentatif, qu'il soit intégré ou non à la syntaxe du message principal (cf. Paveau 2017 ; Jackiewicz & Vidak 2014 ; Mercier 2018). Une fonction sociale est également assurée par ce technomot qui favorise une « affiliation diffuse » (Zappavigna 2011 ; Paveau 2017) à travers la constitution de réseaux interconnectés autour de valeurs et de sujets d'intérêt partagés (Zappavigna & Martin 2018 ; Jackiewicz & Vidak 2014).

Les mentions se composent en revanche d'une suite de caractères alphanumériques précédés de l'arobase. Elles désignent le nom du compte correspondant par le biais de cette sorte de pseudonyme, qui est créé par tout utilisateur au moment de son inscription à la plateforme. Des pratiques technodiscursives spécifiques intègrent la présence de mentions, comme c'est le cas pour les tweets-réponses ou encore pour le partage de contenus externes à la plateforme via les technosignes respectifs¹². Cependant, tout utilisateur peut également se servir de ce technomot pour réaliser plusieurs fonctions discursives. Bigey et Simon (2018 : 66-67) en identifient deux principales : une fonction d'adresse et une fonction justement de mention, au sens de citation. Dans le premier cas, la mention cherche à engager le compte cité en s'adressant directement à lui, comme dans un échange interactionnel, ou bien en le convoquant comme forme d'attestation énonciative. Dans le second, elle « relève d'un usage autonome du signe, où le nom du compte – et par métonymie la personne à

12. Bigey & Simon (2018 : 59) appellent ce type de mentions « serviles » puisqu'elles sont générées automatiquement lors de la rediffusion de contenus externes via Twitter par un bouton de partage intégré. Elles consistent en la mention du compte correspondant à la source du contenu externe partagé – généralement un compte média – précédé souvent de la préposition *via*.

qui appartient ce compte – fait l’objet du discours » (Bigey & Simon 2018 : 66). La mention peut ainsi représenter une ressource dialogique significative, employée dans une orientation aussi bien interlocutive, lors de la recherche d’un contact interactionnel avec le compte mentionné, qu’interdiscursive¹³, lorsqu’elle désigne l’origine énonciative au sein de pratiques citationnelles avec mention.

3.2.2.2. Degrés de pluritechnosémioticit 

Une premi re distinction concerne le degr  d’autonomie s mantique des ressources linguistiques et technolangagi res – telles que les hashtags et les mentions – dans la construction du sens du technodiscours. Les types de ressources s miotiques dont le tweet se compose et la fa on dont il est technodiscursivement structur  n cessitent en effet des activit s d’interpr tation-r ception diff rentes, et en cons quence des approches m thodologiques  galement diff rentes. Voil  pourquoi nous avons distingu  les « tweets simples » des « tweets augment s ». Les *tweets simples* ne comportent qu’une fen tre technodiscursive sans inclure des objets s miotiques d’augmentation (OSA). Correspondant principalement   un message verbo-linguistique, la partie textuelle des tweets simples est ainsi dot e d’un degr  d’autonomie s mantique plus  lev . En revanche, dans les *tweets augment s*, la partie textuelle est augment e par d’autres ressources (pluri)s miotiques (liens URL, images, vid es, GIFs...). Ces composants (pluri)s miotiques participent en effet   la co-construction du sens global du technodiscours, et ils sont susceptibles d’affaiblir l’autonomie s mantique de la partie textuelle.

Form s uniquement d’une fen tre technodiscursive, les *tweets simples* peuvent se composer – selon la d finition que nous proposons – de ressources de nature linguistique, pictographique ( mojis,  motic nes) et/ou technolangagi re (technomots),   l’exclusion des liens URL. Les tweets simples peuvent donc inclure des formes technolangagi res qui entraînent une augmentation  nonciative,   l’instar des hashtags et des mentions, tout en restant plut t autosuffisants d’un point de vue s mantique. En l’occurrence, la s miose du tweet peut  tre saisie de mani re exhaustive m me en dehors de son environnement de production, dans la mesure o  l’on dispose de la presque totalit  des constituants s miotiques du message.

13. Pour plus de d tails sur les diff rentes formes de dialogisme et notamment sur la distinction interlocutif-interdiscursif, on renvoie   Bres *et al.* (2019).

Les *tweets augmentés* se distinguent en revanche par une fenêtre technodiscursive qui est justement augmentée par des éléments pouvant enrichir, prolonger, voire compléter la partie textuelle. C'est le cas d'éléments visuels, audiovisuels et/ou multimodaux (images, vidéos, GIFs, etc.), qui matérialisent donc différents modes sémiotiques interagissant avec le texte du tweet. En outre, la partie textuelle de ces tweets peut être également augmentée par des liens URL, qui peuvent se présenter comme une chaîne de caractères soudée et cliquable, ou ils peuvent également comporter un aperçu de nature textuelle, icono-textuelle et/ou audiovisuelle (*Twitter Card* ou « carte » Twitter) du contenu vers lequel le lien pointe. La forme d'affichage du lien influence ainsi le degré d'interprétation du technodiscours dans sa globalité. Un aperçu du lien fournit en effet plus d'éléments pour l'interprétation du technodiscours, avant de permettre d'accéder éventuellement à la page web correspondante.

La composition pluritechnosémiotique de ces tweets en est le trait distinctif et définitoire, selon la distinction que nous avons proposée. Dans les tweets augmentés, la partie textuelle, qui peut impliquer des éléments linguistiques, technolangagiers et pictographiques (émojis, émoticônes...), interagit avec d'autres modes sémiotiques et avec des éléments technologiques. De cette interaction pluritechnosémiotique découle le sens du technodiscours. Par conséquent, nous définissons comme « objets sémiotiques d'augmentation » (OSA) les contenus pluritechnosémiotiques qui, n'étant pas inclus dans la fenêtre technodiscursive du tweet, augmentent l'énonciation de la partie textuelle, coconstruisant ainsi le sens global du technodiscours. Chaque OSA entraîne une augmentation de la partie textuelle d'ordre aussi bien énonciatif que discursif : plusieurs actes énonciatifs et/ou plusieurs modes sémiotiques interagissent dans la création du message global. Ces relations pluritechnolagagères peuvent en outre se réaliser aussi bien à l'intérieur d'un seul tweet – comme c'est par exemple le cas d'une relation texte-image au sein d'un tweet augmenté d'une photo – qu'à son extérieur – lorsque des éléments d'augmentation mènent à des contenus non inclus dans le tweet, à l'instar des liens URL. En tout cas, l'activité de lecture et d'interprétation est augmentée (Paveau 2017), même si c'est toujours l'instance de réception-interprétation qui trace ses parcours d'« écriture ».

Une différence majeure entre les tweets simples et les tweets augmentés concerne donc le degré d'autonomie sémantique de la partie textuelle qui les compose, ainsi que l'autonomie sémiotique du technodiscours en général. Pour les tweets simples, il est possible de soumettre leur partie textuelle

à des analyses informatisées sans pour autant compromettre l'accès au sens du technodiscours dans sa globalité. En revanche, ce n'est pas toujours le cas pour les tweets augmentés, dont le sens découle de l'interaction entre différentes ressources sémiotiques et technologiques, parmi lesquelles la partie verbo-textuelle. Si une interprétation contextuelle de tout discours est toujours préférable, la quantité des données à analyser ainsi que les objectifs de l'étude peuvent néanmoins rendre nécessaire d'utiliser des outils d'exploration informatisée. Dans ce cas, il est raisonnable de supposer que le traitement textométrique de la partie textuelle des tweets simples est plus fiable que celui de la partie textuelle des tweets augmentés, puisqu'ils présentent des degrés de pluritechnosémioticités différents et, en conséquence, des degrés d'autosuffisance sémantique au niveau verbo-textuel également distincts.

De fait, seuls les émojis – ou encore les émoticônes – peuvent éventuellement modifier la portée pragmatique et donc la valeur sémantique du message des tweets simples. Parmi les diverses fonctions que les émojis peuvent remplir (Magué *et al.* 2020), c'est sans doute leur fonction de modalisation qui est susceptible d'influencer davantage le sens du technodiscours, lorsqu'ils activent une lecture interprétative de second degré. Dans le cas où ils ajoutent une dimension émotionnelle à la partie verbale, voire l'intensifient, ou lorsqu'ils correspondent à des symboles à valeur représentationnelle ou déictique (cf. Magué *et al.* 2020), ils ont des répercussions mineures sur l'autosuffisance sémantique des tweets simples, influençant plutôt leur degré d'expressivité.

On pourrait ainsi se demander s'il aurait été pertinent d'envisager une distinction supplémentaire interne aux tweets simples en fonction de la présence ou non d'éléments pictographiques, à l'instar des émojis. Si, dans notre cas, des limites techniques s'imposent quant à l'identification et au filtrage automatique des tweets incluant des émojis, une démarche combinant des analyses quantifiées et contextualisées peut néanmoins permettre de bien gérer l'étude des tweets simples selon la définition proposée. De fait, notre choix de ne pas faire de distinction supplémentaire à l'intérieur des tweets simples a été dictée par deux autres raisons. En premier lieu, on peut supposer que la présence d'émojis susceptibles de modifier considérablement le sens du tweet a un impact limité dans le traitement de données agrégées, ne biaisant donc pas les résultats quantifiés. En second lieu, l'exploration en contexte de cas représentatifs des analyses agrégées a permis de vérifier les hypothèses avancées et d'attester les interprétations proposées, en prenant en considération également d'éventuels tweets contenant des émojis.

3.2.2.3. Pratiques technodiscursives

Dans le cadre des textualités numériques natives, le dispositif technologique conditionne aussi bien la construction du technodiscours que les processus de réception-interprétation et les activités d'écrilecture. Dans le cas spécifique de Twitter, plusieurs formats de tweets et diverses fonctions ont été développés par la plateforme au fil du temps.

Le partage d'un contenu, correspondant initialement à un message limité à 140 caractères, via la plateforme représente la forme d'origine et la plus canonique de tweet que nous appelons, par souci de clarté, « tweet original », afin de le distinguer des autres types de tweets introduits par la suite. La désignation « tweet original » vise en effet à condenser deux aspects principaux de ce type de tweet : le fait qu'il représente le premier format des technodiscours de la plateforme, dont les caractéristiques et les contraintes ont pourtant évolué progressivement, d'une part, et le statut technodiscursif d'ordre initiatif qui le caractérise, de l'autre¹⁴. Des fonctions introduites par la suite présupposent au contraire une relation matérielle de nature technodiscursive avec un contenu antérieur, dont elles dépendent pour leur activation. Certaines impliquent un acte de rediffusion ou de validation d'un contenu publié, à l'instar du *retweet* et du *j'aime*. D'autres contribuent en revanche à la création d'un contenu additionnel voire complémentaire à un contenu déjà existant dans la plateforme, comme c'est le cas pour les *réponses* et les *retweets avec citation* (« retweet as quote-tweet » dans leur dénomination originale).

D'un point de vue technodiscursif, il est en conséquence possible de distinguer plusieurs « types » de tweets qui correspondent non seulement à des « formats » différents, mais incarnent notamment des pratiques technodiscursives distinctes en fonction aussi de leur finalité pragmatique. Dans le cadre d'une étude technodiscursive, on peut ainsi envisager cinq types de tweets principaux : les tweets originaux, les retweets, les réponses, les threads ainsi que les retweets avec citation. Les quatre premiers peuvent prendre la forme de tweets aussi bien simples qu'augmentés¹⁵, tandis que les retweets avec citation constituent des tweets aug-

14. Dans le cadre de cette dénomination, l'adjectif « original » ne concerne donc pas l'authenticité du contenu publié.

15. Une précision supplémentaire s'impose néanmoins dans le cas des retweets que l'on peut considérer comme « simples » par rapport aux retweets avec citation. Consistant en la rediffusion d'un tweet antérieur, les retweets « simples » impliquent une forme d'augmentation qui n'affecte pas la construction du sens, bien qu'ils véhiculent généralement une forme de validation/appréciation de la part du compte qui le rediffuse. Dans le cadre d'une étude technodiscursive portant sur la construction du sens, il nous

mentés de par leur construction technodiscursive. En effet, les retweets avec citation incluent toujours un tweet antérieur auquel s'ajoute le commentaire du compte qui le rediffuse, avec ou sans autres objets sémiotiques d'augmentation. En l'occurrence, c'est le tweet cité qui fonctionne à l'instar d'un objet sémiotique d'augmentation, puisqu'il interagit avec la fenêtre technodiscursive du tweet citant – c'est-à-dire le commentaire de l'utilisateur qui le rediffuse – en l'augmentant par un acte d'énonciation seconde¹⁶. Les retweets avec citation présentent ainsi une structuration technodiscursive déterminée, présupposant par sa nature même au moins une forme d'augmentation, ce qui n'est pas le cas des retweets « simples » ne comportant aucun commentaire supplémentaire.

Compte tenu des traits macro-génériques des tweets ici mis en avant, l'analyse de la variation intragénérique s'appuiera sur les deux niveaux d'hétérogénéité internes au corpus TW : l'hétérogénéité pluritechnosémiotique, dont relève l'opposition entre tweets simples et tweets augmentés, et l'hétérogénéité technodiscursive, dont relève la distinction des divers types de tweets. Nous précisons dès maintenant que l'étude du corpus TW a été circonscrite aux tweets simples (TWS) pour des raisons de comparabilité intergénérique. Trois types de tweets font ainsi l'objet de l'analyse : les tweets originaux, les réponses et les threads. En revanche, les retweets « simples » ne sont pas pris en compte puisqu'ils représentent des données dupliquées des tweets d'origine correspondants, pouvant donc biaiser les résultats quantitatifs. Nous résumons les caractéristiques technodiscursives des trois types de tweets simples pris en considération afin de mettre en évidence leurs traits énonciatifs et pragmatiques spécifiques.

Le *tweet original* consiste en un message que le twitteur souhaite partager via la plateforme. Il présente une nature discursive initiative, dans la mesure où il peut être comparé à une prise de parole qui n'enchaîne pas sur un message spécifique lui étant matériellement associé. En fonction de la présence éventuelle d'éléments technolangagiers, les tweets originaux simples peuvent toutefois avoir une orientation dialogique plus ou moins marquée et réaliser des fonctions pragmatiques différentes.

Le *tweet-réponse* est en revanche un tweet produit en réaction à un autre tweet. En ce qui concerne son format technodiscursif, il s'ouvre par la mention du ou des compte(s) au(x)quel(s) on répond, et se rattache au tweet

semble approprié de classer les retweets « simples » en fonction du type de tweet rediffusé, ce dernier étant simple ou augmenté.

16. Pour plus de détails sur la structuration technodiscursive des retweets avec citation, on renvoie à Cagninelli (2024c).

antérieur de manière visuelle et technologique par une ligne qui relie les deux profils. Du point de vue des contraintes sous-génériques, un tweet-réponse se caractérise par une orientation interlocutive marquée, puisqu'il s'agit d'un technodiscours réactif qui s'adresse directement au(x) compte(s) cité(s). Si une pareille situation peut se produire également dans un tweet original à travers l'usage de la mention, celui-ci est toutefois orienté en premier lieu vers un destinataire pluriel et indéterminé, ou bien vers une partie de la communauté de la plateforme lorsque le tweet est inscrit interdiscussivement par le hashtag, tout en sollicitant en particulier l'attention des abonnés du compte.

Quant au statut discursif de ces deux types de tweets, le tweet original, en tant que création première, est doté d'une plus grande autonomie sémantique que le tweet-réponse. Le tweet-réponse peut en effet être considéré comme un discours second, car il procède d'un discours antérieur auquel il réagit et dont il dépend sémantiquement. Par conséquent, la nature potentiellement et partiellement incomplète de ce discours pose des difficultés dans la phase d'analyse. Il conviendrait donc de prendre en compte le discours auquel la réponse réagit afin d'avoir accès à tous les éléments contextuels nécessaires à son interprétation. Cette approche n'est toutefois envisageable qu'en fonction d'un nombre limité de discours à analyser, en plus d'éventuelles contraintes imposées par l'extraction des données et par la constitution du corpus. Elle serait possible par exemple lors d'une analyse qualitative, en observant chaque réponse dans son environnement natif en relation avec le tweet auquel on réagit. Cependant, dans ce cas aussi, le tweet auquel la réponse est associée pourrait ne pas être accessible ou avoir été supprimé.

Enfin, les *threads* représentent une sorte d'évolution de la fonction « répondre », à la suite de la tendance des utilisateurs à y recourir dans le but de créer des messages plus longs. Comme le fait remarquer Bibié (2019 : 62), la pratique du thread « s'inspire d'un détournement de l'outil » avec l'objectif de « contourner la politique de Twitter concernant le nombre de caractères ». Les threads sont en effet des « fils discursifs », c'est-à-dire un ensemble de tweets agencés entre eux et publiés en succession par un même utilisateur. Bien qu'ils impliquent un tweet initiatif avec fonction d'ouverture du fil discursif, la nature même de la pratique se définit par l'enchaînement de plusieurs tweets reliés entre eux. Technodiscours caractérisé par un degré de narrativité plus marqué, le thread a un statut discursif et textuel particulier et peut poursuivre diverses visées, ainsi que le montre Longhi (2022).

À la lumière de leurs traits spécifiques, les trois types de tweets simples que nous avons distingués gagnent à être considérés comme des formes sous-génériques différentes lors de la constitution du corpus, afin de donner accès à l'étude des variations intragénériques au sein du technogène des tweets. Dans la phase de constitution du corpus (3.2.4), nous démontrons l'intérêt d'intégrer les deux niveaux d'hétérogénéité à la structure du corpus. Nous limitons ensuite l'analyse (chapitre 5) à un corpus de tweets simples (TWS) afin de favoriser la comparaison intergénérique avec le corpus relevant de la presse en ligne. Nous prenons ainsi en compte une hétérogénéité comparable centrée sur le matériau linguistique des deux formes de textualités numériques sélectionnées. L'analyse linguistique se caractérise toutefois par une approche discursive et écologique qui met en relation la textualité aussi bien avec ses extérieurs extradiscursifs qu'avec les composantes technologiques interagissant dans la construction des parcours de sens. Les enjeux théoriques et méthodologiques posés par la nature technodiscursive de ces formes textuelles sont abordés plus amplement dans les sections suivantes.

3.2.3. *Textualités technodiscursives : quels défis méthodologiques ?*

Le corpus d'étude des technodiscours de la plateforme Twitter (désormais TW) se caractérise par des traits composites et amplifiés par rapport non seulement aux textualités traditionnelles, mais aussi aux textualités numérisées et numériques (Paveau 2015), à l'instar de celles regroupées dans le corpus PN. Ces spécificités technodiscursives posent donc plusieurs défis méthodologiques à la constitution du corpus, tout comme à son exploration.

Ainsi que le remarque Emerit (2016), l'instabilité, la mixité et l'incomplétude propres aux discours numériques natifs se heurtent à la nature nécessairement délimitée et figée d'un corpus. En conséquence, la constitution d'un corpus restituant des technodiscours dans leur écologie représente une entreprise ardue pour l'analyste. Néanmoins, « renoncer à cette instabilité, c'est perdre la nature technolangagière des données pour ne conserver qu'une image de celles-ci » (Emerit 2016, en ligne : §22). La notion de « lieu de corpus » avancée par Emerit (2016, en ligne : §4) permet ainsi de mettre en avant une « potentialité qui [...] est subordonnée » au corpus.

Ce qui paraît toutefois particulièrement significatif de cette notion pour notre étude, c'est la conception et la représentation arborescente du lieu du corpus, qui met en relief sa structure hiérarchique et évolutive (Emerit 2016). En effet, il nous semble que la constitution de corpus multiples et emboîtés –

comme le suggère aussi Emerit (2016) à la suite de Pincemin – peut favoriser, en l'état actuel de la recherche, une exploration écologique complexe de ces productions à la lumière de leurs traits distinctifs et définitoires. Le lieu de corpus représente ainsi le point de départ pour la création de « plusieurs corpus différents qui seront constitués de figements focalisés sur une partie des données accessibles » (Emerit 2016, en ligne : §4). Dans le cadre de ce travail, la création des métadonnées sera un aspect essentiel pour créer des (sous-) corpus spécifiques sur la base des traits distinctifs opératoires pour l'étude de la variation intragénérique (cf. 3.2.4). Les résultats seront enfin comparés avec ceux du corpus PN dans une optique intergénérique.

3.2.3.1. Aborder la complexité technodiscursive : entre extraction et contextualisation

Produit d'une technologie discursive (Paveau 2017), un tweet procède de l'interaction entre langage et technologie, tout en pouvant intégrer également d'autres modes sémiotiques. Vu que le technologique participe matériellement à l'élaboration des technodiscours, une analyse discursive ne peut ignorer ses composantes technologiques et plurisémiotiques, même si elle s'intéresse en premier lieu à la matérialité langagière.

La nature technodiscursive des tweets pose plusieurs défis à la constitution d'un corpus qui rende compte de leur complexité et de leurs traits spécifiques. Néanmoins, le format électronique du corpus permet d'intégrer la représentation des composantes et des spécificités technodiscursives des tweets sous forme de métadonnées en fonction des objectifs de l'étude. Les métadonnées correspondent à des informations que l'on peut associer matériellement aux données au format numérique. La constitution de corpus dans un tel format permet ainsi d'ajouter toute information jugée nécessaire sur les données primaires, en l'intégrant à la structure même du corpus sous forme de métadonnées. Comme nous l'avons développé plus amplement ailleurs (Cagninelli 2024a), diverses catégories de métadonnées peuvent être prévues afin d'enrichir les données, constituant par ailleurs des paramètres à envisager lors de l'analyse et notamment d'une analyse outillée.

Pour tenir compte de la matérialité technolangagière des tweets à la lumière de leurs spécificités et de l'environnement relationnel qui les façonne, il est utile d'aborder la complexité de ces technodiscours à partir de plusieurs angles d'approche. La multiplication des angles d'approche des données favorise non seulement l'exploration de divers niveaux de leur construction technodiscursive, mais elle vise aussi à les appréhender dans

leur globalité et leur complexité par des approches variées et complémentaires, avant une synthèse finale.

3.2.3.2. Outiller l'approche écologique des technodiscours

Les nouvelles pratiques et textualités du web ont amené les chercheurs à réinterroger les cadres théoriques et les méthodes existants, proposant aussi de nouveaux concepts et outils méthodologiques. Le but est de les appréhender et de les décrire en tenant compte de leurs spécificités distinctives, tant langagières que technologiques et sémiotiques. Nous avons ainsi réfléchi à l'élaboration d'un outil permettant de rendre opératoire l'articulation écologique local-global dans les processus de constitution du corpus et de son exploration.

Dans le cadre de notre recherche doctorale, nous avons créé un tel outil, *Visaneco* (Cagninelli & Taglioli 2022)¹⁷, en collaboration avec un informaticien, Massimiliano Taglioli. Notre objectif était de mettre au point une démarche méthodologique outillée qui matérialise une approche écologique des technodiscours, combinant deux perspectives d'appréhension : une perspective extractive et une perspective contextualisée. La figure 3 illustre l'interface de *Visaneco* qui associe ces deux formes d'appréhension des données : à gauche, la visualisation analytique, qui relève de l'extraction, des composantes des tweets et de leurs métadonnées, et à droite, la visualisation écologique du tweet sélectionné, qui relève de la contextualisation. D'une part, la visualisation standard des tweets dans leur environnement natif via un navigateur web favorise une étude contextualisée de ces technodiscours, assurant notamment l'exploration de leur dimension augmentée et relationnelle. De l'autre, une visualisation analytique restitue les diverses informations associées à chaque tweet sous forme de métadonnées. Cette visualisation permet de focaliser l'attention sur des éléments ou des catégorisations spécifiques, en opérant aussi des filtrages des données.

Cet outil met donc en relation les deux approches principales initialement adoptées par les études linguistiques consacrées à ces textualités numériques. Si ces approches étaient d'abord distinctes (Paveau 2013b), plusieurs travaux et outils ont par la suite recherché un équilibre entre les deux (voir par exemple Longhi 2020b, 2021). L'évolution technologique et la mise en place de collaborations entre linguistes et informaticiens ont sans doute

17. Pour une description détaillée de la version plus récente, on renvoie à Cagninelli (2022).

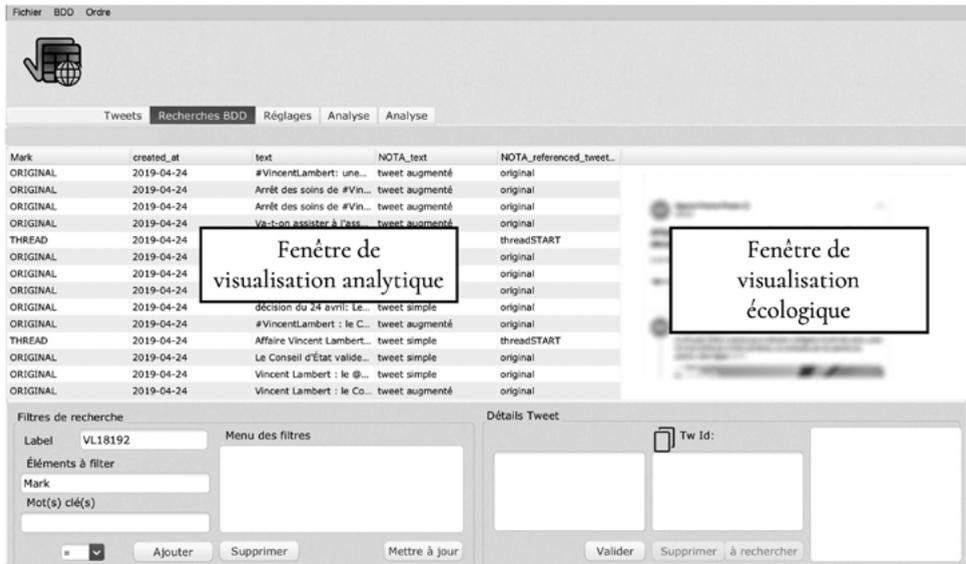


Figure 3. Interface principale de Visaneco.

favorisé la création de nouvelles démarches méthodologiques, permettant de mettre en rapport des angles d'approche apparemment divergents. Le logiciel *Visaneco* a été élaboré sur la base des compétences respectives de ses auteurs¹⁸, résultant ainsi d'une interaction constante entre conceptualisation linguistique et élaboration/réalisation informatique. Il contribue en effet à outiller une démarche écologique informatisée de l'étude des tweets, afin de rendre opératoire une approche théorique et méthodologique spécifique de ces technodiscours. En premier lieu, la visualisation analytique des composants des tweets et de leurs métadonnées, sélectionnées sur la base des objectifs de l'étude, permet d'explorer de manière distincte et ciblée les différents éléments technodiscursifs ainsi que les traits caractéristiques qui leur sont propres. À la lumière des objectifs de la présente étude, il est possible de filtrer, par exemple, les différents types de tweets sur la base des pratiques technodiscursives matérialisées (cf. 3.2.2.3) ou encore de distinguer les tweets simples des tweets augmentés en fonction de la présence éventuelle d'OSA (3.2.2.2). Le logiciel constitue donc un apport significatif à la constitution de corpus et de sous-corpus en exploitant aussi les métadonnées. Ces dernières peuvent restituer tant des informations de

18. Le modèle conceptuel est de notre fait et s'attache à implémenter une vision écologique – à la fois quantifiée et contextualisée (cf. aussi Longhi 2019, 2020b, 2021) – des technodiscours sur Twitter, alors que Massimiliano Taglioli a développé le code et l'interface de l'outil.

nature contextuelle aux technodiscours, qu'en isoler des éléments constitutifs, mais elles peuvent aussi relever d'annotations linguistiques, discursives et technodiscursives établies par l'analyste (cf. Cagninelli 2024a). *Visaneco* crée donc les conditions nécessaires pour effectuer des recherches plus complexes et plus fines, quantifiées et contextualisées, contribuant ainsi à multiplier les approches et les niveaux d'analyse.

En deuxième lieu, la visualisation contextualisée donne accès à la matérialité technodiscursive des tweets, incluant leur dimension plurisémiotique, augmentée et relationnelle. Elle assure le retour au texte dans son environnement de production, qui est nécessaire pour saisir le technodiscours dans sa globalité, tout en permettant de vérifier et de contextualiser les analyses quantifiées. On vise ainsi un équilibre entre des études quantifiées et des études contextualisées, et le retour au texte s'avère essentiel dans l'articulation des deux approches, comme le souligne Longhi (2020b, 2021).

Chacune des deux perspectives d'appréhension offertes par *Visaneco* est ainsi au service de méthodes d'analyse distinctes mais complémentaires, dont l'articulation favorise une analyse plus complète des technodiscours. La perspective extractive vise à rendre la nature complexe et composite des technodiscours plus accessible à des analyses outillées, en isolant leurs différentes composantes. À cette fin, nous avons par ailleurs proposé d'envisager une macro-catégorie de métadonnées spécifique, que nous avons appelée « métadonnées de décomplexification » (Cagninelli 2024a). En séparant les constituants des technodiscours, la perspective extractive permet de soumettre les données ainsi visualisées, de même que les métadonnées associées, à des explorations informatisées. Il est donc possible d'effectuer des analyses textométriques sur la partie textuelle des technodiscours, obtenant ainsi des résultats quantifiables et reproductibles. Des analyses inductives sont également favorisées par ce type d'explorations, permettant d'avoir des aperçus globaux du corpus en fonction de divers paramètres sélectionnés. L'approche extractive des technodiscours crée finalement les conditions pour mettre en place des analyses outillées, favorisant des explorations plus quantifiées (cf. 1.4). L'interprétation de ces résultats requiert toutefois un retour au texte dans son environnement de production, en complément d'analyses plus qualitatives et contextualisées.

En bref, *Visaneco* comporte deux formes de visualisation correspondant à deux approches complémentaires des données, l'une plus extractive et logocentrique et l'autre plus contextualisante, qui s'enrichissent

réciroquement. Il permet une appréhension plus complexe du corpus, en favorisant la mise en relation des niveaux locaux de chaque tweet et/ou du corpus avec le niveau global et contextualisé de chaque technodiscours et, plus en général, du corpus dans son ensemble. En d'autres termes, sous *Visaneco*, la nature délimitée et figée du corpus s'ouvre sur les dimensions numérikée, multimodale, évolutive, interactive et interconnectée du lieu du corpus, mises en relief par Emerit (2016). *Visaneco* combine ainsi des données figées au moment de leur extraction, avec l'exploration actualisée, multimodale et interconnectée, qui permet d'analyser et d'enrichir les données analytiques à la lumière de leur dynamisme et de leur évolution. Les différents niveaux de l'arborescence du lieu du corpus de Emerit (2016) convergent et coexistent dans le logiciel *Visaneco*. En nous inspirant de l'arborescence du lieu du corpus proposée par Emerit (2016), nous illustrons le fonctionnement de *Visaneco* de manière graphique (figure 4), en montrant des reconfigurations impliquées par l'articulation d'une approche extractive-quantifiée et d'une approche écologique-contextualisée.

L'articulation équilibrée des deux formes de visualisation et d'exploration des données contribue ainsi à la création de « corpus réfléchis », à savoir d'« un corpus qui peut effectuer un “retour sur lui-même”, soit dans sa matérialité formelle, soit dans son environnement natif » (Longhi 2021 : 111).

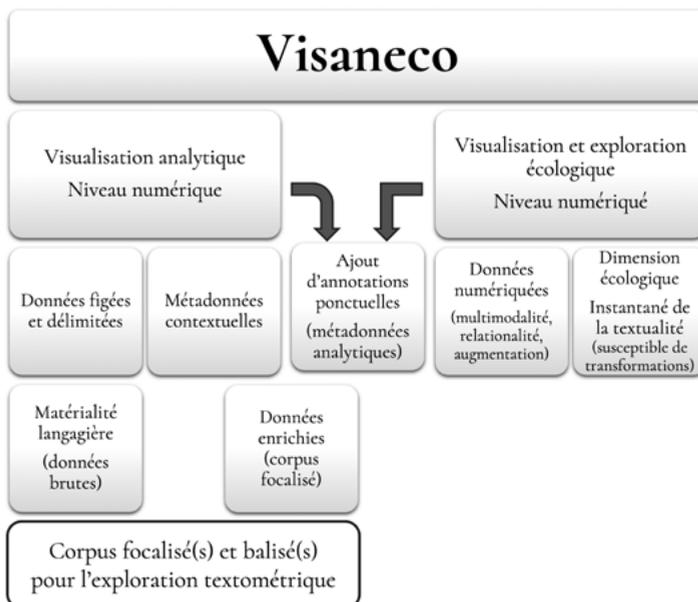


Figure 4. Représentation de la matrice conceptuelle de l'exploration écologique outillée sous *Visaneco*.

3.2.4. *Constituer un corpus de technodiscours selon un principe de variation interne et externe*

Afin d'assurer la comparabilité intergénérique entre les deux corpus d'étude (PN et TW), la sélection des tweets a été effectuée sur la base de critères équivalents à ceux établis pour le sous-corpus PN, tout en les adaptant aux spécificités génériques des technodiscours envisagés. Le critère thématique a donc présidé à la sélection des tweets en fonction de la présence des expressions « Vincent Lambert » ou « VincentLambert », apparaissant non nécessairement dans la partie textuelle, mais aussi au sein d'éléments technolangagiers, tels que des technomots ou des liens URL. Au critère temporel s'est en outre ajouté un critère linguistique circonscrivant la sélection aux tweets écrits en français. Comme on l'a déjà anticipé, les retweets ont été exclus de la recherche pour éviter de biaiser les résultats quantifiés, vu qu'ils ne constituent que des données dupliquées.

Les tweets retenus en fonction de ces critères ont été mis en relation avec trois catégories principales de métadonnées, à la suite d'un travail de sélection et de synthèse eu égard aux objectifs de l'étude :

1. des métadonnées de contextualisation (genre de discours, date de publication, type de tweet) ;
2. des métadonnées concernant la plurisémioticit  du tweet (pr sence/absence de liens URL et d'autres OSA) ;
3. des m tadonn es concernant la dimension technolangag re des tweets (pr sence/absence de hashtags et mentions).

Il ressort qu'une partie des m tadonn es, notamment celles de contextualisation g n rique et temporelle sont ais ment rep rables et comparables selon les diverses textualit s rassembl es dans le corpus CR2. D'autres peuvent en revanche  tre analys es dans une perspective interg n rique seulement si elles identifient des aspects assimilables sur le plan linguistico-discursif. C'est par exemple le cas pour les divers types de tweets qui, puisqu'ils pr sentent des diff rences de nature  nonciative et pragmatique, peuvent constituer un niveau d'h t rog n it  intrag n rique comparable   celui relevant des divergences entre les articles d'information et les articles d'opinion du corpus PN. Cependant, cette comparaison interg n rique pr suppose la cr ation d'un sous-corpus de TW ne rassemblant que les tweets simples (le sous-corpus TWS), car ceux-

ci incarnent une forme de textualité plus proche de celle des articles de journal en ligne, contrairement aux tweets augmentés. La comparabilité entre les articles de journal en ligne et les tweets simples est en effet rendue possible par leur haut degré d'autosuffisance sémantique, déterminé par l'absence d'OSA, qui constitue en revanche le trait définitoire des tweets augmentés (cf. 3.2.2.2).

Enfin, on repère également des métadonnées issues d'annotations liées à des critères intragénériques qui ne sont spécifiques que du corpus TW en vertu de ses traits technodiscursifs distinctifs, à l'instar de la présence d'OSA, de hashtags et/ou de mentions. Elles peuvent donc déterminer divers niveaux d'hétérogénéité intragénérique, sans pour autant assurer la comparaison intergénérique, du moins avec le corpus PN. Quoi qu'il en soit, les métadonnées et les annotations sont incontournables pour rendre compte de la structure interne du corpus, pour établir les niveaux d'hétérogénéité intra- et intergénérique, ainsi que pour constituer des sous-corpus d'étude spécifiques.

3.2.4.1. Hétérogénéité sous-générique d'ordre technodiscursif

Les métadonnées issues des annotations relatives aux quatre types de tweets dont se compose le corpus TW (tweets originaux, tweets-réponses, threads et retweets avec citation) permettent d'avoir un aperçu de la présence quantitative des pratiques technodiscursives respectives qu'ils matérialisent principalement (cf. 3.2.2.2). Plus précisément, il ressort qu'environ deux tiers (68 %) des technodiscours du corpus TW sont des tweets originaux. Les réponses représentent la deuxième pratique technodiscursive la plus fréquente (21 %), alors que les retweets avec citation (8 %) et les threads (3 %) sont plus marginaux au niveau quantitatif.

Cette exploration du corpus par l'annotation (cf. Poudat & Landragin 2017 : 35sq) donne accès à un premier niveau d'hétérogénéité intragénérique d'ordre technodiscursif. On constate en effet que des pratiques généralement à orientation interdiscursive, telles que les tweets originaux et les threads, sont largement dominantes dans ce corpus, tandis que des pratiques à orientation interlocutive, comme les réponses et les retweets avec citation, sont moins présentes. Il serait de ce fait intéressant de comparer ces résultats quantifiés à travers différents corpus thématiques constitués de manière similaire, afin de vérifier s'il s'agit de tendances assez générales liées aux types de tweets ou si elles sont caractéristiques de l'événement discursif dont ces technodiscours relèvent.

Étant donné que la littérature reconnaît une nature polémique plus marquée à la pratique des réponses (cf. par exemple Julliard 2016), on peut en outre émettre l'hypothèse que la capture du débat restituée par le corpus TW n'inclut qu'un nombre limité des échanges interactionnels pouvant se caractériser par un ton polémique. En tant que discours second, les réponses peuvent par ailleurs omettre des éléments de thématisation et de contextualisation du message qui seraient assurés par le tweet auquel on réagit. Pour cette raison, au moins une partie des tweets-réponses liés au débat pourraient avoir été exclus de la sélection par mots-clés. En raison de leur nature initiative et interdiscursive, la plupart des technodiscours du corpus pourraient en revanche relayer des informations relatives à l'événement ou encore le commenter, sans impliquer une relation interlocutive spécifique et ciblée, mais en s'adressant plus largement à la communauté des internautes.

Il est enfin important de tenir compte de ces résultats pour interpréter aussi les analyses quantifiées effectuées sur d'autres niveaux du corpus. On remarque en effet que d'autres rapprochements et divergences caractérisent les comportements des quatre pratiques même au niveau de la composition plurisémiotique, ainsi que l'illustre la section suivante.

3.2.4.2. Hétérogénéité sous-générique d'ordre pluritechnosémiotique

Les annotations concernant la présence d'éventuels OSA permettent de rendre opératoire la distinction entre tweets simples et tweets augmentés (cf. 3.2.2.2) au sein du corpus TW. La mise en place de cette distinction sous forme de métadonnées matérialise le niveau d'hétérogénéité nécessaire à l'analyse de la variation intragénérique d'ordre plurisémiotique. Pour ce faire, les différents types de tweets ont été mis en relation avec la composition technodiscursive qui les caractérise. Cette exploration, qui s'appuie sur les métadonnées, montre tout d'abord une présence plus élevée de tweets augmentés (62 %) que de tweets simples (38 %). Cela témoigne donc d'une tendance marquée à la construction de technodiscours augmentés et, éventuellement, plurisémiotiques.

Ce croisement d'annotations fait en outre ressortir des différences intéressantes quant à la présence d'OSA en fonction des divers types de tweets retenus, comme l'illustre le graphique de la figure 5. À cet égard, il convient toutefois de noter que, dans le cas des retweets avec citation, l'analyse a porté sur la présence/absence d'OSA supplémentaires au tweet cité dans un souci de comparaison intragénérique d'ordre technosémiotique et

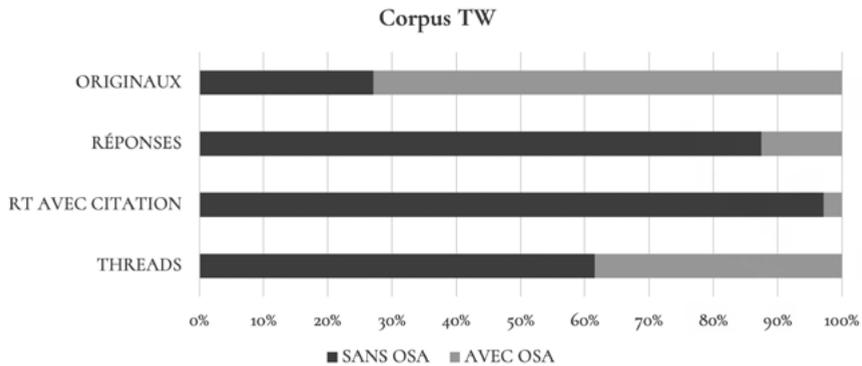


Figure 5. Présence/absence d'OSA en fonction des divers types de tweets.

sous-technogénérique. Contrairement aux autres types de tweets envisagés, les retweets avec citation sont en effet le seul format qui, de par sa nature même, comporte un élément d'augmentation correspondant au tweet cité.

On constate ainsi des tendances opposées entre les tweets originaux, constituant généralement une pratique initiative et interdiscursive, et les pratiques plus réactives et interlocutives, telles que les réponses et les retweets avec citation. Alors que la grande majorité des tweets originaux sont augmentés par un ou plusieurs OSA, les réponses et les retweets avec citation ne les incluent que dans un nombre limité de cas.

On peut donc supposer que ces comportements différents dépendent de deux facteurs : le statut discursif des types de tweets ainsi que leur composition technodiscursive. Pour ce qui concerne le statut discursif, les tweets originaux sont des productions primaires, car ils correspondent à un acte initiatif qui ne s'enchaîne pas directement et matériellement sur un tweet antérieur. Ils sont ainsi souvent augmentés de divers OSA pour atteindre plusieurs finalités pragmatiques. Il faut néanmoins rappeler que la diffusion de contenus externes à la plateforme peut se réaliser également par le bouton intégré sur des sites tiers, qui comporte la création d'un tweet original incluant le lien vers la page correspondante. Au contraire, les réponses et les retweets avec citation constituent des productions secondaires, dans la mesure où ils procèdent d'un autre tweet, à la fois pour leur production et leur interprétation. Constituant une sorte de prolongement d'un technodiscours antérieur, ces pratiques ont moins souvent recours à des OSA qui les prolongeraient à leur tour.

Quant à la construction technodiscursive des différents types de tweets, elle peut à son tour influencer sur la présence d'OSA et cela notamment dans

le cas des retweets avec citation. Comme nous l'avons précisé plus haut, les retweets avec citation ne peuvent correspondre qu'à des tweets augmentés du fait de leur structuration même, car le tweet cité qu'ils intègrent par définition est tout à fait assimilable à un OSA. Par conséquent, il est moins fréquent que ces technodiscours incluent des objets sémiotiques d'augmentation supplémentaires, qui complexifieraient davantage la construction du sens.

Les threads constituent enfin un cas à part, à mi-chemin entre les tweets originaux, d'un côté, et les réponses et les retweets avec citation, de l'autre. En effet, bien qu'ils correspondent principalement à des tweets simples dans le corpus, leur tendance à être augmentés par des OSA est plus marquée que dans les pratiques d'origine secondaire à orientation interlocutive. Dans ces fils discursifs, ce sont justement les tweets consécutifs au premier qui en représentent un prolongement. Compte tenu de leur statut discursif, on peut toutefois supposer que, dans les threads, les relations entre la partie textuelle et l'OSA sont similaires au cas des tweets originaux.

L'observation des tendances principales relatives à la composition pluri-sémiotique des différents types de tweets étaye davantage encore les divergences de leurs fonctionnements technodiscursifs, corroborant la nécessité de les appréhender de manière distincte à la fois épistémologiquement et méthodologiquement. Dans le cadre du présent ouvrage, l'étude est donc limitée au corpus de tweets simples (désormais TWS) afin d'assurer la comparaison intergénérique. L'autosuffisance sémantique (et sémiotique) plus élevée des tweets simples (cf. 3.2.2.1) favorise en effet l'étude contrastive avec les textualités du corpus PN, les deux pouvant être explorés selon une approche similaire, mais à la lumière de leurs spécificités génériques respectives.

3.3. Démarches d'exploration du corpus : une approche multidimensionnelle combinée

Le corpus dont l'on vient de présenter la structuration est de type textuel (Mayaffre 2005), et a été construit en fonction d'une problématique discursive. Il s'agit en outre d'un corpus contextualisé, puisqu'il porte en soi des éléments relatifs aussi bien à ses conditions de production – matérialisées sous forme de métadonnées – qu'à sa dimension interdiscursive, en fonction de sa composition interne. Les différents niveaux d'hétérogénéité (Pince-min 2012a) constitutifs de ce corpus sont au service d'une démarche d'ex-

ploration également hétérogène, qui présuppose l'articulation de plusieurs méthodes et niveaux d'analyse pour avoir accès à la variation (Garric 2012), en multipliant les angles d'approche des données (Longhi 2018).

La mise en forme électronique du corpus et le balisage de ses niveaux d'hétérogénéité grâce à la présence des métadonnées sont indispensables pour permettre l'utilisation des différents outils d'exploration en fonction des objectifs de l'analyse. Dans le cadre de notre étude, l'apport de l'informatique et du numérique¹⁹ s'avère incontournable, aussi bien dans la phase de collecte des données et de structuration du corpus, que dans les phases d'analyse par le biais d'outils textométriques. En l'occurrence, le rôle de l'informatique et du numérique devient significatif dans la relation avec les processus de réflexion et d'interprétation engagés par l'analyste. Il s'agit en effet d'une interaction constante humain-machine. D'une part, le corpus procède des hypothèses de recherche de l'analyste et donne corps à son cadre théorico-méthodologique. De l'autre, l'analyste accède au terrain d'étude dans la forme du corpus (Longhi 2018), s'y réfère pour (re)construire son sens à travers de multiples parcours d'exploration et d'interprétation, en traversant les divers paliers qui le composent (Rastier 2001, 2011 ; Mayaffre 2002a). C'est un échange réciproque et interdépendant qui a lieu entre l'analyste et le corpus. D'un côté, le corpus résulte des hypothèses et des objectifs qui président à sa constitution et n'a de sens qu'à la lumière d'une subjectivité interprétante (Mayaffre 2002a). De l'autre, l'analyste s'appuie sur le corpus comme un objet heuristique dans le cadre d'une approche herméneutique (Mayaffre 2002b, 2005) ; le corpus lui permet de justifier ses interprétations et d'avoir accès à des parcours de sens qui ne seraient pas possibles à la lecture du texte hors de la contextualisation assurée par le corpus (Rastier 2001, 2011 ; Mayaffre 2002a).

Comme nous l'avons déjà souligné (1.2), l'exploration outillée du corpus permet d'objectiver tant l'accès au texte que les méthodes d'analyse sur lesquelles la réflexion et l'interprétation de l'analyste prennent ensuite appui. Elle favorise la mise en relief de structures sous-jacentes ainsi que de phénomènes récurrents (Mayaffre *et al.* 2019). Pour ce faire, les métadonnées ajoutées en phase de constitution du corpus jouent un rôle essentiel (cf. aussi Cagninelli 2024a), notamment pour aborder les divers types de variation envisagés.

19. Rappelons par ailleurs que, dans notre recherche, les matériaux analysés eux-mêmes ont une nature numérique, ce qui présuppose de prendre en considération, dans l'analyse linguistico-discursive, le rôle de l'environnement numérique dans la co-construction de ces textualités, comme le soutient aussi Paveau (2017).

À la lumière de ces principes, notre méthodologie d'exploration du corpus se veut à la fois multidimensionnelle et combinée : multidimensionnelle, car elle envisage plusieurs niveaux d'analyse, et combinée en raison des diverses démarches et méthodes d'analyse exploitées afin de multiplier les angles d'approche des données. Elle se fonde, plus précisément, sur la combinaison d'analyses outillées quantifiées et contextualisées (cf. aussi Longhi 2020b, 2021), afin de dépasser l'opposition épistémologique et méthodologique entre quantitatif et qualitatif (Longhi 2019 ; Meunier 2020), d'où notre choix d'employer le terme *quantifié* (cf. 1.4). L'intégration d'approches complémentaires vise ainsi à favoriser une appréhension plus complexe de l'objet d'étude, comme nous l'avons souligné plus haut (1.2 et 1.4).

Cette démarche se déroule en trois étapes. La première concerne l'exploration du niveau global du corpus par le biais de méthodes inductives-quantifiées. Il s'agit de mettre en évidence les spécificités et la structuration interne du corpus, à travers des analyses thématiques et des calculs contrastifs en fonction des métadonnées retenues comme variables d'analyse, tout en faisant ressortir les similitudes et les divergences entre les diverses sous-parties et les niveaux d'hétérogénéité sélectionnés dans une perspective intragénérique. Ces opérations amènent également à identifier de nouvelles pistes d'exploration à approfondir et à formuler des hypothèses interprétatives à tester.

La deuxième phase s'intéresse en revanche aux niveaux locaux du corpus, qui sont explorés sur la base des résultats inductifs à travers des méthodes contextualisées appliquées à différentes « grandeurs intermédiaires » (Rastier 2011). Ces explorations permettent de creuser les résultats inductifs et de vérifier les hypothèses qui s'ensuivent à travers l'observation de phénomènes locaux d'ordre micro- et mésotextuel, interprétés à la lumière des contextes et des différents facteurs de variation. Plus précisément, l'étude porte sur les formes de référenciation des objets de discours ainsi que sur leurs procédés de mise en discours, en proposant une analyse fine des éléments linguistiques soumis à la variation et de leurs implications pragmatiques.

La troisième phase vise enfin une synthèse interprétative en articulant les résultats des deux phases précédentes ; les hypothèses interprétatives formulées lors de l'exploration inductive du corpus sont ainsi réinterrogées à la lumière des résultats des études contextualisées.

Cette démarche en trois étapes a pour objectif de donner accès aux interrelations existant entre les niveaux macrostructurels du corpus (le genre de discours et les sous-genres discursifs) et les niveaux micro- et mésos-

structurels (l'unité linguistique, l'énoncé...). En d'autres termes, il s'agit de mettre en relation les dimensions linéaire et réticulaire de la textualité (Adam 2020a) avec la matérialisation du contexte interdiscursif assuré par le corpus (Mayaffre 2002a ; Rastier 2011). L'étude se caractérise donc par des allers-retours constants entre le global du corpus et le local, entre les divers niveaux d'analyse micro- méso- et macro-textuels. Des allers-retours entre des analyses quantifiées et statistiques et des observations qualitatives en contexte sont également effectués. Le but est d'analyser les processus de construction discursive de la controverse à travers divers contextes et parcours d'exploration, en prenant en considération les déterminants sociodiscursifs d'ordre générique. Cette pluralité et complémentarité de démarches d'analyses visent, d'une part, à identifier des spécificités du corpus qui seraient difficilement perceptibles à une lecture linéaire et, de l'autre, à justifier les interprétations. Ainsi, l'articulation de méthodes différentes mais complémentaires ne déstabilise pas l'étude, au contraire elle l'enrichit. En effet, « le quantitatif et le qualitatif ne s'opposent aucunement : seule une analyse qualitative peut rendre significatifs des phénomènes quantitatifs remarquables » (Rastier 2001 : 214).

L'articulation complémentaire entre démarches quantifiées et qualitatives en présuppose également une deuxième, à savoir l'articulation entre le processus d'interprétation assuré par l'analyste, d'une part, et l'informatique, en tant que support et outil pour l'analyse, de l'autre. Ainsi que le souligne Pincemin,

Le lecteur humain est évidemment le meilleur interprète d'un texte [...]. Les atouts de l'ordinateur sont sa vitesse de calcul et sa mémoire : l'intérêt est de les mettre à profit pour aider la lecture humaine, lui suggérer des points d'appui, des pistes d'investigation, lorsque le volume des textes dépasse les capacités cognitives (Pincemin 2012b : 5, note 12).

Les divers outils et méthodes ainsi combinés concourent à objectiver le plus possible l'accès aux données, en délimitant la subjectivité de l'analyste – nécessaire et inévitable – par la multiplication des angles d'approche. L'exploration outillée est guidée tantôt par les outils tantôt par les interprétations et les intuitions formulées en fonction des résultats obtenus. Dans notre travail, l'analyse outillée du corpus a une triple fonction : une fonction d'objectivation de l'approche des données, une fonction heuristique, visant à faire émerger de nouvelles pistes d'investigation, et une fonction d'étayage des interprétations.

Compte tenu des critères d'exploitabilité du corpus ainsi que des caractéristiques discursives et génériques propres à ses sous-parties, chaque sous-corpus (PN et TWS) est exploré, dans un premier temps, de façon séparée eu égard à ses spécificités génériques et à son volume afin d'en saisir les traits distinctifs et les variations intragénériques. La démarche exploratoire reste cependant similaire et comparable. Un bilan contrastif entre les deux sous-corpus ainsi analysés est finalement proposé en tant que synthèse conclusive des variations intergénériques.

Analyse du corpus

partie II

Chapitre 4

Analyse du corpus du discours journalistique

À la lumière de la démarche combinée que l'on vient de décrire, le corpus PN est tout d'abord exploré au niveau global, dans sa structuration interne, dans le but d'en identifier la répartition thématique et les traits énonciatifs principaux. Il s'agit ainsi de faire ressortir les contrastes entre ses sous-parties et les niveaux d'hétérogénéité sélectionnés dans une perspective intragénérique. L'exploration inductive-quantifiée portant sur le niveau global du corpus est aussi susceptible de mettre en relief d'autres aspects significatifs qui n'avaient pas été envisagés en amont, favorisant l'émergence de nouvelles pistes de recherche à approfondir lors de l'analyse qualitative et contextualisée. Les principes de base des différentes fonctionnalités textométriques mobilisées au fil de l'étude seront brièvement décrits lors de la première utilisation de chacune.

4.1. Aborder le niveau global du corpus : hétérogénéité sous-générique et métaénonciatrice

Cette première phase d'analyse a pour objectif de faire ressortir les variations internes du corpus en exploitant les différents niveaux d'hétérogénéité matérialisés par sa structuration grâce aux métadonnées. Les traits caractéristiques identifiés lors de ces analyses inductives-quantifiées, réalisées au niveau global du corpus, seront ensuite mis en relation avec les analyses qualitatives effectuées sur des phénomènes linguistiques aux niveaux locaux.

4.1.1. Aperçus sémantique et énonciatif

L'exploration du niveau global du corpus s'intéresse à deux dimensions principales : la dimension sémantique, par l'entrée lexicale, et la dimension énonciative, notamment par le biais des pronoms personnels, sans toutefois se limiter à ceux-ci. Les pronoms personnels peuvent en effet être considérés comme des éléments d'accès privilégié à l'inscription de la subjectivité dans le discours (cf. Benveniste 1966), jouant ainsi un rôle significatif dans la polarisation énonciative-pragmatique. Les résultats des aperçus sémantique et énonciatif du corpus dans son ensemble sont en outre mis en relation avec sa structuration interne afin d'identifier les convergences et les divergences entre les sous-parties du corpus, tant au niveau métaénonciatif que pragma-énonciatif.

4.1.1.1. Thématiques principales

La première analyse effectuée vise à donner accès à une vue d'ensemble du corpus et notamment des différents aspects thématiques participant à la construction discursive de l'événement dans le discours journalistique. Il s'agit d'une classification hiérarchique descendante (CHD), à savoir une classification de type lexical qui se base sur la méthode Reinert (1990, 1993), effectuée grâce au logiciel *Iramuteq*¹ (Ratinaud 2014, 2018). Cette fonctionnalité permet de classer des segments de texte ou des textes entiers, en fonction des objectifs de l'étude, sur la base de la distribution du vocabulaire. Par la CHD, le corpus est en effet décomposé en classes qui regroupent des segments de texte ou des textes s'apparentant entre eux sur la base de calculs statistiques relatifs à la distribution des mots en contexte. Les différentes classes correspondent ainsi aux « mondes lexicaux » (Reinert 1993) du corpus. Comme l'explique Pincemin (2022 : 379), ces classes « peuvent être lues comme approchant des univers sémantiques, décrits par les mots caractéristiques que l'on peut automatiquement associer à chaque classe par un calcul de spécificités ou une mesure statistique telle que le chi-2 ». Les résultats de la classification sont ensuite illustrés sous la forme d'un dendrogramme, comme en donne un exemple la figure 6. Ce type de graphique rend compte également de la hiérarchisation des découpages successifs effectués sur les unités analysées.

La CHD réalisée sur le corpus PN met en évidence cinq mondes lexicaux principaux, soit les cinq thématiques dominantes ressortissant de la sélection d'articles relatifs à la médiatisation du cas de Vincent Lambert² (figure 6).

1. <http://www.iramuteq.org>.

2. Dans les sections relatives à l'analyse du corpus, on emploiera l'abréviation VL par souci de concision.

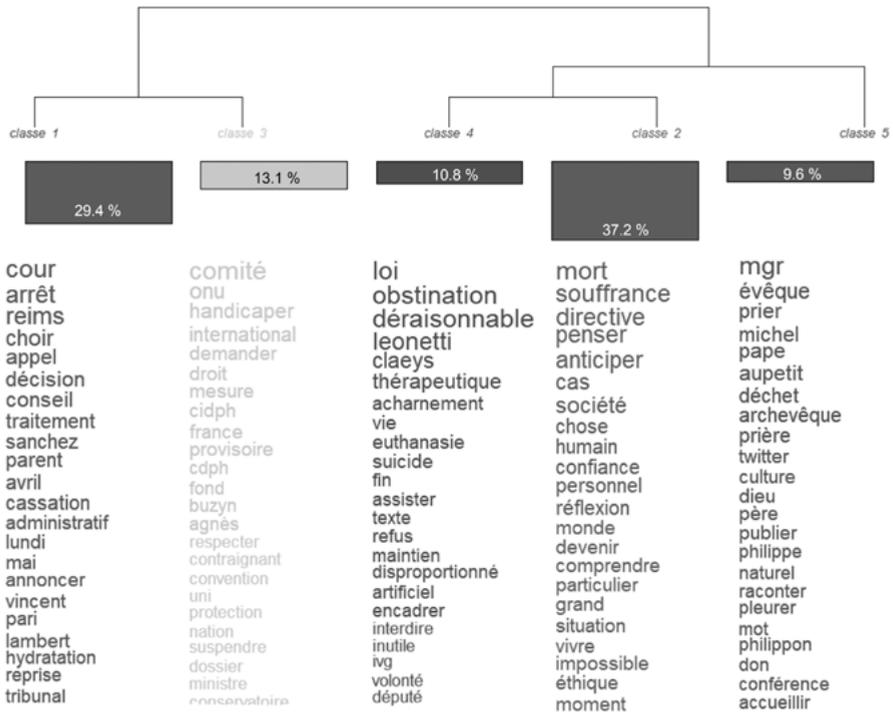


Figure 6. CHD dans le corpus d'étude PN.

La représentation hiérarchisée permet de dégager les relations entre les diverses classes, dans la mesure où chaque regroupement – c'est-à-dire chaque nœud de la structure arborescente du graphique – résulte des différences entre les classes ainsi séparées. En effet, comme le résumait Marchand et Ratinaud (2017 : 78), « on regroupe les profils semblables et on oppose les profils différents ». Par conséquent, les classes se différencient de plus en plus à chaque nœud du diagramme qui identifie un niveau hiérarchique nouveau. Au-dessous de chaque classe, on affiche en outre les formes les plus surreprésentées, c'est-à-dire les formes dont la fréquence dans la classe est supérieure à la fréquence attendue³.

À partir de la figure 6, on peut donc remarquer que les cinq classes de la CHD sur le corpus PN sont articulées selon trois nœuds majeurs : à gauche, le nœud regroupant les classes 1 et 3 ; au centre, le nœud avec les classes 2 et 4 ; ces dernières étant aussi en relation avec la classe 5, affichée à gauche du graphique. Du point de vue thématique, la partie gauche de la classification rassemble les segments de texte relatifs aux différentes démarches

3. Cf. plus loin le calcul des spécificités.

judiciaires de l'affaire VL, les procédures médicales et les acteurs principaux (classe 1 – *décision/arrêt de la Cour d'appel, le Conseil d'État, parents de Vincent Lambert, décision d'arrêt des traitements, hydratation et la nutrition*) en relation avec les références aux institutions engagées (classe 3 – *Comité des droits des personnes handicapées, ONU, France, Convention internationale relative aux droits des personnes handicapées*). Les classes de la partie droite relèvent en revanche des enjeux éthiques et sociaux de la fin de vie (classe 2 – la dimension *humaine* de la question, la *dignité* de la *mort* et les *souffrances* que la fin de vie peut entraîner, le *respect* de la *volonté* individuelle), de même que de l'encadrement législatif des pratiques qui peuvent être mises en œuvre en situation de fin de vie (classe 4 – *loi Claeys-Leonetti, acharnement thérapeutique, obstination déraisonnable, euthanasie, suicide assisté*). La classe 5 regroupe enfin les segments de texte incluant des références à une instance énonciative déterminée, l'institution religieuse, comme le montre notamment la présence de titres et de termes d'adresse religieux (*Mgr, évêque, prêtre, pape*) (classe 5).

L'analyse factorielle des correspondances (AFC) proposée par *Iramuteq* (figure 7) en relation avec la CHD aide ensuite à mieux illustrer la bipartition thématique principale du corpus. Cette analyse relève d'un usage particulier de l'AFC développée par Benzécri en 1969. Comme l'explicitent Leblanc *et al.* (2017 : 152), ce type d'analyse contribue à « identifier les faits saillants d'un corpus en termes de distribution du stock lexical. Il s'agit alors de réduire un grand nombre de variables à quelques facteurs structurants ». L'AFC s'appuie sur des tableaux statistiques réalisés à partir des matériaux linguistiques du corpus, et elle permet d'« extraire les faits saillants du tableau lexical en produisant une représentation graphique des profils lignes et colonnes de ce tableau » (Leblanc *et al.* 2017 : 153)⁴.

Dans le cas de l'AFC liée à la CHD sous *Iramuteq*, elle se limite en revanche à croiser entre elles les formes retenues par la CHD, offrant une représentation spatiale des relations entre les mots regroupés dans les diverses classes de la CHD.

L'opposition entre la partie gauche et la partie droite du graphique relatif à la dimension lexico-sémantique du corpus semblerait donc rendre compte, respectivement, du côté informatif-factuel, et du côté commentaire-opinion de la médiatisation du cas de VL dans le discours journalistique. Les mots affichés dans la partie gauche concernent les faits – notam-

4. Pour plus de détails sur les caractéristiques et le fonctionnement de l'analyse factorielle des correspondances, on renvoie à l'approfondissement proposé par Jean-Marc Leblanc (2017 : 218-228) dans l'ouvrage collectif consacré aux *Méthodes et outils informatiques pour l'analyse des discours*, sous la direction de Émilie Née.

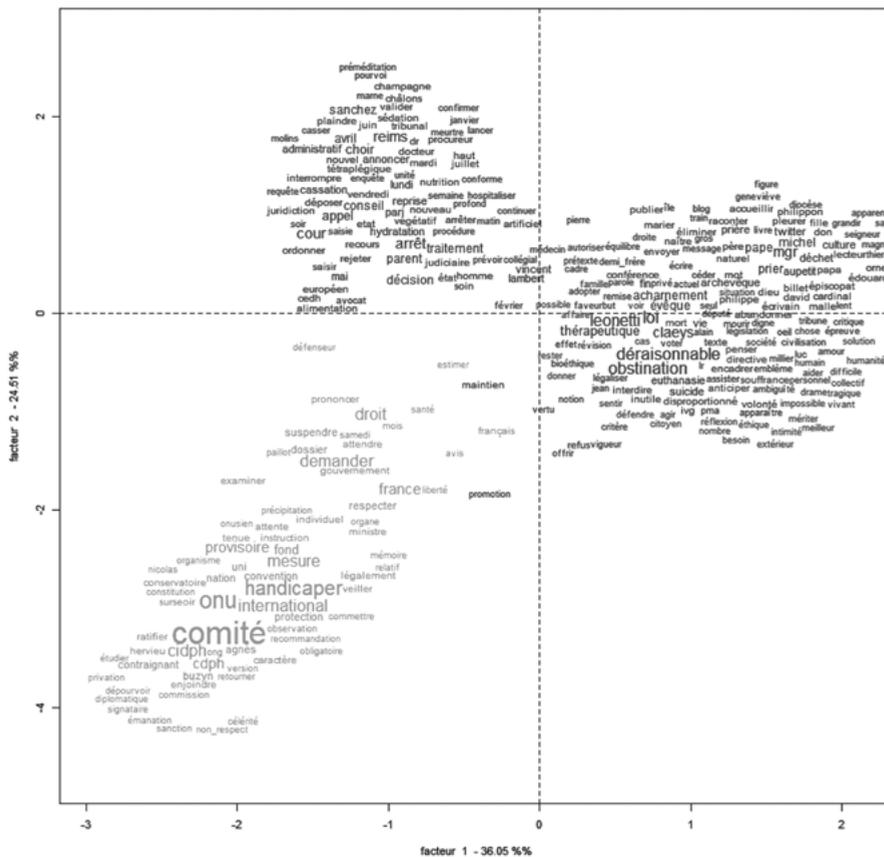


Figure 7. AFC des classes identifiées par la CHD.

ment judiciaires – qui ont scandé cette affaire, correspondant à la classe 1, dans le carré en haut à gauche, et à la classe 3, dans le carré inférieur gauche. La partie droite du graphique illustre pour sa part les mots employés notamment dans les segments de texte de commentaire ou d'analyse de l'événement, se concentrant dans les classes 2, 4 et 5 de la CHD⁵.

On remarque en outre que les deux mondes lexicaux de la partie gauche du graphique (classes 1 et 3) présentent un degré de dispersion marqué, alors que les trois mondes de la partie droite (classes 2, 4 et 5) se rapprochent plus les uns des autres. Il s'agit d'indications supplémentaires sur les relations entre les classes ainsi réparties. On peut donc supposer des divergences plus accentuées entre les segments de texte rapportant (au sens de Charaudeau

5. Le format noir et blanc de l'image ne permet toutefois pas de bien distinguer les différences de couleurs entre les classes 2, 4 et 5.

2011) l'événement, contrairement aux segments de texte commentant les faits et véhiculant des prises de position plus explicites.

À partir de ces résultats, le contraste plus fort au sein du corpus PN relèverait de la polarisation sur laquelle repose la classification sous-générique des textes du corpus, qui s'appuie néanmoins sur des critères énonciatifs et pragmatiques. Il est ainsi nécessaire de vérifier si cette opposition lexicosémantique est également confirmée au niveau énonciatif.

4.1.1.2. Dimension énonciative-pragmatique

Afin de tester si la dimension énonciative et pragmatique du corpus confirme l'interprétation de la bipartition principale de la CHD comme une opposition d'ordre sous-générique, notre analyse se concentre sur la distribution des pronoms personnels. On peut en effet considérer l'emploi de certains pronoms personnels plutôt que d'autres comme un élément significatif de la polarisation énonciative-pragmatique des deux sous-genres du discours journalistique. Comme l'a bien montré Benveniste, les pronoms personnels représentent un élément fondamental pour l'inscription de la subjectivité dans l'acte langagier : « Les pronoms personnels sont le premier point d'appui pour cette mise au jour de la subjective dans le langage » (Benveniste 1966 : 262).

Benveniste met toutefois en évidence une série de différences à l'intérieur du système des pronoms personnels eux-mêmes. Il est ainsi utile de rappeler quelques-unes des oppositions identifiées par le linguiste afin de montrer le rôle des pronoms personnels dans l'opposition objectivité-subjectivité sur le plan énonciatif. Une première opposition concerne les pronoms *je* et *tu*, d'une part, et le pronom de la troisième personne, de l'autre. Benveniste soutient en effet que seulement *je* et *tu* impliquent une « référence constante et nécessaire à l'instance de discours » (p. 253), c'est-à-dire l'acte d'énonciation spécifique « par [lequel] la langue est actualisée en parole par un locuteur » (p. 251). Les deux pronoms ne comportent donc pas de référence « constante et "objective" » (*ibid.*, p. 252), mais ils renvoient à un être animé humain, qui change en fonction de l'instance d'énonciation, et présupposent aussi une prédication sur celui-ci (p. 228). Au contraire, le pronom *il* constitue la « non-personne » :

la non-personne est le seul mode d'énonciation possible pour les instances de discours qui ne doivent pas renvoyer à elles-mêmes, mais qui prédisent le procès de n'importe qui ou n'importe quoi hormis l'instance même, ce n'importe qui ou n'importe quoi pouvant toujours être muni d'une référence objective (Benveniste 1966 : 255-256).

Quant à l'opposition singulier-pluriel relative aux pronoms personnels, Benveniste affirme qu'elle ne concerne effectivement que la troisième personne en raison du statut de la référence qui la caractérise. Il soutient en effet que les pronoms *nous* et *vous* renvoient plutôt à « une personne amplifiée et diffuse », ce qui ne constituerait donc pas une forme véritable de pluriel (Benveniste 1966 : 235).

Sur la base de ces considérations, l'étude statistique de l'emploi des pronoms personnels s'avère significative pour distinguer le ton énonciatif *objectivisé* ou *subjectivisé* (Moirand 2007) caractérisant les textes du corpus. Le calcul des spécificités (Lafon 1980) permet en effet d'observer la distribution des unités linguistiques à travers les diverses sous-parties du corpus, en favorisant une perspective comparative-contrastive. Ce calcul identifie les formes qui sont suremployées ou sous-employées dans chaque sous-partie, correspondant en l'occurrence aux cinq classes obtenues via la CHD illustrée dans la figure 6. Quatre paramètres sont pris en considération dans le calcul des spécificités : le nombre des occurrences du corpus, le nombre d'occurrences dans la partie, la fréquence de chaque unité dans le corpus, et la fréquence de chaque unité dans la partie, comme l'expliquent Leblanc *et al.* (2017). Dans le premier cas, la fréquence des unités dans la partie est supérieure à la fréquence théorique attendue (cf. par exemple les mots affichés sous les diverses classes de la CHD), dans le second, leur fréquence est en revanche inférieure. Cette fonctionnalité constitue donc « un moyen de caractérisation d'une partie d'un corpus » (Pincemin *et al.* 2010 : 344).

L'analyse des spécificités sur la distribution des pronoms personnels en fonction des cinq mondes lexicaux dégagés par la CHD (figure 8) semble confirmer la lecture interprétative de la CHD en termes d'une bipartition thématique d'ordre sous-générique selon la polarisation pragma-énonciative. Le calcul des spécificités montre à son tour des tendances divergentes entre les deux ramifications du graphique. Dans la ramification droite, la classe 2 se caractérise par la surreprésentation de la plupart des pronoms, signalant aussi bien une prise en charge à la première personne du singulier – le *nous* étant souvent employé comme complément – que l'emploi du *on* à valeur générique.

Cela étant, il est toutefois improbable que tous les segments de texte regroupés dans les classes 2 et 5 n'appartiennent qu'à des articles d'opinion, d'autant plus que ces derniers ne comptent que pour un tiers des articles totaux (cf. 3.1.3.2). Ce premier aperçu du corpus renforce par conséquent la nécessité de traiter séparément les articles qui ont un ton informatif-ex-

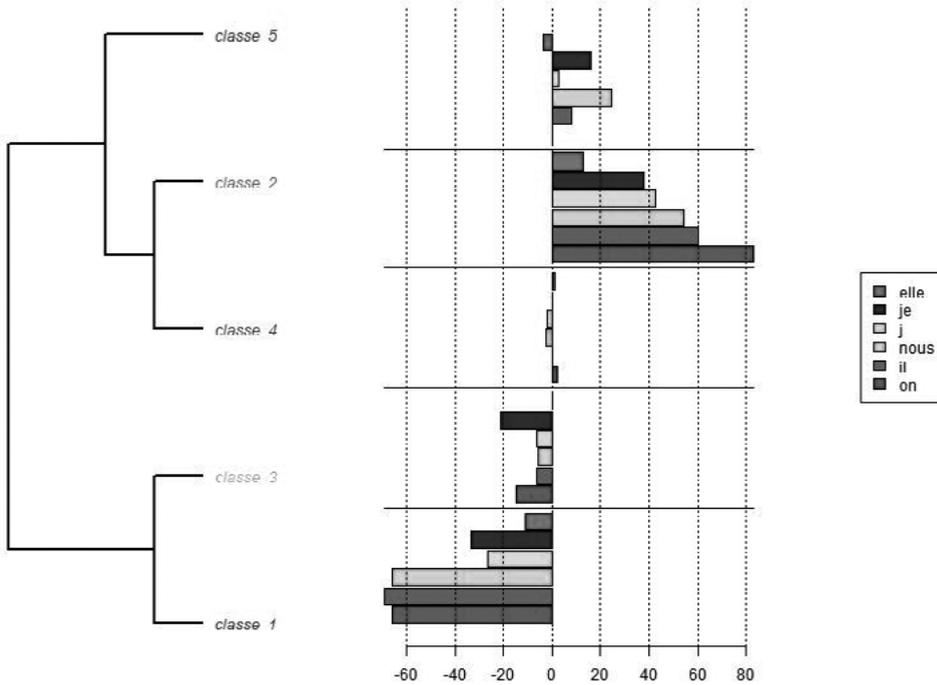


Figure 8. Spécificités sur les pronoms personnels par classe en fonction de la CHD.

plicatif (PN-INFO) et ceux qui commentent l'actualité ou présentent des opinions (PN-OP), en raison de leurs différences aussi bien au niveau lexical et sémantique qu'au niveau énonciatif-pragmatique, comme on l'avait prévu lors de la conception et structuration de ce corpus d'étude (cf. 3.1.2.2 et 3.1.3.2).

Étant donné la présence remarquable de segments de texte se rapprochant du pôle « implication-commentaire » (Adam 1997), en quantité supérieure à ce que l'on aurait pu attendre du fait de la structuration interne du corpus, il s'avère en outre intéressant d'observer comment de possibles segments aux traits plus subjectivisés (au sens de Moirand 2007) sont intégrés au sein des discours visant un effet d'objectivation. À ce propos, on peut émettre l'hypothèse que le recours à des modes autonimiques de représentation des propos d'autres énonciateurs contribue de manière significative à la présence de ces passages textuels aux traits subjectivisés.

4.1.1.3. L'hétérogénéité sous-générique et métaénonciatrice à l'épreuve de l'exploration outillée

Afin de vérifier et de creuser les divergences lexico-sémantiques entre les articles d'information et ceux d'opinion, une AFC a été effectuée sur les formes des sous-parties sous-génériques des quatre journaux. À la différence du cas précédent, l'AFC réalisée ici relève du modèle mis au point par Benzécri (1969, 1973). Comme le résume Pincemin, ce type d'analyse

opère une synthèse globale des relations entre mots (ou plus généralement traits linguistiques, motifs) et textes (ou plus généralement partie du corpus). Les mots sont comparés les uns aux autres sur la base des textes qui les emploient ; et réciproquement les ressemblances et écarts entre textes sont évalués par le vocabulaire qu'ils mobilisent (Pincemin 2022 : 378).

La représentation graphique associée à l'AFC permet en outre d'illustrer dans le plan cartésien « les rapprochements et oppositions les plus forts entre mots, entre textes, et indirectement entre mots et textes » (Pincemin 2022 : 378).

Comme le montre la figure 9, les sous-parties qui regroupent les articles d'information s'opposent de manière nette aux sous-parties rassemblant les articles d'opinion. Ceci justifie le choix d'aborder séparément les deux sous-genres d'articles, en tenant également compte des écarts au niveau sémantique.

Il faut également mettre en évidence que la disposition spatiale des sous-parties regroupant les articles d'information témoigne de divergences plus marquées du fait de leur distance plus élevée, alors que les sous-parties des articles d'opinion sont moins éloignées. On s'attend par conséquent à des différences moins accentuées ainsi qu'à quelques points de convergence entre les sous-parties les plus proches. Ce sera ensuite l'analyse déductive et contextualisée de phénomènes locaux qui clarifiera mieux les éléments linguistiques sur la base desquels se réalisent les écarts et les rapprochements aux niveaux de l'hétérogénéité sous-générique et métaénonciatrice, représentant deux des paramètres choisis pour étudier la variation intragénérique.

En bref, cette approche inductive du corpus a permis de vérifier la pertinence d'une hypothèse théorico-méthodologique formulée en phase d'élaboration et matérialisée ensuite grâce à un niveau d'annotation, tout en faisant émerger des pistes de réflexion et d'exploration.

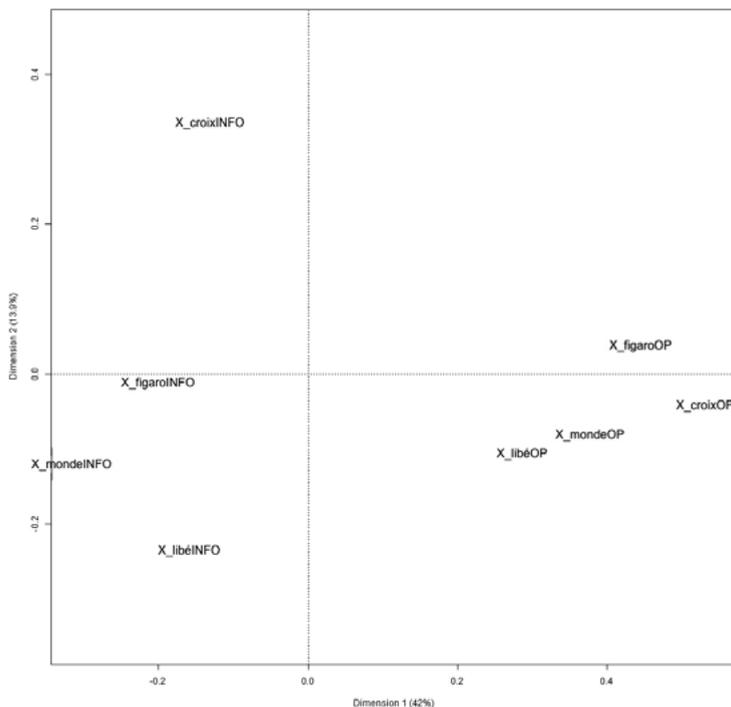


Figure 9. AFC sur les formes dans les sous-parties PN-INFO et PN-OP.

4.1.2. Variations intragénériques en fonction de la polarisation énonciative-pragmatique

Il s'agit à présent d'approfondir l'étude de la variation interne du corpus journalistique en fonction de l'hétérogénéité sous-générique. Deux angles d'approche complémentaires sont adoptés en raison de leur rôle fondamental dans la mise en discours d'un événement depuis une perspective déterminée : le niveau lexico-sémantique et le niveau énonciatif-pragmatique. Ces deux niveaux sont étroitement interreliés et participent des positionnements discursifs à l'égard dudit événement.

4.1.2.1. Variations lexico-sémantiques

Afin de saisir les traits lexicaux et sémantiques caractéristiques des deux sous-parties du corpus (PN-INFO et PN-OP) et les variations intragénériques qui en découlent, chacune a été soumise à son tour à une classification hiérarchique descendante ultérieure. Dans PN-INFO, la CHD (figure 10) met en

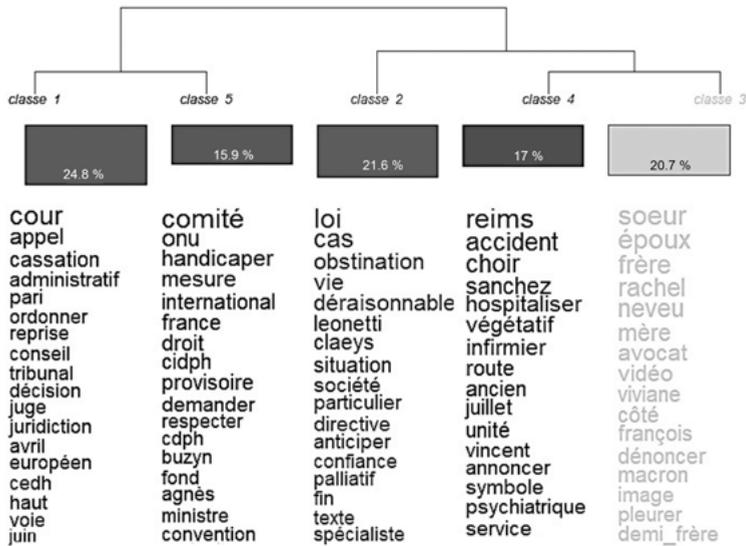


Figure 10. CHD dans le sous-corpus PN-INFO.

relief, dans ce cas aussi, cinq classes principales, relevant de trois questions majeures de nature judiciaire, législative et médicale, soulevées par le cas de VL.

Dans la ramification gauche, on repère les passages concernant les diverses procédures judiciaires engagées et les différentes institutions nationales et internationales sollicitées à se prononcer sur ce cas (classes 1 et 5). On retrouve ensuite, dans la classe 2, les passages relatifs aux lois en vigueur en matière de fin de vie (la loi Leonetti de 2005 et plus encore la loi Claeys-Leonetti de 2016), qui illustrent l'encadrement législatif de référence, en abordant en particulier les questions de la reconnaissance de l'obstination déraisonnable, des directives anticipées et des soins palliatifs. Comme on peut le voir dans les exemples suivants, un ton informatif-explicatif caractérise ces extraits qui rendent compte et clarifient des aspects du cadre législatif ayant un rôle clé dans cette situation spécifique.

- (1) Cette procédure est encadrée par la loi Claeys-Leonetti de 2016 [...]. Selon cette loi, les traitements peuvent être « suspendus » lorsqu'ils « apparaissent inutiles, disproportionnés ou lorsqu'ils n'ont d'autre effet que le seul maintien artificiel de la vie » (« Le Figaro », 20/05/2019, actualitésociété, Figaro, PN-INFO)⁶.

6. Une sélection de métadonnées pertinentes pour l'analyse est fournie entre parenthèses pour chaque extrait, selon l'ordre suivant : source énonciative (nom du journal), date de l'article, rubrique du journal, auteur de l'article, sous-partie sous-générique du corpus (PN-INFO/PN-OP).

- (2) Au titre de la non-obstination déraisonnable, la loi Claey's-Leonetti du 2 février 2016 prévoit un arrêt de tout traitement de maintien artificiel en vie : [...] mais aussi traitement de survie comme la nutrition et l'hydratation artificielles (« Le Monde », 20/05/2019, société, Béguin, PN-INFO).

L'exemple 1 précise l'encadrement législatif de référence pour la procédure mentionnée et en rapporte des courts extraits entre guillemets. L'exemple 2 signale également le texte législatif de référence, en résumant ses dispositions au sujet de la non-obstination déraisonnable. Les deux cas présentent une énonciation objectivée à la troisième personne et au présent de l'indicatif, ce qui confère une dimension informationnelle et explicative. On remarque toutefois des différences dans le choix du mode de représentation du discours législatif.

Une relation plus étroite caractérise les classes 3 et 4, qui portent respectivement sur le conflit familial et sur différents aspects liés au domaine médical. Les références au CHU⁷ de Reims où VL est hospitalisé, au médecin qui l'a en charge, mais aussi à son passé d'infirmier et à la cause de son état actuel (*accident*) sont en effet regroupées dans cette classe. On pourrait supposer que le lien entre ces thématiques procède du fait que les diverses prises de position des membres de la famille de VL sont construites en s'alignant ou en s'opposant aux décisions prises par l'équipe médicale du CHU.

En ce qui concerne en revanche les articles d'opinion (figure 11), la CHD dégage une variété plus large de sujets, distinguant sept classes thématiques. Dans ces discours, le cas de VL est abordé et commenté en ouvrant aussi la réflexion sur d'autres questions associées.

Si la structuration hiérarchique du graphique favorise l'identification de plusieurs articulations thématiques, l'analyse contextuelle des segments de texte ainsi classifiés permet en outre de remarquer des différences intragénériques. En effet, bien que le niveau hiérarchiquement supérieur du dendrogramme (classes 6 et 7) présente des affinités avec les classes du sous-corpus PN-INFO concernant les aspects les plus factuels de l'affaire, les questions législatives et les aspects judiciaires sont ici expliqués, discutés et évalués. Ils font l'objet de prises de position à travers l'emploi de ressources lexicales

7. Dans les articles, l'acronyme CHU est généralement utilisé pour faire référence à l'hôpital de Reims. Cependant, le logiciel *Iramuteq* lemmatise ce terme en le considérant comme une forme fléchi du verbe *choir*. D'où la surreprésentation de ce verbe dans la classe 4. Un retour au contexte a pourtant permis de dégager cette catégorisation du traitement automatisé qui ne correspond pas à l'emploi contextuel de la forme.

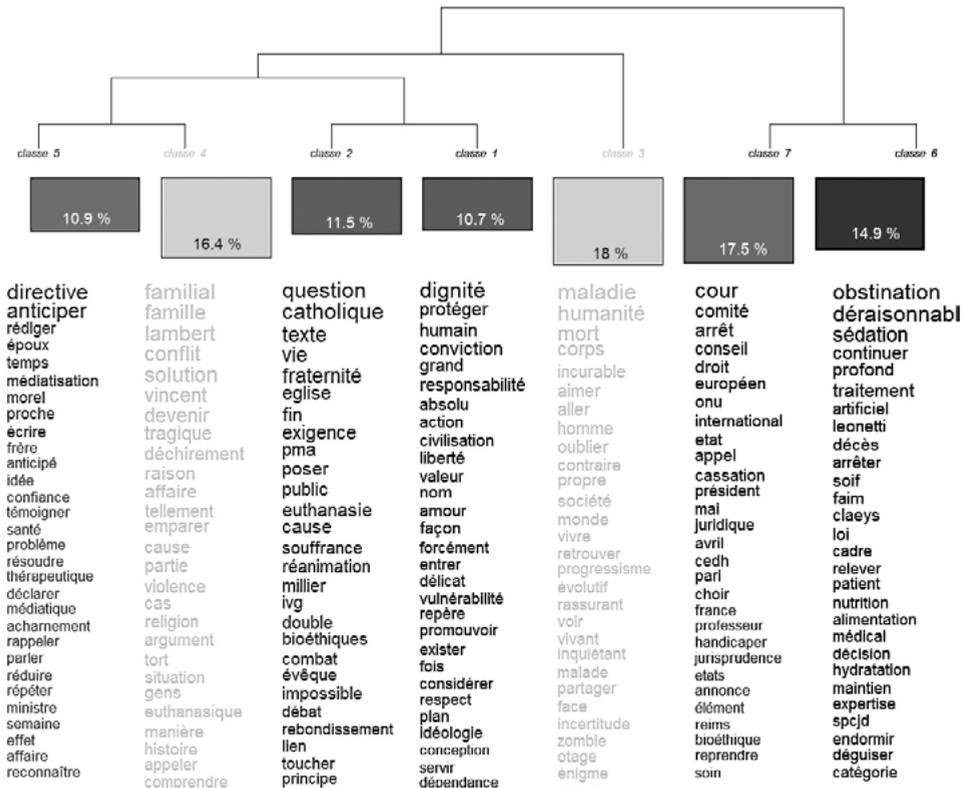


Figure 11. CHD dans le sous-corpus PN-OP.

ouvertement axiologiques (*scandale, absurdité, inacceptable*) ou bien de procédés argumentatifs d'ordre dialogique, à l'instar de la négation polémique (*elle n'a pas pour objet le handicap*) ou des questions rhétoriques (*Mais la nutrition et l'hydratation sont-elles vraiment des « traitements » ?*), comme on peut en observer des exemples dans les extraits ci-dessous⁸.

- (3) [...] le geste qui mettrait actuellement fin à sa vie **ne pourrait qu'être interprété comme un geste euthanasique provoquant le scandale** dans une partie de la famille et auprès de beaucoup d'autres personnes [...]. (« La Croix », 13/05/2019, blog, Saintôt, PN-OP)
- (4) Était-on dans une « obstination déraisonnable » autorisant, selon les termes de la loi, qu'on mette fin aux traitements ? **Mais la nutrition et**

8. L'emploi du gras dans les exemples est de notre fait. Il vise à mettre en relief les éléments principaux sur lesquels porte l'analyse.

l'hydratation sont-elles vraiment des « traitements » ? (« Le Figaro », 11/07/2019, éditorial, de Montety, PN-OP).

- (5) La loi française, **je le redis**, est une loi sur l'obstination déraisonnable et la volonté du patient, **elle n'a pas pour objet le handicap**. [...] **C'est évidemment inacceptable** (« Libération », 21/05/2019, interview, E1 : Favereau, e2 : Denis Berthiau, PN-OP).

Les exemples 3-5 se distinguent de manière significative des exemples 1 et 2. Il ne s'agit pas ici d'apporter des compléments d'information ou des précisions tirées du texte institutionnel (ex. 1-2), mais plutôt d'exprimer un point de vue ou d'offrir une interprétation déterminée. Selon les cas, la perspective de l'énonciateur se traduit soit par une mise en discours orientée, qui se rapprocherait de l'un des deux positionnements principaux de la controverse, soit par une prise de position assumée. Dans l'exemple 3, l'énonciateur propose une catégorisation en termes de « geste euthanasique », ce qui implique une prise de position et une évaluation. L'emploi du conditionnel modalise l'affirmation, en la présentant comme une possibilité. Celle-ci est toutefois renforcée par la négation restrictive qui circonscrit les interprétations possibles du geste mentionné. L'exemple 4 se caractérise en revanche par la présence de mots entre guillemets qui marquent une forme de *modalisation autonymique d'emprunt* (Authier-Revuz 2020). En l'occurrence, l'énonciateur remet en question les catégorisations proposées, en (se) demandant si elles ont été effectuées correctement. Ces questions ne sont toutefois qu'apparentes car l'énonciateur semble suggérer la bonne réponse à travers la structuration même de ces interrogations totales, leur enchaînement et notamment l'emploi du connecteur d'opposition *mais* et de l'adverbe *vraiment*. Enfin, l'extrait 5 illustre un exemple de prise en charge énonciative à la première personne du singulier. Dans le cadre de cet entretien, l'interviewé, Denis Berthiau, tient à préciser les sujets qui sont règlementés par la loi et ceux qui en sont exclus. L'emploi du présent de l'indicatif marque le degré de certitude de l'affirmation, que l'énonciateur assume ouvertement en signalant la réitération du propos de sa part. Il exprime ensuite sa position contre la mise en relation du handicap avec l'obstination déraisonnable, en la jugeant comme « évidemment inacceptable ».

En descendant la hiérarchie de la CHD, on repère ensuite plusieurs sujets de discussion abordés en relation avec le cas de VL, tels que la place et la conception de la mort dans la société contemporaine, ainsi que celles du corps, de la maladie et de la souffrance (classe 3), alors que le dernier niveau

de la classification (classes 1, 2, 4 et 5) relève des arguments principaux de la polarisation liée à la controverse.

Le conflit familial dont le cas de VL a fait l'objet est traité en particulier dans les segments de la classe 4, en relation avec ses répercussions sur l'opinion publique et la société française. Il en ressort la valeur symbolique de ce cas d'actualité devenu une *cause*, contribuant à relancer le débat public sur la fin de vie et à s'interroger sur la législation actuelle et notamment sur la question des directives anticipées (classe 5). La médiatisation de cette « affaire » devient ainsi exemplaire pour une réflexion collective. Les dimensions tragique et intime de ce cas ainsi que les enjeux sociaux, éthiques et politiques liés au débat public qu'il a relancé sont également soulignés.

Les segments de texte rassemblés dans la classe 1 traitent en revanche de la question de la dignité humaine et en particulier du respect pour la fragilité et la vulnérabilité des malades en état de dépendance, de la conception de la vie, du respect de la volonté du patient ainsi que de la liberté de pouvoir choisir sa propre fin de vie. La classe 2 regroupe enfin des segments qui abordent d'autres sujets aux enjeux éthiques tels que l'euthanasie, l'interruption volontaire de grossesse (IVG), la procréation médicalement assistée (PMA), en les rapportant à la position que le monde catholique pourrait assumer à leur égard.

Pour résumer, l'analyse comparative-contrastive des mondes lexicaux des deux sous-parties de PN a mis en relief des divergences intragénériques au niveau lexico-sémantique, concernant non seulement la variété des sujets abordés, mais aussi la façon de focaliser l'attention sur des aspects différents de l'événement et de la controverse associée. On peut donc s'attendre que ces différences soient encore plus marquées sur le plan énonciatif-pragmatique, en raison des traits distinctifs des deux sous-genres, comme le suggère aussi l'analyse contextualisée au niveau mésotextuel par les quelques exemples illustrés.

4.1.2.2. Variations énonciatives-pragmatiques

Dans le but d'analyser les divergences entre PN-INFO et PN-OP au niveau énonciatif-pragmatique, une AFC a été réalisée sur les catégories grammaticales, c'est-à-dire sur les « types » dans les options de *Iramuteq* (Ratinaud 2014). Cela confirme une opposition plutôt nette entre les deux sous-parties internes des quatre journaux, qui se concentrent, respectivement, dans la moitié gauche et dans la moitié droite du plan cartésien (figure 12).

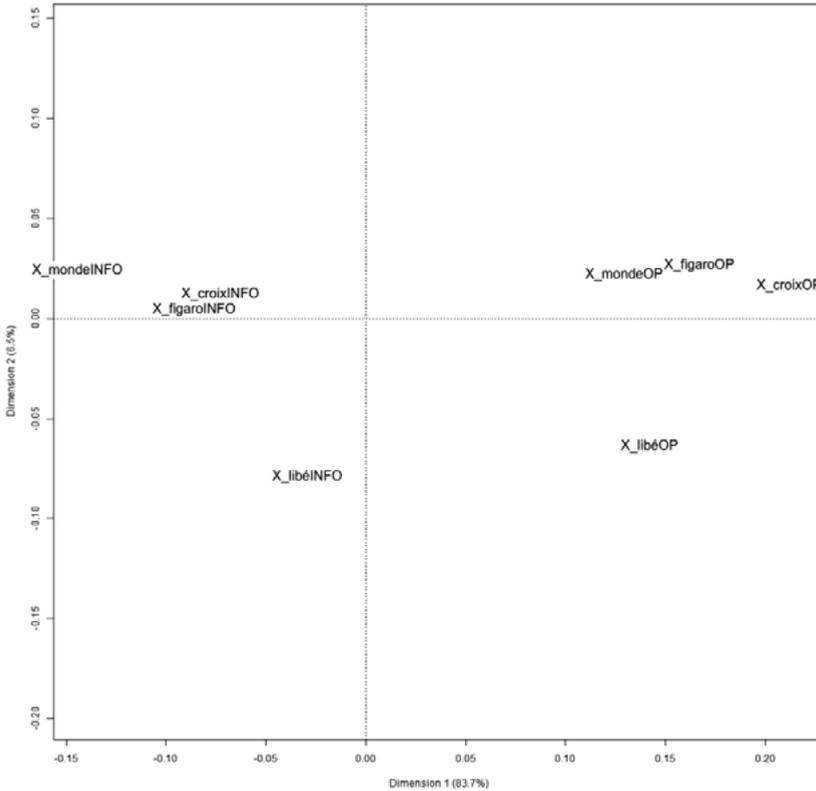


Figure 12. AFC sur les types des sous-parties PN-INFO et PN-OP.

À la différence de l'AFC sur les formes (figure 9), les sous-parties des articles d'information ne sont pas aussi éloignées les unes des autres. Cela indique des emplois moins divergents au niveau des catégories grammaticales, dont les usages diffèrent toutefois de manière remarquable par rapport aux articles d'opinion, ce qui confirme une plus forte opposition au niveau sous-générique. On constate néanmoins que les deux sous-parties rassemblant les articles de « Libération » (libéINFO et libéOP) s'éloignent des sous-parties correspondantes des autres quotidiens, tout en gardant entre elles l'opposition principale.

L'analyse des spécificités (figure 13), effectuée sous TXM⁹ (Heiden *et al.* 2010), permet de clarifier quelques éléments comme étant responsables des divergences énonciatives-pragmatiques entre les sous-parties des quatre journaux. Elle montre que, globalement, les articles d'opinion recourent da-

9. <https://txm.gitpages.huma-num.fr/textometrie/>.

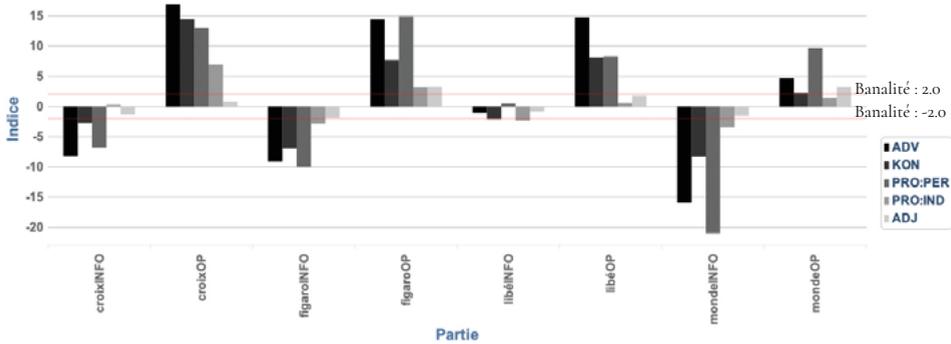


Figure 13. Spécificités sur les catégories grammaticales dans les sous-parties de PN.

vantage aux pronoms personnels (PRO:PER) pour marquer la subjectivité en tant qu'instance énonciative, aux pronoms indéfinis (PRO:IND) pour renvoyer de manière générique et vague à des référents notamment personnels, à des éléments à valeur argumentative et modalisante tels que les adverbes (ADV) et les conjonctions (KON), ainsi qu'à des catégories plus axiologiques tels que les adjectifs (ADJ).

Sur la base de ces résultats, il est possible d'avancer des hypothèses sur le style général des journaux examinés et, en conséquence, sur les variations intragénériques en fonction des quatre métaénonciateurs. La tendance à garder un style énonciatif moins subjectivisé (Moirand 2007) est plus marquée dans les articles d'opinion du « Monde », où les cinq catégories analysées sont moins surreprésentées que dans les sous-parties correspondantes des autres quotidiens. De plus, elles sont largement sous-représentées dans les articles d'information du « Monde », signalant ainsi un style qui évite autant que possible des formes subjectives et/ou argumentatives.

Les articles d'opinion de « La Croix » suremployent en revanche plus que les autres les catégories grammaticales envisagées. Ces dernières sont toutefois assez limitées dans la sous-partie informative de ce journal.

En ce qui concerne enfin « Libération » et « Le Figaro », on remarque des emplois assez similaires dans les articles d'opinion, alors que des tendances différentes se dégagent dans les articles d'information. « Le Figaro » tend à sous-employer les cinq catégories grammaticales analysées par rapport à ses articles d'opinion ainsi qu'à ceux de « La Croix » et de « Libération ». Ce dernier les utilise en revanche plus fréquemment que les autres quotidiens dans ses articles à visée informative, qui présentent notamment un large recours aux pronoms personnels.

On peut ainsi conclure que l'utilisation de méthodes contrastives telles que l'AFC et les spécificités, réalisées sur la base de deux critères d'hétérogénéité – l'hétérogénéité sous-générique (PN-INFO/PN-OP) et l'hétérogénéité métaénonciatrice (entre les quatre journaux) – permet non seulement de vérifier les oppositions supposées et de saisir des divergences et/ou des affinités nouvelles, mais aussi d'identifier les éléments linguistiques qui sont à l'origine des variations.

4.1.3. *Variations intragénériques en fonction de phénomènes autonymiques*

Dans le discours journalistique, il arrive souvent que les propos d'autres énonciateurs, qui ne coïncident pas avec l'auteur de l'article, y soient rapportés. C'est également le cas pour les articles à visée informative qui, tout en se caractérisant par un style énonciatif objectivisé (Moirand 2007), peuvent inclure aussi les propos d'autres énonciateurs que le journaliste. Ces propos peuvent être englobés au sein de l'énonciation du journaliste à travers divers « modes de représentation de discours autre », pour reprendre la terminologie de Authier-Revuz (2020). Suivant la proposition de l'auteure, il est possible de distinguer des formes de « (re)formulation paraphrastique » (Authier-Revuz 2020 : 226sq.), qui visent à rendre compte du contenu de l'énoncé à représenter, d'une part, des « modes autonymiques de la représentation de discours autre », qui comportent « le plan d'une matérialité signifiante "mentionnée" » (Authier-Revuz 2020 : 311sq), de l'autre.

L'intégration des propos d'autrui au sein de l'énonciation du journaliste et notamment l'emploi de modes autonymiques peuvent avoir des répercussions significatives sur la portée pragmatique et argumentative du discours (cf. 3.1.2.2), jouant également un rôle dans la construction discursive de l'événement et, en conséquence, des positionnements. Le deuxième paramètre pris en considération pour analyser les variations intragénériques concerne donc la présence, les caractéristiques linguistiques et la dimension pragmatique de mots et de segments textuels entre guillemets.

Par souci de clarté et de concision, nous distinguons les mots et/ou les segments textuels étant placés entre guillemets comme des « discours autonymiques » et, inversement, les mots et/ou les segments textuels n'étant pas entre guillemets comme des « discours non autonymiques ». En l'occurrence, le mot *discours* relève d'un emploi singulier et moins spécifique par

rapport au concept linguistique élaboré en ADF¹⁰. Cette classification rend en effet compte de fragments textuels de taille variée qui participent d'une production discursive déterminée, étant appréhendée à partir de la forme textuelle de l'article de journal. La distinction entre *discours autonomiques* et *discours non autonomiques* matérialise donc un angle d'exploration du corpus qui transcende l'unité textuelle. Elle donne accès à des fragments de discours qui soit correspondent à l'énonciation première de l'article, soit font l'objet d'une rupture énonciative, marquée graphiquement par l'emploi des guillemets. Les guillemets constituent en effet l'un des moyens pour signaler différents modes autonomiques de la représentation de discours autre, dont le discours direct et la modalisation autonome d'emprunt, en suivant encore Authier-Revuz (2020). Comme le remarque l'auteure, les guillemets peuvent marquer des phénomènes distincts qui partagent toutefois le trait de l'autonomie : « l'on retrouve la possibilité de conjointre entre guillemets des éléments hétérogènes [...] sur la base de leur commune dimension autonome » (Authier-Revuz 2020 : 310).

Si la bipartition entre *discours autonomiques* et *discours non autonomiques* répond à une exigence méthodologique (à savoir la nécessité d'automatiser l'identification d'une partie des phénomènes signalant une hétérogénéité énonciative), elle repose néanmoins sur un critère permettant de discriminer efficacement entre ce qui relève de l'autonomie et ce qui n'en relève pas. Ce sera ensuite le retour au texte et l'analyse en contexte qui permettront de distinguer les divers modes autonomiques, leurs fonctions pragmatiques ainsi que leurs effets argumentatifs.

À travers cette opposition matérialisée par les métadonnées, nous avons calculé le rapport entre le nombre de mots des segments textuels (mots, syntagmes, phrases) qui constituent des phénomènes autonomiques signalés par les guillemets (DA – discours autonomiques) et ceux qui n'en incluent pas (DNA – discours non autonomiques) à l'intérieur de chaque sous-partie des quatre journaux (cf. 3.1.3.2). De cette façon, on peut vérifier le poids quantitatif des phénomènes autonomiques selon les deux partitions sous-génériques.

À partir des graphiques illustrés dans la figure 14, on remarque qu'environ un tiers des mots constituant les articles d'information font l'objet, en moyenne, de modes autonomiques de représentation des propos d'autrui marqués par les guillemets. Au contraire, le recours à ces formes de repré-

10. Pour une définition du *discours* depuis la perspective de l'ADF, on renvoie à Charaudeau & Maingueneau (2002) ainsi qu'à Maingueneau (2014a).

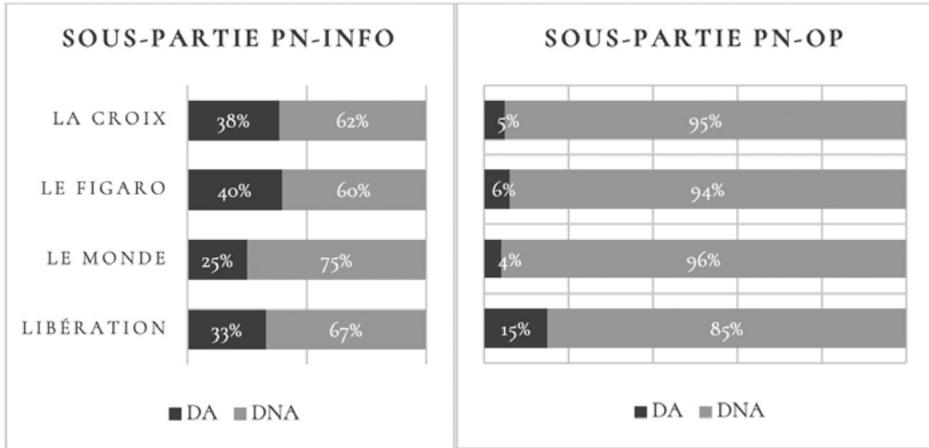


Figure 14. Répartition des discours autonomiques et des discours non autonomiques en fonction du nombre de mots.

sentation est beaucoup moins fréquent dans les articles d'opinion. Cela n'est du reste pas surprenant dans la mesure où, dans PN-OP, l'énonciateur premier, auteur de l'article, peut soit exprimer son propre avis (comme c'est le cas d'un éditorial), soit céder la parole à l'énonciateur second avec lequel il interagit directement (à l'instar, par exemple, d'une interview). Une utilisation plus limitée des formes autonomiques de représentation caractérise donc ces discours qui sont centrés en particulier sur une seule instance énonciative, que ce soit l'éditorialiste, l'auteur d'une tribune ou l'interviewé. Par contre, dans les articles d'information, la construction de l'événement repose davantage sur des confrontations discursives relevant d'une pluralité d'instances énonciatives, dont les propos peuvent faire l'objet de plusieurs formes de représentation.

Si l'on peut rattacher les proportions différentes à l'intérieur des deux sous-parties aux finalités pragmatiques et aux styles discursifs divers qui les caractérisent, le pourcentage moyen de la présence de représentations autonomiques entre guillemets dans les articles à visée informative nous paraît un aspect saillant à approfondir.

En fonction de l'hétérogénéité métaénonciatrice dans le sous-corpus PN-INFO, on peut constater, à partir du graphique de la figure 14, l'emploi privilégié des modes autonomiques de représentation par « La Croix » et « Le Figaro », où cette pratique concerne environ 40 % du nombre total des mots, alors que « Le Monde » l'utilise moins que les autres quotidiens. Compte tenu des implications argumentatives liées aux représentations autonomiques des discours cités, cette tendance minoritaire dans « Le

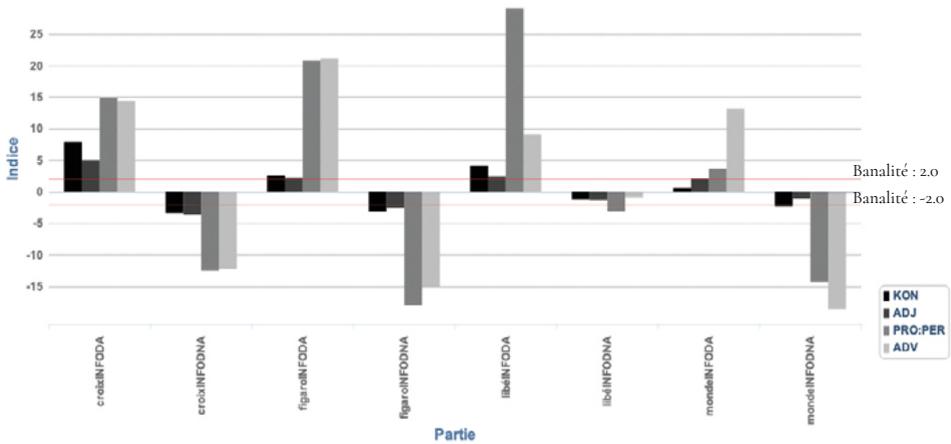


Figure 15. Spécificités sur les catégories grammaticales dans PN-INFO selon la distinction entre discours autonomiques et discours non autonomiques.

Monde » pourrait être interprétée comme l'indice d'un degré plus élevé d'objectivité énonciative, qui résulte de la prédominance d'un ton principalement narratif-descriptif. À l'inverse, une plus grande présence de segments entre guillemets pourrait favoriser un ton plus dramatique, notamment dans le cas de propos d'instances d'énonciation personnelles¹¹, ce qui se rapproche davantage du style des articles d'opinion.

Afin de vérifier l'interprétation de ces résultats quantitatifs du point de vue linguistique, l'analyse des spécificités peut clarifier la distribution de quelques catégories grammaticales significatives : conjonctions, adjectifs, pronoms personnels et adverbes. En effet, ces catégories participent à la construction argumentative du discours, ainsi qu'à la mise en évidence ou à l'effacement de la subjectivité du sujet-énonciateur. Le calcul des spécificités sur ces catégories (figure 15, ci-dessus) montre en effet que les segments entre guillemets dans les articles d'information tendent à suremployer les mêmes catégories que les articles d'opinion, alors qu'elles sont sous-représentées dans les discours non autonomiques.

Les discours autonomiques signalés par les guillemets représentent un lieu d'altérité énonciative. Celle-ci est exhibée en mettant en avant la subjectivité qui en est à l'origine et dont les traces ressortent des choix énonciatifs attribués par le journaliste à un autre énonciateur. L'alternance entre segments

11. La citation d'extraits de textes institutionnels, caractérisés par une énonciation impersonnelle, n'aurait pas les mêmes répercussions. Il s'agit généralement d'une citation d'autorité qui peut servir à renforcer le propos, mais elle peut également faire l'objet de critiques.

textuels non autonymiques et autonymiques relève donc d'une alternance entre l'énonciation plus objectivée de l'énonciateur principal de l'article (le journaliste) et l'énonciation plus subjectivée de l'énonciateur second. Cela concerne aussi bien la prise en charge énonciative que la validation modale par le biais de formes adverbiales et adjectivales. D'où l'importance de vérifier d'éventuelles différences internes entre les discours autonymiques et non autonymiques, notamment au sein des articles d'information (cf. 4.2.2).

4.2. Analyse déductive et contextualisée des niveaux locaux : hétérogénéité métaénonciatrice dans PN-INFO

La deuxième phase de la démarche d'analyse se veut plus déductive-contextualisée, dans le but d'approfondir des aspects saillants repérés lors de la première phase. Le corpus est ainsi exploré à différents niveaux locaux afin de vérifier les hypothèses et les interprétations formulées lors de l'étude inductive, mais aussi pour procéder à une caractérisation plus fine des phénomènes identifiés.

L'analyse des niveaux locaux vise la comparaison contrastive de la construction discursive de deux des objets de discours principaux dont relève la polarisation de la controverse. Par *objet de discours*, nous entendons un référent discursif d'ordre thématique en relation étroite avec le sujet controversé sur lequel portent les productions rassemblées en corpus¹². Plus précisément, l'analyse concerne les constructions discursives¹³, qui résultent des choix de référenciation et des procédés de mise en discours, de deux objets de discours parmi les plus abordés et les plus controversés dans le débat. Il faut en effet souligner qu'un même objet de discours peut être appréhendé et mis en discours par diverses ressources linguistiques et par plusieurs procédés discursifs qui véhiculent un point de vue spécifique sur

12. L'expression « objet de discours » est généralement utilisée dans les études discursives pour référer à ce dont on parle dans le discours envisagé ; cf. par exemple Charaudeau & Maingueneau (2002). Cependant, des concepts spécifiques ont également été théorisés en termes d'*objet de discours* chez Sitri (1996) et d'*objet discursif* chez Longhi (2015, 2018). D'une part, notre usage du terme « objet de discours » ne recoupe pas ces concepts, bien qu'il puisse y être mis en relation. De l'autre, il ne se rapproche qu'en partie de l'usage de « objet-de-discours » fait par Rabatel (2021). Dans notre étude, l'opérationnalité du terme « objet de discours » (sans tirets) est déterminée à la fois du point de vue thématique (les différents objets de discours constituent des référents thématiques qui font l'objet de débat) et du contexte communicationnel envisagé (le débat autour d'un sujet controversé).

13. Compte tenu des objectifs de l'ouvrage, le cadre épistémologique et théorique sur lequel se fonde notre conception de l'activité discursive et des processus de construction du sens ne sera pas présenté ici. Pour plus de détails à ce propos, on renvoie à Cagninelli (2022).

l'objet, participant ainsi à la construction des positionnements discursifs dominants. On peut donc envisager qu'un même objet de discours – en tant que référent discursif de nature thématique – est susceptible de donner lieu à une pluralité d'« objets discursifs » (Longhi 2015, 2018), condensant des appréhensions phénoménologiques différentes de la question. Celles-ci sont négociées et se stabilisent au niveau interdiscursif – à la fois intra- et intergénérique – participant de positionnements discursifs différents.

L'analyse des constructions discursives des objets de discours sélectionnés est réalisée sur deux niveaux d'hétérogénéité intragénérique. Tout d'abord, nous examinerons les variations liées à des instances journalistiques différentes, que nous avons proposé de nommer « hétérogénéité métaénonciatrice » (cf. 3.1.3.1) à la suite du concept de *métaénonciateur* (Maingueneau 2021). Ce niveau d'hétérogénéité permettra donc de faire ressortir par contraste des constantes propres au discours journalistique, tout comme les traits spécifiques à chaque journal dans la construction de ces objets. Dans un premier temps, l'étude sera donc conduite principalement sur les articles d'information puisqu'ils sont censés relever davantage de la ligne éditoriale générale du journal.

Le deuxième niveau d'hétérogénéité analysé concernera en revanche les variations sous-génériques selon la polarisation énonciative et pragmatique à l'origine de la distinction entre les articles d'information et ceux d'opinion. Ces deux sous-corpus seront soumis à des analyses complémentaires en fonction de leurs différences pragma-énonciatives. Les articles d'opinion seront explorés dans un deuxième temps, en focalisant l'attention notamment sur les instances énonciatives auxquelles chaque journal donne la possibilité de s'exprimer sur la question ainsi que leur alignement (cf. 4.3.2).

Parmi les objets de discours qui jouent un rôle remarquable dans la construction discursive de la controverse associée au cas de Vincent Lambert et des positionnements discursifs polarisés qui en découlent, deux sont approfondis à travers une analyse qualitative-déductive : les choix de référencement pour désigner l'individu qui est au centre de la polémique et son état de santé, ainsi que la représentation de la mort en relation avec la fin de vie ou avec son cas spécifique.

Comme nous l'avons anticipé plus haut, l'analyse déductive-contextualisée porte d'abord sur les formes de référencement privilégiées par les divers journaux pour construire ces objets de discours, pour explorer ensuite en contexte les différentes façons dont ils sont mis en discours. Pour ce faire, nous aurons recours initialement à l'analyse des spécificités pour comparer les fréquences des différentes formes de référencement selon les journaux.

Ensuite, l'analyse des concordances et des cooccurrences permettra d'en saisir les contours référentiels et d'en identifier les constructions discursives et les polarisations.

L'étude des concordances et des cooccurrences permet en effet deux types d'observation contextuelle des formes et des procédés saillants. D'une part, les concordances assurent la visualisation de chaque occurrence d'une forme ou d'une expression dans son environnement contextuel immédiat. Elles donnent ainsi lieu à « une présentation synthétique et structurée » des données et favorisent le repérage « des récurrences de constructions ou de contextes locaux, en gardant le lien avec les contextes globaux » (Mayaffre *et al.* 2019 : 106). D'autre part, le calcul des cooccurrences se fonde sur l'« association statistiquement significative de deux unités linguistiques (en général deux mots) dans une fenêtre déterminée (en général le paragraphe) » (Mayaffre *et al.* 2019 : 107). Nous suivons ainsi la proposition de Mayaffre (2008 : 60) de considérer la cooccurrence comme la « forme minimale du contexte ».

À travers ces analyses donnant accès à différents niveaux locaux du corpus, il sera possible de mieux comprendre la dimension micro- et méso-textuelle des constructions discursives principales mises en relief au niveau statistique. En effet, comme le soulignent Mayaffre *et al.* (2019 : 106), l'« interprétation des décomptes doit être éclairée par l'observation en contexte des faits comptabilisés ».

4.2.1. *Construction des objets de discours : formes de référenciation*

Les formes de référenciation employées pour désigner VL participent à la construction de l'un des objets de discours principaux de la controverse. Leur analyse s'avère ainsi significative afin de saisir la mise en perspective de l'événement et, en conséquence, le positionnement discursif principal auquel chaque journal peut être apparenté d'un point de vue linguistico-discursif.

Comme le résume Sarrazin (2017 : 330) dans *Termes et concepts pour l'analyse du discours*, la référenciation indique l'« [a]cte qui consiste, pour un énonciateur, à désigner un référent à travers l'actualisation d'une séquence linguistique, et [le] résultat de cet acte ». Suivant Rabatel (2021), nous envisageons l'acte de référenciation comme intrinsèque à l'énonciation elle-même, contribuant à la construction d'une représentation linguistique du référent extralinguistique à partir d'une perspective spécifique. Il s'agit donc d'une représentation discursive qui repose également sur le

rapport expérientiel du sujet de l'énonciation avec le référent. Mondada et Dubois (1995 : 288) soulignent en outre la dimension intersubjective et interdiscursive de la référenciation en tant que processus de « construction collaborative d'objets de discours – c'est-à-dire d'objets dont l'existence est établie discursivement, émergeant de pratiques symboliques et intersubjectives ». L'acte de référenciation présuppose donc une nature énonciative, contextualisée, subjective et intersubjective, et finalement interdiscursive. Comme le met en évidence Longhi (2015 : 131), le « processus de référenciation [...] ne se fonde pas sur une conception objective du rapport langage / monde, mais sur la médiation sémiotique des ressources langagières dans la mise en mots du réel selon un certain point de vue ». Les formes de référenciation jouent ainsi un rôle important dans la construction des positionnements discursifs.

Nous avons choisi d'utiliser le terme « formes » de référenciation pour prendre en compte la multiplicité des éléments et des configurations par lesquels la référenciation peut se réaliser. Dans le corpus envisagé, ces formes peuvent en effet correspondre à des substantifs, à des syntagmes ou à des propositions identificatoires souvent en apposition au nom, à des reprises référentielles notamment anaphoriques, ou bien à des expressions utilisées dans le cadre de l'identification d'autres individus sur la base de leur relation avec VL (par ex. *les parents du patient* ou *le neveu du jeune homme*).

Parallèlement à la variabilité des « formes » qui actualisent la référenciation, un deuxième critère nous a amenée à privilégier ce terme. Dans le cadre d'études de corpus et notamment lors d'explorations informatisées, le terme *forme* permet de renvoyer aux occurrences du corpus telles qu'elles y sont effectivement orthographiées. Il s'agit donc d'adopter une perspective déterminée sur les données, à la différence d'une exploration par lemme, par exemple. Ces deux approches ne sont toutefois pas exclusives l'une de l'autre. Au contraire, elles peuvent être adoptées à différentes étapes de la démarche ainsi que de manière complémentaire, en fonction des nécessités et des objectifs de l'étude.

L'identification des formes de référenciation de VL dans le corpus a été possible grâce à une analyse contextualisée des concordances du syntagme « Vincent Lambert » et des paragraphes consécutifs. Les formes (*jeune*) *homme*, *patient*, *quadragénaire* et *infirmier* se sont avérées parmi les plus fréquentes. Une analyse des spécificités a ensuite relevé des distributions différentes selon les journaux, comme le montre le graphique de la figure 16.

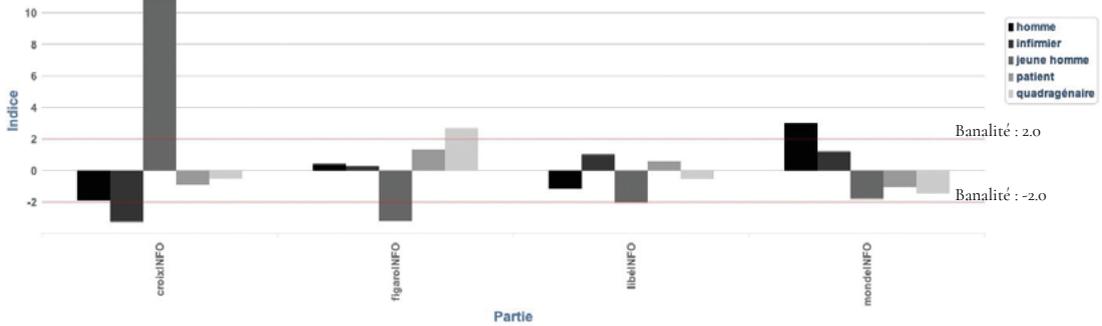


Figure 16. Spécificités de homme, infirmier, jeune homme, patient et quadragénaire.

L'analyse des concordances et des cooccurrences de ces formes fait en outre ressortir des implications argumentatives significatives, qui participent à la construction des positionnements discursifs dominants. On peut constater que « La Croix » désigne VL en recourant à l'expression *jeune homme*, qui est largement surreprésentée dans cette sous-partie par rapport aux articles d'information des autres journaux. Cette forme de référencement peut donc être considérée comme significative et caractéristique des articles d'information de « La Croix ». Il s'agit d'une forme de référencement à valeur générique qui identifie VL en tant qu'être humain adulte de sexe masculin. Elle souligne aussi son jeune âge, ce qui implique moins souvent une situation de fin de vie. Par cette désignation, « La Croix » attire donc l'attention sur la dimension humaine de la question et, en conséquence, du sujet qui est au centre des positions conflictuelles. Le choix de caractériser VL par un terme générique contribue en outre à aborder la question de la fin de vie dans sa dimension plus générale, en interrogeant aussi la conception même de l'être humain. La portée implicitement argumentative de cette forme se rapproche d'ailleurs d'une autre qui est en revanche surreprésentée, parmi d'autres, dans « Le Figaro » : il s'agit de *quadragénaire*. Comme c'est le cas pour le désignatif *jeune homme* dans « La Croix », l'emploi de *quadragénaire* met l'accent sur le jeune âge de VL, représentant d'ailleurs l'un des éléments que l'équipe médicale prend en compte dans l'évaluation du cas et des pratiques à mettre en œuvre.

Des formes de référencement similaires sont en revanche utilisées par « Le Monde » et « Libération ». « Le Monde » désigne principalement¹⁴

14. Précisons à cet égard que l'analyse contextualisée de la forme *homme*, surreprésentée dans les articles d'information du « Monde », a montré qu'elle fait surtout l'objet de références et de désignations institutionnelles.

VL en fonction de son statut socio-professionnel : c'est un *ancien infirmier* qui se trouve en *état végétatif* ou *pauci-relationnel*. L'élément référentiel privilégié est donc la profession, qui relève d'une catégorisation d'ordre social. Cependant, celle-ci perd, au moins partiellement, sa valeur distinctive dans l'état de santé actuel, comme le souligne l'adjectif *ancien*. Le terme *infirmier* tend à être employé souvent dans « Libération » aussi, bien qu'il ne soit spécifique¹⁵ dans aucun des deux cas. Dans les deux journaux, la forme de référencement adoptée souligne l'opposition entre le statut socio-professionnel de VL avant l'accident et son état actuel, qui l'empêche d'exercer sa profession. Tandis que « Le Monde » utilise presque toujours le terme en cooccurrence avec l'adjectif *ancien*, en mettant l'accent sur le changement entre l'avant-accident et l'après, « Libération » alterne le recours à l'adjectif *ancien* et l'emploi du verbe *être* à l'imparfait. Dans les deux cas, le contraste passé-présent est néanmoins dominant.

« Le Figaro » fait au contraire référence à VL en précisant souvent son état socio-sanitaire par le terme *patient* (65 occurrences sur les 81 totales dans la sous-partie INFO du « Figaro » réfèrent à VL). Cette forme de référencement permet d'attirer l'attention en premier lieu sur le domaine institutionnel qui est censé s'occuper de lui, à savoir l'institution médicale. De fait, ce terme figure aussi bien dans les avis et les décisions médicaux et juridiques concernant VL (31 %), que dans des constructions appositives de caractérisation de VL (23 %) ou de personnes en relation avec lui (26 %) (cf. figure 17). En deuxième lieu, la forme *patient* souligne également la dimension passive de la situation du référent ainsi désigné, dans le sens que toute décision est subie par VL, indépendamment de sa volonté. Elle contribue donc à véhiculer une mise en perspective déterminée, dotée aussi d'une valeur argumentative.

L'analyse contextualisée du terme *patient* dans les articles du « Figaro » a par ailleurs fait ressortir son usage en association avec un syntagme de spécification qui précise l'état de santé. VL est en effet défini comme étant un *patient tétraplégique* (17 occurrences sur 29 dans le sous-corpus PN-INFO) *en état végétatif* (6 occurrences) ou, plus fréquemment, *en état de conscience minimale* (14 occurrences). Ce procédé de mise en discours de la forme de référencement *patient* est évocateur d'une position dépendante des soins et des traitements, et il laisse également entendre

15. L'indice de spécificité est en effet inférieur au seuil de banalité fixé conventionnellement à +2/-2 sous TXM (Heiden *et al.* 2010) (cf. figure 16).

l'impossibilité d'action et de décision de la part du patient lui-même. De plus, le suremploi de ce terme favorise l'encadrement de cette situation comme un cas d'invalidité ou de handicap grave et non pas comme une situation de fin de vie, ce qui représente l'un des aspects principaux de la polarisation.

Les autres journaux utilisent également la forme *patient* en tant que terme spécialisé médico-juridique pour désigner VL, bien qu'elle ait une saillance différente selon les cas. Une étude qualitative contextualisée des occurrences de la forme *patient* a ainsi permis de distinguer les différentes valeurs référentielles principalement actualisées, dont les résultats sont résumés dans le graphique de la figure 17.

La distribution des valeurs référentielles assumées par la forme *patient* en fonction de ses usages en contexte dégage des tendances différentes dans les quatre journaux. « Le Figaro » et « Le Monde » privilégient la forme *patient* pour définir la condition de VL, mais aussi pour identifier les personnes qui sont censées prendre des décisions face à sa situation, qu'il s'agisse des proches ou des spécialistes de l'institution médicale ou éthique. Dans ce deuxième cas, l'emploi de *patient* participe à la mise en évidence de la situation de passivité vécue par VL, qui se trouve dans l'impossibilité de décider de son sort. En revanche, dans « La Croix » et « Libération », le terme *patient* est plutôt employé pour mettre en relation ce cas spécifique avec une situation générique à laquelle il pourrait être rattaché, et cela notamment dans « Libération ». L'utilisation de *patient* avec la fonction de renvoyer anaphoriquement à VL est assez fréquente respectivement dans « La Croix », « Le Figaro » et « Libération ».

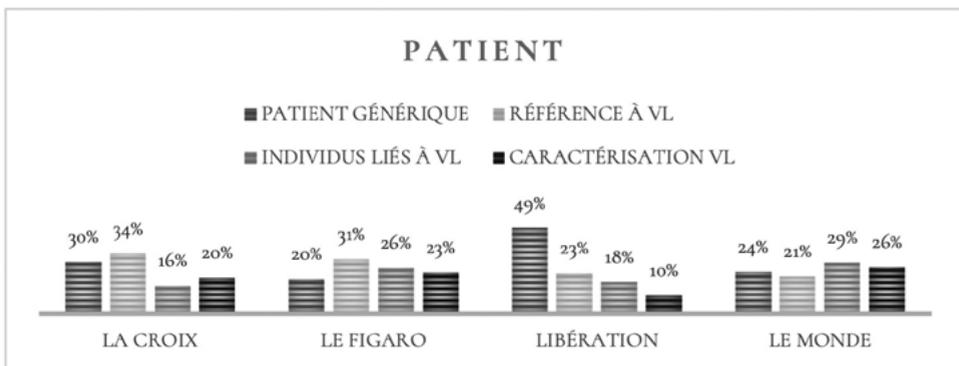


Figure 17. Valeurs référentielles de la forme *patient* (sous-corpus : PN-INFO).

4.2.2. Construction des objets de discours : procédés discursifs

4.2.2.1. Procédés de référencement et de mise en discours de l'état de santé

La catégorisation de l'état de santé de VL représente un autre aspect clé dans la mise en perspective de l'événement et dans la polarisation des positionnements, qui est abordée différemment dans les quatre journaux. En l'occurrence, la construction de l'objet de discours – l'état de santé – s'opère à travers non seulement plusieurs choix de référencement, mais aussi divers procédés de mise en discours de ceux-ci. Trois formes sont principalement employées : *végétatif* (138 occurrences), *tétraplégique* (65 occurrences) et *pauci-relationnel* (23 occurrences). L'analyse des spécificités montre toutefois que chaque journal tend à recourir davantage à une forme plutôt qu'aux autres, ainsi que l'illustre la figure 18.

Comme on l'a déjà anticipé plus haut, la forme *tétraplégique* est largement utilisée dans « Le Figaro » aussi bien en tant que substantif qu'adjectif. À la différence des autres journaux, le large recours à cette forme de référencement dans les articles du « Figaro » permet d'insister sur la catégorisation de l'état de VL comme un cas de tétraplégie et donc de

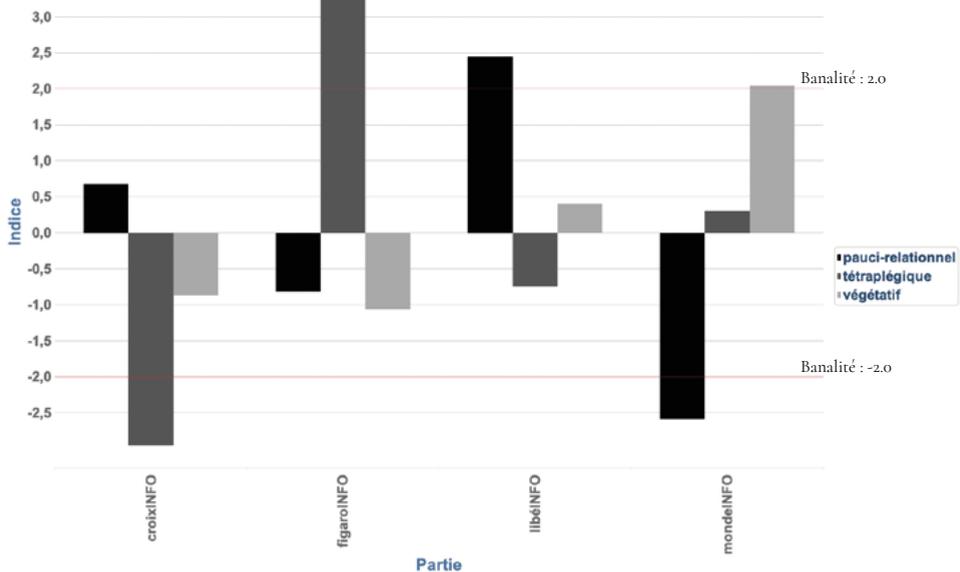


Figure 18. Calcul des spécificités sur les formes pauci-relationnel, tétraplégique et végétatif.

handicap, ce qui a des répercussions argumentatives. Cette mise en discours semblerait en effet justifier indirectement le maintien en vie de VL, en soulignant certes son invalidité, mais en réfutant en même temps d'assimiler sa situation à un cas de fin de vie.

Au contraire, « Le Monde » emploie le terme *tétraplégique* surtout en cooccurrence avec l'adjectif *végétatif* (13 occurrences sur 15), qui souligne l'absence de conscience accompagnant la parésie. Par ailleurs, le suremploi du syntagme prépositionnel *en état végétatif*, qui est spécifique du « Monde » (cf. figure 18), apparaît souvent en cooccurrence avec l'adverbe *depuis*, employé 33 fois sur les 44 formes totales, ainsi qu'avec les adjectifs *chronique* (8 cooccurrences) et *irréversible* (9 cooccurrences). Ces procédés de contextualisation de l'expression catégorisant l'état de santé contribuent de fait à mettre en relief la stabilité d'une situation qui perdure dans le temps. Cette mise en discours de l'état de VL souligne l'absence de conscience propre à l'état végétatif – ce qui évoque d'ailleurs la distinction entre une vie physique et une vie psychique – en insistant également sur l'immutabilité de la situation. La catégorisation prédominante dans « Le Monde » en termes d'état végétatif favorise l'assimilation à un cas de fin de vie, dont la survie ne dépend que des machines qui l'alimentent et l'hydratent, sans reconnaître aucune forme d'intentionnalité. Ces procédés de mise en discours du syntagme *état végétatif* sembleraient donc se rapprocher des constructions discursives et des arguments caractéristiques d'un positionnement favorable à l'interruption des traitements.

Enfin, l'expression *en état pauci-relationnel* est suremployée dans « La Croix » et plus encore dans « Libération », tout en étant la moins exploitée des trois. Le choix entre *pauci-relationnel* et *végétatif* est particulièrement significatif, car les deux impliquent, respectivement, la présence ou l'absence d'un état de conscience minimale, comme on peut le lire aussi sur les sites de plusieurs associations s'occupant de l'accompagnement de patients dans ces états ainsi que de leurs proches¹⁶. Par conséquent, l'emploi de *pauci-relationnel* peut favoriser une mise en discours de la situation de VL en tant que malade en état de *handicap* (forme également suremployée par « La Croix ») et non pas en fin de vie, déterminant ainsi le rapprochement du positionnement opposé par rapport à celui suremployant l'expression *état végétatif*.

16. Voir par exemple les définitions proposées sur <https://www.avc-normandie.fr/espace-professionnels/evc-epr/etat-vegetatif-chronique-etat-pauci-relationnel,4306,5629.html> et sur <https://aftc-gironde.org/etat-vegetatif-chronique-evc-etat-pauci-relationnel-epr/> (dernière consultation le 10/05/2023).

Par ailleurs, une même expression de catégorisation de l'état de santé peut également faire l'objet de mises en discours différentes, impliquant des perspectivations argumentatives opposées, comme c'est le cas pour *pauci-relationnel*. L'analyse des concordances de cette forme permet en effet de mettre en relief des différences entre les usages de « La Croix » et ceux de « Libération », entraînant des répercussions sur la polarisation de la construction discursive de l'objet de discours « état de santé ». Dans le cas du quotidien d'inspiration catholique, la forme *pauci-relationnel*, tout en n'étant pas spécifique de la sous-partie, est employée à la fois pour définir, dans un nombre limité de cas, l'état de VL, en en soulignant notamment la régression vers l'état végétatif (2 occurrences), et pour identifier une catégorie de patients ainsi que les services spécifiques qui les prennent en charge. En l'occurrence, le recours à l'expression *patients en état végétatif ou pauci-relationnel*¹⁷ tend à lisser, voire annuler les différences entre les deux états. Ainsi, le degré de conscience du patient, qu'il soit minimal ou absent, ne représente pas un critère significatif. L'attention se dirige plutôt sur la dimension humaine de VL et, plus en général, des personnes vivant dans des conditions de santé similaires, au nom de la défense de l'accompagnement et du respect de la vie, indépendamment de sa condition.

Dans les articles de « Libération » aussi, la forme *pauci-relationnel* est mise en relation avec la question de la conscience du patient. En outre, elle fait souvent l'objet d'une mise en discours entre guillemets, relevant d'un usage autonymique à valeur modalisante. Ce procédé signale la nature technique du mot tout en s'accompagnant parfois d'une prise de distance de la part de l'énonciateur premier, comme on peut le voir dans l'exemple ci-dessous tiré d'un article signé par les journalistes Favereau et Albertini.

- (6) Plongé dans un état dit « pauci-relationnel », c'est-à-dire sans lien avec l'extérieur, Vincent Lambert était depuis des années nourri et hydraté artificiellement (« Libération », 20/05/2019, récit, E. Favereau/D. Albertini, PN-INFO).

La périphrase explicative insérée par L₁/E₁ signale que, dans les faits, l'existence de VL est dépouillée de toute relation avec ce qui l'entoure, son maintien en vie ne dépendant que de la nutrition et de l'hydratation

17. Les deux adjectifs peuvent apparaître également en ordre inversé, sans que cela change la portée sémantique et pragmatique de la référénciation.

artificielles. L'absence de conscience ainsi mise en relief attire l'attention sur la dimension biologique d'une vie, présentée comme entièrement dépendante de ces traitements et donc comme privée de son cours naturel. Cette mise en discours fait d'ailleurs appel à la réflexion – certes profondément subjective – sur la conception de la vie et sur la notion d'une « vie digne ».

On remarque finalement que l'expression catégorisante *en état pauci-relationnel* revêt des connotations et des implications différentes d'un point de vue argumentatif dans les articles des deux journaux. Comme c'est souvent le cas, l'analyse inductive-quantifiée aide à faire ressortir des éléments significatifs, sans pour autant qu'il soit possible d'en identifier le rôle et la portée non seulement sémantique mais aussi pragma-discursive, ce qui exige en revanche une étude déductive-qualitative, menée aux niveaux local et global.

Au-delà de *tétraplégique*, *végétatif* et *pauci-relationnel*, on peut identifier une quatrième expression employée pour qualifier l'état de VL. Il s'agit de la négation de l'état de fin de vie selon la structure syntaxique suivante : *ne + être + pas + en fin de vie*, qui est attestée souvent au sein des discours autonomiques entre guillemets. En effet, le refus que la situation de VL puisse être assimilée à un cas de fin de vie constitue l'un des arguments défendus par les partisans de son maintien en vie. D'un point de vue argumentatif, il est donc possible de rattacher l'emploi de cette expression de dénégation au recours au terme *handicapé*, puisque les deux dénie – explicitement ou implicitement selon le cas – la possibilité de catégoriser ce cas comme une situation de fin de vie. On peut repérer l'utilisation de cette structure de dénégation notamment dans les déclarations des parents, rapportées par tous les journaux, souvent à travers des modes de représentation autonymique entre guillemets. Cependant, une analyse des spécificités réalisée à la fois sur la séquence *pas + en fin de vie* et sur le terme *handicapé* montre que, d'un point de vue quantitatif, l'expression dénégative est particulièrement présente et donc mise en évidence dans « La Croix » et « Le Figaro », comme on peut l'observer dans le graphique de la figure 19.

Dans « Le Figaro », il arrive en outre que la négation de l'état de fin de vie s'accompagne de la qualification de VL de *handicapé*, revenant à constituer une sorte de binôme associatif, ce qui est moins souvent le cas tant dans « Libération » que dans « Le Monde ». La fréquence élevée de la mise en discours de cet argument contribue à mettre en relief une conception de l'état de VL qui étayerait le refus de l'arrêt des traitements, orientant ainsi la construction discursive de l'événement et, en conséquence, son interprétation.

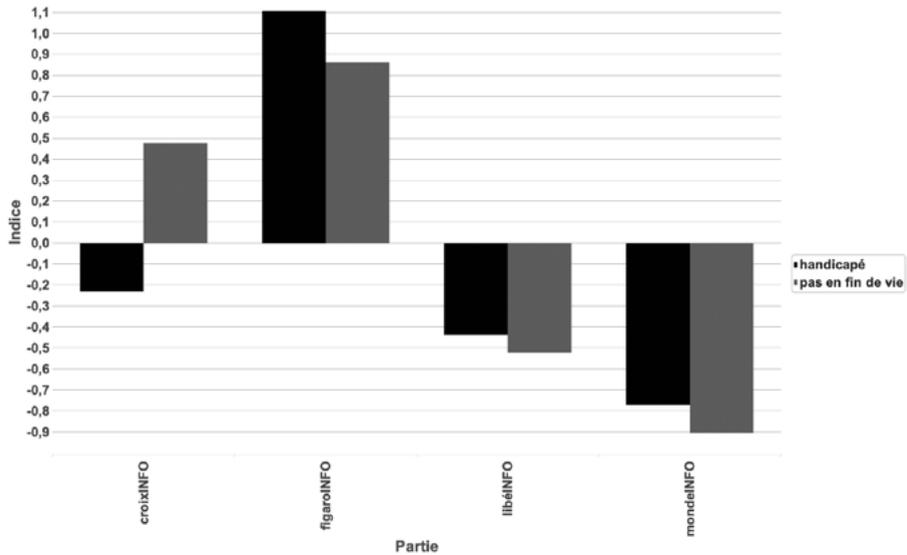


Figure 19. Calcul des spécificités sur les formes handicapé et pas en fin de vie.

4.2.2.2. Procédés de mise en discours de la mort

La mort représente un autre thème central des discours relatifs à l'« affaire Vincent Lambert » et plus en général au débat sur la fin de vie. Dans le cadre du sous-corpus PN-INFO, ce sujet est toutefois traité en relation avec deux questions en particulier : le déroulement du processus médical qui précède la mort et l'annonce de la mort de VL.

Si des différences significatives ne s'observent ni dans les choix de référence, ni au niveau de cooccurrences récurrentes, les quatre journaux proposent toutefois des mises en discours divergentes de ces questions, en attirant l'attention sur des aspects divers. Les articles de « La Croix » abordent la situation de VL et la décision d'arrêter les traitements¹⁸ dans leur valence générale, c'est-à-dire en fonction des enjeux socio-éthiques qu'elles entraînent. Dans l'exemple suivant (7), le journal cite les propos du pape François et met ainsi en avant une critique de la société contemporaine, accusée de ne pas prendre suffisamment en charge les personnes les plus fragiles. Un événement singulier devient donc une occasion pour solliciter la réflexion sur la question plus générale de la fin de vie et de l'accompagnement des malades.

18. Si l'opposition *soins-traitements* peut parfois être évocatrice de la polarisation des avis dans le débat (Cagninelli 2022), nous utilisons toutefois le terme *traitements* non pas pour sa valence argumentative, mais en tant que terme spécialisé, employé par les institutions médicale et juridique.

- (7) [...] le pape François a invité ce mercredi 10 juillet à prier « pour les **malades abandonnés et qu'on laisse mourir** » (« La Croix », 10/07/2019, religion, Houdaille, PN-INFO).

Le choix de l'énonciateur second (le pape) et les mots qui en ont été rapportés permettent de mettre en relief l'une des prises de position sur la question, privilégiant une perspectivation déterminée que l'on pourrait considérer comme plus proche d'un positionnement vis-à-vis de la controverse que de l'autre. La citation rapportée met en discours la question de la mort en attirant l'attention sur une catégorie particulière – à valeur englobante par rapport au cas d'actualité – qui est concernée : les personnes malades. Elle souligne en outre la dimension passive de l'abandon des malades à leur sort, impliquée par le choix du verbe *laisser mourir*.

Un deuxième aspect caractérise également la mise en discours de la mort dans les articles de « La Croix ». Différemment de l'exemple 7, d'autres articles mettent plutôt en évidence la dimension active de l'arrêt des traitements qui va *conduire* à la mort, la présentant ainsi comme un événement provoqué. Cela se réalise non seulement par l'emploi de verbes qui expriment l'accomplissement de l'action et présupposent une valeur agentive, tels que *provoquer* et *conduire*, mais aussi en dressant une comparaison – de manière indirecte – avec l'exécution d'une peine capitale, ainsi que le montre l'exemple suivant.

- (8) Sur Twitter et avec une tonalité plus virulente Mgr Marc Aillet, évêque de Bayonne, s'est lui indigné : « Nul n'a le droit de **provoquer la mort d'autrui**, ni pouvoir médical ni pouvoir judiciaire. **La peine de mort a été abolie.** [...] » (« La Croix », 16/05/2019, religion, Bevilacqua, PN-INFO).

Comme dans l'exemple précédent, on remarque ici que la construction polarisée de la mort fait l'objet des segments textuels entre guillemets, rapportant les propos d'énonciateurs seconds. En général, le champ sémantique de la mort provoquée et délibérée traverse les articles de « La Croix », faisant ressortir l'aspect émotionnel des déclarations concernant ce cas. Conformément à la ligne éditoriale du journal, les propos rapportés sont en particulier ceux des personnalités religieuses, mais aussi des parents de VL ou de leurs avocats. Des expressions fortes sont ainsi représentées au sein de segments entre guillemets, comme c'est le cas pour « crime d'État », « tentative d'assassinat » et encore « meurtre avec préméditation », utilisées notamment par les parents de VL et par leurs avocats pour référer à

sa mort. Si tous les journaux rendent compte de ces expressions, ils le font toutefois avec des focalisations différentes en ce qui concerne aussi bien la représentation quantitative de ces expressions que leur sélection. Par *focalisation*, nous désignons le fait que, dans le cadre de la construction d'un même objet de discours, l'on met en évidence plus fréquemment certains aspects au détriment d'autres, tout comme le fait qu'on donne plus de place à la représentation des propos de certaines instances énonciatives plutôt qu'à celle d'autres. Il s'ensuit que des focalisations ciblées contribuent à leur tour à une construction orientée de la question débattue. La focalisation participe donc de la « dimension argumentative » (Amossy 2016) du discours et résulte d'une présence quantitativement significative d'éléments comparables au niveau paradigmatique, c'est-à-dire d'éléments potentiellement interchangeables ou alternatifs du point de vue référentiel, chacun étant néanmoins doté d'une spécificité sémantique et susceptible donc d'activer des points de vue et des valeurs argumentatives différentes. L'étude des focalisations repose donc nécessairement sur des analyses quantifiées et contrastives effectuées sur corpus, afin d'examiner la distribution d'éléments comparables en fonction d'un niveau d'hétérogénéité.

Le graphique suivant (figure 20) illustre les résultats de l'analyse des spécificités réalisée sur les formes *assassinat*, *crime* et *meurtre*, pouvant être considérées comme trois choix de référencement alternatifs pour la construction de l'objet de discours « mort » que l'on peut retrouver dans les déclarations rapportées. Il s'agit donc d'un point de départ efficace pour l'étude d'éventuelles focalisations engendrées par la forme de référencement privilégiée.

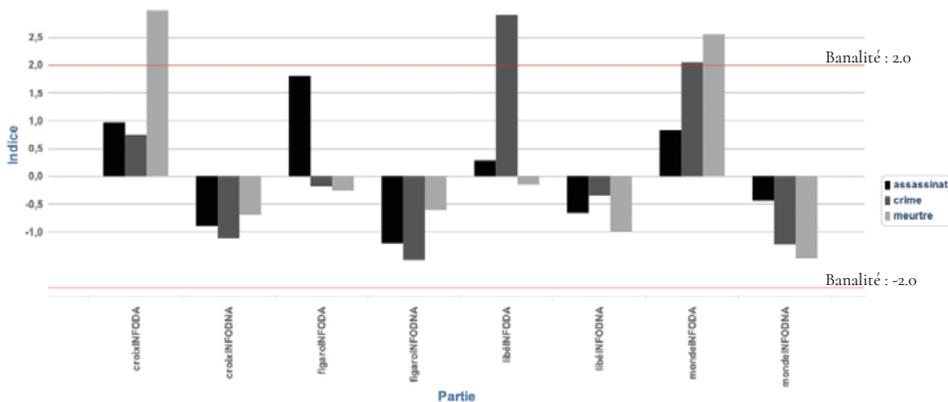


Figure 20. Spécificités sur les formes *assassinat*, *crime* et *meurtre*.

Dans tous les cas, les trois formes apparaissent dans les segments textuels relevant d'une représentation autonymique entre guillemets. En l'occurrence, les guillemets ne marquent pas seulement une rupture énonciative introduisant les propos d'un autre énonciateur que l'énonciateur journalistique, mais ils signalent aussi une forme de modalisation. Il s'agit en effet de segments textuels que les quotidiens ne prennent pas en charge à leur compte en raison de leur portée axiologique et argumentative. Il est néanmoins intéressant de remarquer qu'ils les rapportent avec des fréquences différentes, selon les aspects qu'ils souhaitent développer davantage.

La mise en discours des déclarations relatives au thème de la mort repérée dans « La Croix » présente des similitudes avec celle du « Figaro », le journal qui leur donne le plus de place. En l'occurrence, la dimension active de l'arrêt des traitements se dégage de sa représentation en tant qu'acte qui va donner la mort – dans le cas de VL – et non pas qui laisse partir le patient. À plusieurs reprises, des déclarations axiologisées négativement telles que *condamnation à mort* et *mise à mort* font l'objet d'une représentation autonymique entre guillemets, signalant le rôle de rapporteur du journaliste.

« Le Figaro » rapporte l'avis de quelques spécialistes des soins palliatifs qui, « même s'ils sont loin de tous partager le même point de vue » (Leclair, 19/05/2019), réfutent l'assimilation de l'« arrêt des soins » à un acte délibéré de donner la mort. Cette réfutation ne contribue cependant pas à offrir une représentation équilibrée des deux positions. La cooccurrence entre *provoquer* et *mort* est ainsi employée dans des contextes opposés, permettant de donner voix aux deux positions antagonistes (tableau 1). La représentation de la dimension active de l'acte de donner la mort s'avère néanmoins prédominante du point de vue quantitatif.

En ce qui concerne « Le Monde », en dehors des segments textuels entre guillemets, le journal utilise les trois expressions principalement associées à la mort – *mort*, *mourir*, *décès* – de manière équilibrée. Si l'emploi de *décès* ne se révèle pas particulièrement significatif, dans la mesure où il apparaît très souvent en relation avec des questions médicales ainsi que dans les discours rapportés de spécialistes, les formes *mort* et *mourir* sont en revanche utilisées pour rendre compte de positions antagonistes. À la différence de « La Croix » et du « Figaro », dans « Le Monde » l'arrêt des traitements est présenté comme un acte qui va *mener* à la mort, plutôt que *provoquer* ou *conduire*. Alors que *provoquer* souligne le caractère actif et délibéré de l'acte, le choix de *mener* véhicule plutôt l'image d'un accompagnement vers la mort, comme on peut le voir dans les concordances présentées dans le tableau 2.

Tableau 1. Concordances de provoquer en cooccurrence avec mort (sous-corpus : « Le Figaro » – PN-INFO).

Références corpus « Le Figaro » (PN-INFO)	Contexte gauche	Pivot	Contexte droit	Polarisation
<i>figaroAFP, 24/04/2019</i>	n'est donc pas définitive et la décision de	provoquer la mort	de Vincent Lambert ne peut pas être exécutée »,	<i>Faire mourir</i>
<i>Lutard, 26/04/2019</i>	Il n'a pas le droit de	provoquer délibérément la mort	». « Cela fait partie des choses sur lesquelles nous réfléchissons »,	<i>Faire mourir</i>
<i>Leclair, 19/05/2019</i>	eau dans l'organisme va entraîner une insuffisance rénale et	provoquer la mort	de manière naturelle.	<i>Laisser mourir</i>
<i>Leclair, 19/05/2019</i>	L'objectif est de soulager et non de	provoquer la mort	. Il ne s'agit en aucun cas d'une euthanasie »,	<i>Laisser mourir</i>

Tableau 2. Extrait des concordances de mort et mourir (sous-corpus : « Le Monde » – PN-INFO).

Références corpus « Le Monde » (PN-INFO)	Contexte gauche	Pivot	Contexte droit
<i>Béguin, 02/07/2019</i>	œuvre mardi par l'équipe médicale doit mener à la	mort	de Vincent Lambert en quelques jours.
<i>Jacquin, 28/06/2019</i>	l'alimentation et de l'hydratation devant mener à la	mort	de cet ancien infirmier.
<i>Jacquin, 25/06/2019</i>	l'arrêt de l'alimentation et de l'hydratation, devant mener à la	mort	du patient.

Le verbe *mener* permet d'éviter la connotation péjorative entraînée par *provoquer*, qui a été utilisé aussi bien pour dénoncer la nature euthanasique du protocole médical (selon les parents), que pour la réfuter (notamment de

la part des spécialistes en soins palliatifs), ainsi que l'ont montré les extraits du « Figaro » (tableau 1).

Dans « Le Monde », la mort est en outre présentée comme un événement qui survient naturellement, ainsi qu'il ressort des mots des responsables du premier arrêt des traitements de VL (mai 2013), dont le journal rapporte les propos (Béguin, 20/05/2019). La décision de l'époque est en effet justifiée par l'intention de « le laisser partir lentement, de sa belle mort », tout en reconnaissant en même temps l'erreur procédurale accomplie à l'époque quand une hydratation minimale avait été maintenue. Dans l'un des derniers articles du corpus consacrés à l'affaire, « Le Monde » présente enfin VL non plus comme le symbole du débat sur la fin de vie, mais comme le symbole de l'une des deux positions, à savoir la défense d'un « droit à mourir dans la dignité » (Favre Le Cadre, 11/07/2019).

Les procédés discursifs employés dans les articles de « Libération » se différencient en revanche des tendances relevées dans les autres journaux. Tout d'abord, on ne repère pas de cooccurrences significatives entre *mort* et des verbes déterminés. Les déclarations connotées négativement de la part d'énonciateurs défendant une position favorable au maintien des traitements ne sont pas non plus particulièrement représentées. Par contre, plus d'espace est réservé aux réactions suivant l'annonce de la mort, qui la définissent comme un « meurtre avec préméditation » ou encore comme un « crime d'État ». En l'occurrence, la mise en discours de ce sujet est principalement narrative-factuelle. Sans préciser la dimension active ou passive du processus précédant cet avènement, la construction discursive de la mort attire en particulier l'attention sur la sédation profonde qui préviendra d'éventuelles souffrances.

Si dans les autres journaux la mise en discours du thème de la mort semblerait se rallier à l'une des deux positions, dans « Libération » le positionnement discursif se fonde en revanche sur l'absence des éléments caractéristiques repérés dans les autres sous-parties d'articles d'information. De fait, seule l'analyse contrastive entre les divers métaénonciateurs permet de dégager la construction d'un positionnement discursif par soustraction des éléments qui sont récurrents et spécifiques des autres instances de comparaison.

Par ailleurs, l'infinitif *mourir* est le seul élément suremployé dans un contexte expliquant, du point de vue médical, l'avènement de la mort du patient à la suite de l'arrêt de l'hydratation et de la nutrition artificielles. À cette question, « Libération » a consacré un article entier de la rubrique

appelée *CheckNews*¹⁹. À travers la reproduction des mots de la *Société française d'accompagnement et de soins palliatifs*, l'article précise deux aspects de l'arrêt des traitements en question : il constitue « un arrêt du maintien en vie artificiel » qui sera « suivi d'une intensification des soins ». Cela permet donc de refuser de manière indirecte deux des arguments principaux du positionnement adverse, à savoir que VL n'est pas en fin de vie et qu'il va être privé de soins. Le journaliste prend ensuite la parole pour prévenir le lecteur de ne pas assimiler cette procédure à « une euthanasie », avant de céder de nouveau la parole au discours expert pour en donner une définition :

- (9) Attention à **ne pas confondre cet arrêt du maintien en vie avec une euthanasie**, qui consiste à « mettre fin à la vie d'une personne malade à sa demande » et reviendrait à lui injecter un produit létal (« Libération », 20/05/2019, ServiceChecknews, PN-INFO).

Dans la conclusion de l'article, on explique enfin que le patient mourra de manière naturelle et non pas de faim et de soif. La progression textuelle permet ainsi de faire ressortir le contraste entre la dimension artificielle du maintien en vie de VL, qui ne dépendrait que des traitements assurés, et la mort qui surviendra au contraire naturellement.

Cette mise en discours de la question – qui, d'une part, n'entraîne pas de références à une présentation active de la mort en relation avec l'arrêt des traitements et, de l'autre, met en relief l'avènement naturel de la mort, sans aucune souffrance pour le patient qui bénéficiera « d'une intensification des soins » – semblerait ainsi offrir une perspectivation de la question qui penche plutôt pour une position favorable à l'arrêt des traitements.

4.3. Synthèse comparative et articulation du local au global

La nature informative et objectivée des articles rassemblés dans le sous-corpus PN-INFO empêche la présence d'une prise de position explicite que l'on pourrait attribuer à l'énonciateur premier de l'article ou bien au journal. Cependant, l'analyse déductive-contextualisée a montré que les procédés de mise en discours, les choix de référencement, leurs fréquences, la mise en avant quantitative de certains aspects et la sous-représen-

19. Pour une description détaillée de cette rubrique, on renvoie à la page officielle : <https://www.liberation.fr/checknews-about/> (dernière consultation le 27/04/2023).

sentation d'autres peuvent contribuer à construire l'événement selon une perspectivation déterminée. Celle-ci favorise une représentation orientée de la question, qui peut être plus conforme à une position plutôt qu'à une autre dans la polarisation caractérisant tout conflit polémique, dont l'affaire VL est un exemple. En effet, ainsi que l'affirme Amossy, la polémique entraîne une

rhétorique de la polarisation consist[ant] à établir des camps ennemis [...]. Il s'agit de rallier à un groupe constitutif d'une identité [ou d'un positionnement – dirions-nous], ou de présenter les choses de façon à ce que ceux qui se sentent au départ solidaires d'un groupe donné se mobilisent en faveur de la thèse qui le renforce (Amossy 2014 : 59).

Si, comme le montre Amossy, la polarisation peut reposer sur des procédés explicites identifiant et marquant l'opposition de groupes adverses, l'analyse conduite jusque-là sur le sous-corpus PN-INFO met en relief le rôle des procédés de mise en discours de l'événement. Ces derniers offrent en effet des perspectivations différentes des objets du discours, ainsi que des focalisations de l'attention sur certains aspects au détriment d'autres. Cela favorise une représentation orientée de l'événement qui peut se rapprocher, du point de vue discursif, de l'un des deux positionnements de la polarisation.

En croisant les résultats des deux phases de l'analyse, il est possible de dégager un positionnement discursif dominant dans chaque sous-partie métaénonciatrice. L'articulation des résultats inductifs-quantifiés avec les explorations déductives-contextualisées représente en effet l'étape finale de notre démarche. L'analyse de la construction discursive de deux des objets de discours fondamentaux de la controverse a finalement permis de mieux comprendre les écarts et les apparentements mis en évidence par les AFC réalisées aussi bien sur les formes que sur les catégories grammaticales. Grâce à l'articulation de ces analyses, il a ainsi été possible de mettre en évidence les constantes et les différences tant au niveau de l'hétérogénéité métaénonciatrice que de l'hétérogénéité sous-générique.

4.3.1. *Rapprochements et divergences de nature métaénonciatrice*

L'analyse des choix de référenciation et des procédés de construction discursive des deux objets de discours analysés fait ressortir plusieurs convergences et divergences entre les articles d'information des quatre journaux

sélectionnés. On peut remarquer tout d'abord des similitudes entre les formes et les procédés de mise en discours employés par « La Croix » et par « Le Figaro ». Les choix de référencement privilégiés par les deux journaux participent en effet à une mise en perspective de l'événement qui se rapproche d'un positionnement « pro-maintien », c'est-à-dire un positionnement favorable au maintien des traitements administrés à VL. Néanmoins, les deux journaux mettent en relief des aspects différents. Alors que « La Croix » attire l'attention sur l'humanité de VL (*jeune homme*) et de toute personne dans un état de souffrance, « Le Figaro » définit principalement VL en tant que *patient*, contribuant ainsi à mettre en évidence le domaine institutionnel de référence de la question, le domaine médical. Cette forme de référencement souligne également son rôle passif en raison de son impossibilité d'action et de décision. L'une des différences principales entre les deux journaux concerne toutefois la tendance de « La Croix » à éviter de trancher de manière générale sur la question, en soulignant la nécessité d'envisager ces situations au cas par cas.

D'autres aspects partagés ont trait à la mise en discours du thème de la mort, soulignant notamment sa dimension active en tant qu'acte provoqué. Cette construction discursive est quantitativement persistante dans les articles du « Figaro », bien qu'elle soit présente surtout au sein des discours autonymiques entre guillemets. On retrouve aussi une représentation active de la mort dans les déclarations rapportées dans les articles de « La Croix », caractérisées généralement par le refus des pratiques euthanasiques actives, en particulier de la part d'instances d'énonciation religieuses. Quoique cette position soit prédominante dans le domaine religieux, « La Croix » souligne également l'existence d'avis divergents dans le monde catholique. La question de la mort y est abordée dans le cadre de la fin de vie selon une perspective générale, qui souligne les enjeux sociaux et éthiques de ces situations, ainsi que la nécessité de défendre et d'accompagner ces personnes. Si une perspective polarisée se dégage de la mise en discours de la mort dans les articles de « La Croix », il ne s'agit pas d'une prise de position tranchée.

Au contraire, les mises en discours de l'état de santé de VL et les constructions discursives de la mort dans « Libération » et « Le Monde » semblent s'apparenter davantage à un positionnement « pro-arrêt », c'est-à-dire un positionnement favorable à l'arrêt des traitements. En ce qui concerne l'identification de VL, les deux journaux mettent en relief son statut socio-professionnel, par le recours à *infirmier* et notamment à *ancien infirmier*. Cette expression focalise l'attention sur la fracture temporelle qui distingue deux phases bien différentes de sa vie, pouvant par

ailleurs faire écho à la réflexion sur la conception de la vie en relation par exemple avec le rôle social de l'individu et de sa possibilité d'être actif. La construction discursive de l'état de VL est également significative dans la mise en perspective de l'événement qui, dans les deux cas, souligne l'absence de conscience qui le caractériserait, tout en employant des formes de référenciation et des procédés discursifs différents. Ces perspectivations semblent ainsi se rallier, du point de vue discursif, à un positionnement favorable à l'interruption des traitements.

Quant à la construction discursive du thème de la mort, les articles de « Libération » présentent une mise en discours d'ordre narratif-factuel, sans donner beaucoup de place à la représentation de déclarations prononcées à ce sujet par les partisans du positionnement pro-maintien, où l'on peut repérer des éléments linguistiques axiologisés. L'attention est dirigée en particulier sur la nature de la mort suivant l'arrêt des traitements, présenté comme le prélude d'un événement naturel. Cette perspectivation amène donc à interpréter la représentation de la mort dans « Libération » comme se rapprochant des discours pro-arrêt. La construction discursive du « Monde » encadre en revanche l'arrêt des traitements comme une forme d'accompagnement vers la mort, une mort qui est présentée comme survenant de manière naturelle. Certes moins ouvertement que dans le cas de « Libération », la représentation offerte par « Le Monde » semble se rattacher cependant à son tour à un positionnement pro-arrêt lors de ce cas spécifique.

Quatre positionnements discursifs dominants peuvent être ainsi identifiés selon une opposition graduelle entre les deux positions polarisées, en donnant lieu à des positionnements discursivement plus mitigés pour « La Croix » et « Le Monde » et à d'autres plus marqués pour « Le Figaro » et « Libération ».

4.3.2. *Comparaison sous-générique et hétérogénéité métaénonciatrice*

Une analyse circonscrite et ponctuelle des articles d'opinion en fonction de leur hétérogénéité métaénonciatrice permet à ce point de vérifier les différentes constructions discursives des objets de discours repérées dans les articles d'information, ainsi que les positionnements discursifs qui en ressortent sur la base des rapprochements et des divergences dégagés jusqu'à présent. Le sous-corpus PN-OP est ainsi utilisé comme une sorte de « corpus de contrôle » dans le cadre d'une comparaison sous-générique. Bien qu'elle soit très délimitée, l'exploration du corpus suit, dans ce cas aussi, une

démarche allant du global au local, avant que les résultats ne soient mis en relation dans une synthèse finale.

4.3.2.1. Aperçu des rapprochements et oppositions selon l'hétérogénéité métaénonciatrice dans PN-OP

Les écarts et les convergences attestés par l'analyse déductive-contextualisée de PN-INFO au niveau des métaénonciateurs semblent être confirmés tout d'abord par l'AFC effectuée sur les formes des articles d'opinion (figure 21). Le graphique de l'AFC montre en effet que les quatre journaux s'apparentent deux par deux selon la polarisation déjà mise en évidence à partir des articles d'information : « Le Monde » et « Libération », d'une part, et « Le Figaro » et « La Croix », de l'autre. En l'occurrence, une distance plus marquée sépare néanmoins ces derniers et surtout le quotidien d'obédience catholique.

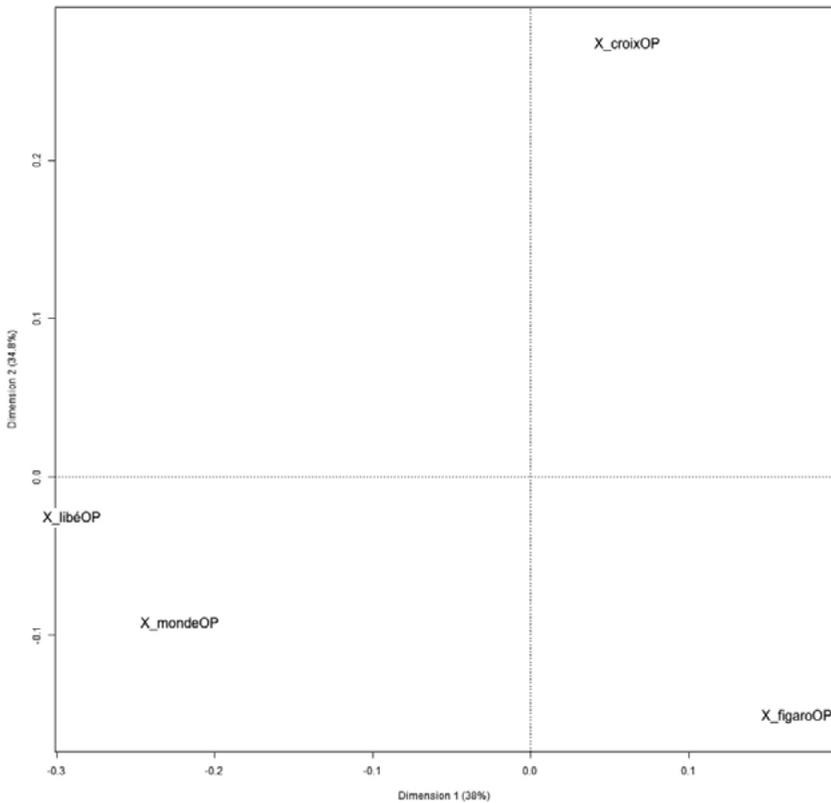


Figure 21. AFC sur les formes (sous-corpus PN-OP).

Les métadonnées concernant, respectivement, la classification sous-générique des articles et les discours autonomiques entre guillemets permettent de repérer, du point de vue quantitatif, quelques tendances liées aux deux journaux affichés dans la partie droite du graphique. D'une part, grâce à l'annotation sous-générique, on peut en effet constater que « La Croix » et « Le Figaro » sont les deux journaux qui ont consacré le plus d'espace à l'affaire VL dans leurs articles à la fois d'information et d'opinion (cf. 3.1.3.2). De l'autre, l'annotation des discours autonomiques entre guillemets témoigne également d'une fréquence plus élevée de ce procédé de représentation des mots d'autrui dans « Le Figaro » (cf. 4.1.3), alors que « La Croix » ne se différencie pas des autres quotidiens à cet égard. On peut ainsi supposer que la position éloignée de « La Croix » dans l'AFC est déterminée par d'autres facteurs, dont les spécificités de sa ligne éditoriale, qui donne plus d'espace que d'autres journaux aux questions et aux avis religieux.

4.3.2.2. Sources énonciatives : types, distribution et points de vue dans PN-OP

Compte tenu de la nature pragmatico-énonciative et de la visée communicative des articles d'opinion, un aspect significatif à prendre en considération dans l'analyse de ces discours est la typologie des instances énonciatives qui s'y expriment. La répartition quantitative des instances énonciatives dans le sous-corpus PN-OP et l'étude qualitative des positions qu'elles défendent peuvent en effet servir de contre-essai utile pour confirmer, intégrer ou revoir les positionnements discursifs auxquels les différents journaux semblent se rattacher d'après l'analyse des articles d'information.

Pour favoriser une exploration contrastive entre les articles d'opinion au niveau de l'hétérogénéité métaénonciatrice, ceux-ci ont été regroupés en trois catégories distinctes selon la nature de l'énonciateur premier (L_1/E_1 – l'auteur de l'article), de l'éventuel énonciateur second (e_2 – en particulier la personne interviewée lors d'entretiens ou de témoignages), ainsi qu'en fonction du degré d'engagement de l'auteur de l'article. Ce dernier peut soit correspondre à l'instance énonciative et source principale de l'événement commenté/provoqué, soit jouer le rôle de médiateur des propos d'un énonciateur second, comme c'est le cas dans une interview. En prenant appui sur la classification proposée par Charaudeau (2011) (cf. aussi 3.1.2.2), nous distinguons avant tout deux types d'ins-

tances énonciatives principales qui peuvent assumer la fonction d'énonciateur premier : les sources internes au journal et les sources externes. Nous considérons les articles écrits par des journalistes appartenant à la rédaction du journal ou y collaborant de manière constante (comme c'est le cas des chroniqueurs) comme des textes produits par des sources internes au journal. Les articles dont l'énonciateur premier n'est ni un journaliste du quotidien ni un collaborateur habituel sont envisagés en revanche comme des textes rédigés par des sources externes, à l'instar des articles qui paraissent dans le journal sous forme de tribunes ou de blogs comme dans « La Croix ».

En ce qui concerne les articles relevant d'instances énonciatives internes au journal, une distinction supplémentaire s'impose néanmoins quant à la nature du point de vue principalement véhiculé. Nous séparons ainsi les articles qui favorisent l'expression de l'opinion du journaliste lui-même en tant que L1/E1, tels que les éditoriaux, les analyses et les chroniques, des articles où le journaliste met en discours les propos d'un énonciateur second, à l'instar des interviews ou des témoignages. Ces articles mettent donc en exergue le point de vue d'une source externe au journal qui est toutefois médié par l'auteur et énonciateur premier de l'article.

Les articles d'opinion ont été finalement regroupés selon trois types de sources : sources internes, sources externes médiées par le journaliste et sources externes. La figure 22 en donne une illustration synthétique.

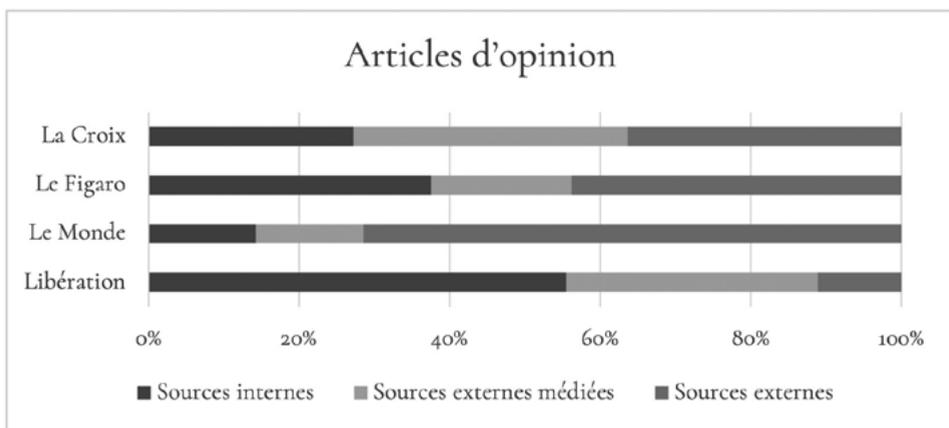


Figure 22. Classification des sources énonciatives principales des articles d'opinion.

Sur la base de cette tripartition, il est tout d'abord intéressant de constater que la plupart des articles d'opinion du sous-corpus PN du « Monde » ont été rédigés par des sources externes. Dans cette sous-partie, on peut en effet repérer cinq tribunes qui procèdent d'une variété d'instances énonciatives, dont la plupart sont des spécialistes du domaine médical. Quant aux mises en perspective de l'événement et aux prises de position qui en ressortent, quatre tribunes s'opposent à la plupart des arguments et des constructions discursives caractérisant les discours pro-maintien. Dans plusieurs tribunes, la représentation passive de la mort est abordée en relation avec les questions de la dignité des personnes en fin de vie et du respect de la volonté du patient, sans pour autant manquer de souligner la complexité de la décision à prendre et les enjeux éthiques qui y sont associés. Par ailleurs, la mort peut être abordée en relation avec la question plus large de l'euthanasie, dont l'acceptation peut également être analysée et clarifiée, comme c'est le cas dans l'extrait suivant :

- (10) Il ne s'agissait pas d'euthanasie, au sens d'une mort sciemment demandée par quelqu'un qui en est encore loin mais qui se sait atteint d'une maladie incurable (« Le Monde », tribune, Fournier, 11/07/2019, PN-OP).

En bref, c'est un positionnement pro-arrêt qui semble se dégager des procédés discursifs mobilisés dans ces tribunes et que l'on peut repérer aussi dans le seul éditorial de cette sous-partie. Il s'agit toutefois d'un positionnement relatif à ce cas spécifique, qui n'implique aucune généralisation. On retrouve par ailleurs des prises de position contraires à l'euthanasie ainsi que la défense de la loi Claeys-Leonetti. Néanmoins, seuls deux des sept articles analysés ne se rattachent pas à ce positionnement quantitativement dominant.

« Libération » est en revanche le journal qui présente le plus grand nombre d'articles d'opinion rédigés par des sources internes, dont deux chroniques, une analyse, un éditorial et un billet. Si quelques similitudes dans la construction des objets de discours avaient été identifiées avec « Le Monde » lors de l'analyse des articles d'information, les articles d'opinion de « Libération » mettent en relief une divergence majeure concernant la prise de position face à la loi Claeys-Leonetti. Contrairement au traitement qui caractérise le sous-corpus du « Monde », dans « Libération », cette loi est souvent remise en question, tant dans les articles d'information (de façon plus mitigée) que dans ceux d'opinion (de manière plus ouverte). En l'occur-

rence, c'est plutôt une législation plus avancée en matière de fin de vie qui est visée. Par exemple, le caractère équilibré de la loi est considéré comme une absence de prise de position tranchée sur des aspects clés de la fin de vie, comme en donne une illustration le passage suivant tiré de la chronique de la philosophe et psychanalyste Sabine Prokhoris :

- (11) Encore une aberration de la peu courageuse loi Claeys-Leonetti, une loi en forme de « ni, ni », toute d'évitements (« Libération », 06/06/2019, chronique, Prokhoris, PN-OP).

On peut en effet remarquer l'axiologisation négative véhiculée ouvertement par les éléments introduisant la loi à travers le substantif « aberration » et l'adjectif « peu courageuse », critiquant ensuite la nature floue qui la caractériserait.

Quant aux sources externes du sous-corpus PN-OP de « Libération », des spécialistes des domaines médical et juridique sont sollicités dans le cadre de deux entretiens pour approfondir les aspects les plus techniques de l'affaire et pour éclaircir certains concepts clés. C'est par exemple le cas de la catégorisation de l'état de santé de VL en termes d'« état pauci-relationnel », dont les précisions fournies signalent une prise de distance de la part de « Libération », comme cela avait déjà été relevé dans les articles d'information du même journal.

La prédominance de sources internes au journal qui véhiculent une prise de position assez ouverte favorise une mise en perspective et un positionnement déterminé sur la question. Par ailleurs, une perspectivation spécifique peut également ressortir de la façon dont les sources énonciatives impliquées – et/ou leurs propos – sont mises en discours et qualifiées. Dans l'exemple suivant (12), on peut remarquer que la mise en discours de la position de l'instance religieuse contribue à la connoter négativement. La journaliste signale d'abord l'évolution de la position défendue : elle était initialement « assez mesurée », puis elle s'est « durci[e] », ce qui présuppose un processus de polarisation de plus en plus marqué. Ensuite, le choix de l'adjectif à valeur péjorative « ultraconservatrice » pour qualifier la « vision » de l'archevêque participe à son tour à une mise en perspective négative de l'une des positions s'affrontant sur la question, orientant ainsi la perception de celle-ci.

- (12) Assez mesurées au départ, les instances catholiques ont peu à peu durci leurs positions sur la fin de vie au fil de l'affaire Lambert. [...] L'ar-

chevêque de Paris Michel Aupetit semble désormais imposer sa vision ultraconservatrice [...] (« Libération », analyse, Sauvaget, 23/06/2019, PN-OP).

Seul un article se distingue du reste du sous-corpus et se caractérise par la présence de certains aspects mis en relief davantage par d'autres journaux.

En ce qui concerne le positionnement opposé, c'est le sous-corpus de « La Croix » qui comporte de nombreux articles donnant la parole à des sources externes ou externes médiées. Deux sources ont été particulièrement sollicitées : des représentants de l'institution religieuse et des personnes travaillant dans le domaine de la fin de vie. À cela, on ajoute aussi deux articles de blogue associés à « La Croix » et signés par Daniel Bounoux, professeur émérite à l'Université Stendhal de Grenoble, agrégé de philosophie.

Différents avis du monde catholique et d'experts de questions éthiques font l'objet de ces articles d'opinion et participent à la mise en évidence de la complexité de la fin de vie, en soulignant parfois l'impossibilité de trancher sur la question. La plupart des sources externes convoquées défendent une position pro-maintien, bien que cela ne soit nécessairement pas la règle. Si la dimension humaine reste un aspect mis en avant dans ces discours, la singularité de chaque cas en constitue un deuxième trait caractéristique. Par ailleurs, le cas d'actualité peut également constituer le point de départ d'une réflexion plus générale sur la vie, en s'interrogeant sur ce que l'on peut considérer comme une vie, mais aussi sur la conception de sa propre vie et des conditions de son déroulement, comme le résume cette affirmation tirée de la tribune de Baudoin de Beauvais, aumônier d'hôpital.

- (13) La question qui m'habite n'est pas de vivre coûte que coûte, mais de vivre toujours mieux mes relations (« La Croix », tribune, de Beauvais, 21/05/2019, PN-OP).

Les articles d'instances internes au journal, qui s'expriment dans une chronique et deux éditoriaux, présentent en revanche une mise en perspective proche de celle repérée lors de l'analyse des articles d'information. Ils traitent en effet de la question de la valeur et de la dignité de la vie humaine, en soulignant à la fois l'unicité de chaque cas et l'importance de la rédaction de directives anticipées car « tout un chacun doit réfléchir sur ce qu'il souhaite pour lui-même » (« La Croix », 11/07/2019, éditorial, Goubert).

Parmi les quotidiens analysés, « Le Figaro » est celui qui a publié le plus de tribunes sur l'affaire VL au cours de la période considérée et selon les

critères de sélection adoptés. Des experts dans le domaine éthique ainsi que des représentants des domaines politique, philosophique, littéraire et religieux ont exprimé leur avis et leur position sur l'affaire et, plus en général, sur la fin de vie. Les propos de sources appartenant à ces mêmes domaines d'expertise ont par ailleurs été relayés dans des articles d'entretien. Deux éditoriaux et trois articles d'analyse ont enfin commenté l'affaire d'après la perspective d'instances internes.

La plupart de ces articles mettent en discours la question de la fin de vie en se rapprochant d'une position pro-maintien, comme c'est le cas pour la tribune signée par Damien Le Guay (publiée le 20/05/2019 et mise à jour le 03/07/2019). Le spécialiste des questions d'éthique de la mort, tout en s'opposant à l'arrêt des traitements de VL, en refuse l'assimilation à un « meurtre » ou encore à un « assassinat », ce qui reviendrait à « jeter l'opprobre sur les pratiques palliatives ». Dans la tribune de Emmanuel Hirsch (03/05/2019), l'arrêt des traitements de VL est en revanche qualifié d'« euthanasie consentie » et VL est considéré être « en état végétatif » et non pas en fin de vie (Hirsch) ou encore « en état pauci-relationnel », comme on peut le lire dans la tribune signée par Jean-Frédéric Poisson (14/05/2019). Dans les deux cas, on présuppose un état de maladie ou de handicap qui ne serait pas destiné à l'interruption des traitements, comme il est ressorti aussi de l'analyse des articles d'information. Cependant, il faut également signaler la présence de certains articles où le locuteur-énonciateur premier ou l'énonciateur second s'abstient de prendre position.

En conclusion, dans les quatre journaux, il est possible de rattacher la plupart des articles d'opinion à l'un des deux positionnements discursifs dominants, selon la polarisation relevée dans les articles d'information. Si les articles d'opinion se caractérisent par des formes axiologiques plus marquées et des prises de position plus ouvertes, c'est encore leur fréquence et leur distribution quantitative qui déterminent l'émergence de focalisations spécifiques et d'une perspectivation prédominante. En revanche, les types de sources privilégiés semblent influencer davantage sur la portée argumentative des points de vue exprimés et/ou représentés, notamment en fonction de l'autorité énonciative de l'instance convoquée, ainsi que sur le style discursif du journal. En particulier, la diversification des sources et le recours à des instances externes au journal pourraient être interprétés comme un effort d'une construction plus objectivée de l'événement à travers la multiplication des points de vue convoqués et non médiés par le « métaénonciateur » (Maingueneau 2021).

4.3.2.3. Perspectives principales selon l'hétérogénéité métaénonciatrice dans PN-OP

Si l'on revient au niveau global du sous-corpus PN-OP par une AFC (figure 23) réalisée à partir de la CHD illustrée plus haut (cf. figure 11, 4.1.2.1), il est possible de vérifier les perspectives différentes adoptées par les quatre sous-parties, confirmant par ailleurs les tendances déjà relevées dans les articles d'information.

L'AFC de la figure 23 affiche les formes les plus spécifiques des articles d'opinion des quatre journaux en fonction de la répartition des sous-parties attestée dans la figure 21. On peut ainsi remarquer que, sur le côté gauche, les formes les plus spécifiques des articles du « Monde » (mots en gris foncé au centre et au bas de la partie gauche du graphique) et de ceux de « Libération » (mots en gris clair dans les carrés à gauche) mettent en évidence

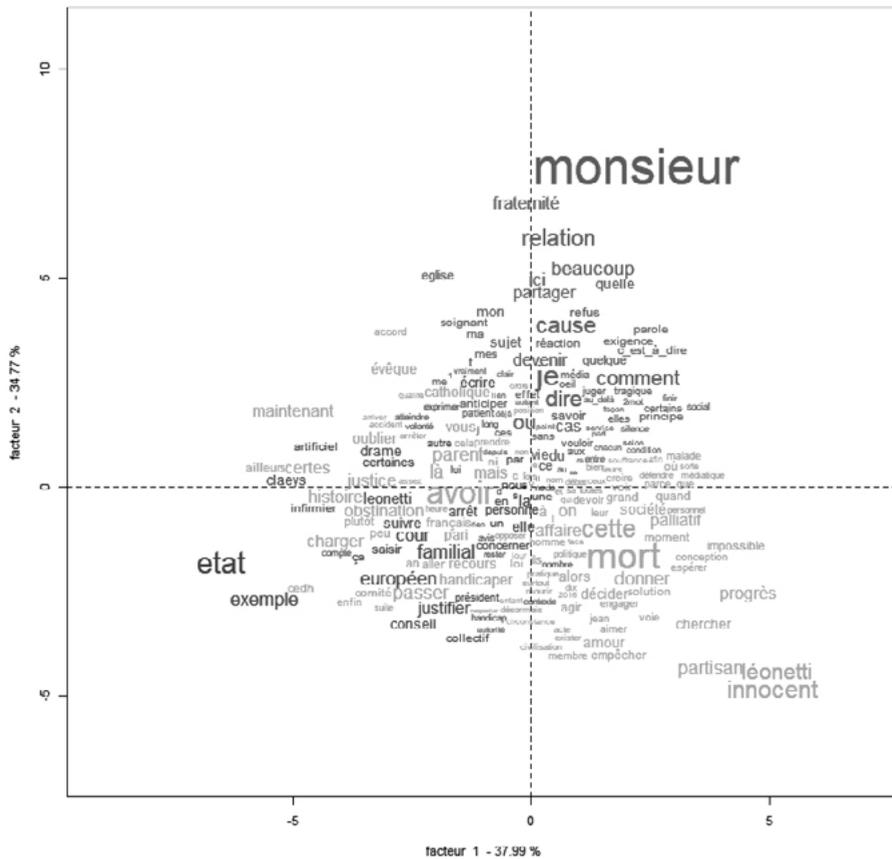


Figure 23. AFC sur le sous-corpus PN-OP.

les aspects socio-politiques et institutionnels de la fin de vie : la gestion de l'affaire de la part des institutions (*état, européen, cour*), le conflit familial (*drame, familial*), le cadre législatif et la procédure médicale ainsi que la valeur exemplaire assumée par ce cas. L'analyse qualitative de passages représentatifs identifiés par la CHD montre toutefois que les articles de « Libération » abordent plus en général la question de la fin de vie en souhaitant des avancées législatives, alors que les articles du « Monde » se concentrent plutôt sur les aspects de ce cas spécifique, en défendant la loi en vigueur.

Sur l'autre côté du graphique, le côté droit, l'objet du discours largement surreprésenté dans les articles d'opinion du « Figaro » (mots en gris clair dans le carré en bas à droite) est en revanche la mort en tant que sujet de réflexion général, mais aussi l'événement de la mort de VL, dont la mise en discours correspond à une prise de position ouverte.

Les articles d'opinion de « La Croix » (mots en gris foncé dans le carré supérieur à droite) se distinguent enfin du reste du sous-corpus PN-OP même dans une perspective énonciative par le suremploi du pronom personnel de la première personne du singulier. La dimension symbolique du cas de VL est particulièrement abordée, au point que l'on souligne un changement de statut : il ne s'agit plus d'une « affaire » mais il est devenu une « cause ». La dimension humaine de la question mise en avant dans les articles d'information est soulignée aussi dans les articles d'opinion.

L'analyse de la distribution des formes de référencement employées pour désigner VL et son état de santé dans PN-OP met finalement en évidence quelques convergences mais aussi des différences du point de vue sous-générique (cf. figure 24). D'une part, on constate des choix similaires entre les articles d'information et ceux d'opinion du « Monde » et de « Libération », qui continuent à privilégier la mise en relief de son statut socio-professionnel, à travers les expressions *ancien infirmier* et *ex-infirmier*. De l'autre, si « La Croix » a encore recours, dans les articles d'opinion, au terme *homme*, celui-ci n'est pourtant pas employé en cooccurrence avec l'adjectif *jeune*. Il est plutôt introduit soit par le démonstratif soit par l'article indéfini, en mettant dans les deux cas au centre l'humanité de VL. De même, les articles d'opinion du « Figaro » emploient davantage des dénominations pointant la dimension humaine et non pas le cadre médical et la condition de passivité mis en relief par le terme *patient*, suremployé en revanche dans ses articles d'information.

Tous les journaux emploient enfin le terme *patient* surtout dans sa valeur générique (37 occurrences sur 46). Il représente ainsi une forme de référencement minoritaire pour désigner VL dans les articles d'opinion, de même

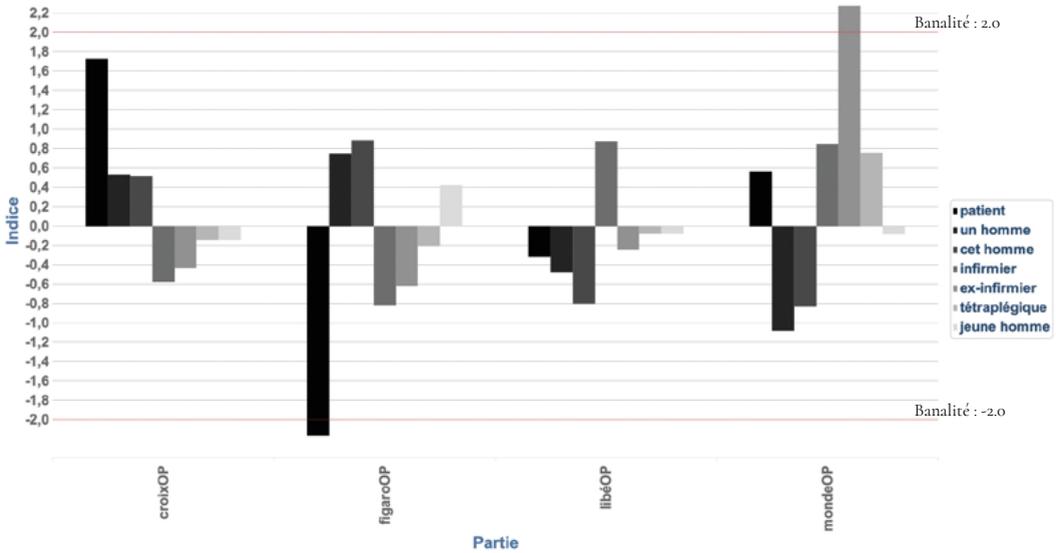


Figure 24. Spécificités sur les formes dans le sous-corpus PN-OP.

que *jeune homme* (1 occurrence) et *tétraplégique* (1 occurrence). Par contre, la forme *quadragénaire* est absente.

En bref, si des tendances constantes se révèlent au niveau de l'hétérogénéité sous-générique dans les cas du « Monde » et de « Libération », quelques variations caractérisent en revanche les partitionnements sous-génériques de « La Croix » et du « Figaro ». Ces variations se rattachent toutefois à un même positionnement discursif dominant, ayant ainsi une fonction de confirmation.

4.3.3. Variations sous-génériques et constantes métaénonciatrices

L'articulation des résultats de l'analyse du sous-corpus PN-INFO avec l'étude des articles d'opinion rassemblés dans le sous-corpus PN-OP permet globalement de valider les quatre positionnements discursifs dominants identifiés à partir des formes de référence et des constructions discursives principalement employées dans les articles d'information du corpus.

Les différences énonciatives et pragmatiques sur lesquelles repose la classification sous-générique des articles semblent ne pas influencer la mise en perspective de l'événement privilégiée par les divers journaux, bien qu'elle soit plus ou moins explicite selon les deux sous-genres. Plus précisément,

l'analyse inductive-quantifiée a montré que la variation intragénérique au niveau énonciatif-pragmatique concerne notamment le recours à une énonciation plus subjectivée dans les articles d'opinion, à travers des connecteurs argumentatifs et des ressources lexicales polarisées et connotées (dénominations et adjectifs). Cependant, ces mêmes caractéristiques ont émergé également de l'étude des discours autonymiques entre guillemets au sein des articles d'information, mettant ainsi en avant le rôle de ces procédés de représentation des propos d'autrui.

Les CHD effectuées sur les deux sous-corpus sous-génériques (PN-IN-FO et PN-OP) ont en outre permis de dégager des divergences dans les traitements discursifs des thématiques principalement abordées, en révélant également une variété thématique majeure dans les articles d'opinion. La plus grande variété de thèmes caractérisant ces derniers peut procéder de la subjectivité et de la diversité des sources énonciatives engagées, mais aussi des choix des journaux de concentrer l'attention et la réflexion sur certains aspects de l'affaire plutôt que sur d'autres. En effet, les articles d'opinion mettent en relation certains aspects-clés de l'affaire VL avec les questions sociales, éthiques et politiques d'intérêt public qui leur sont liées. Les avis de journalistes ou d'experts sélectionnés sont exposés en proposant en conséquence une perspective déterminée et inévitablement orientée sur les questions envisagées. La subjectivité est ainsi plus prééminente dans ces discours. En revanche, même si les articles d'information ne se limitent pas quant à eux à rendre compte des événements qui ont ponctué ce cas ayant longuement marqué l'actualité française, leur dimension narrative-factuelle favorise une certaine uniformité dans les thèmes abordés.

Au niveau global des dimensions lexico-sémantique et pragmatico-énonciative plusieurs variations intragénériques peuvent être repérées en fonction de l'hétérogénéité sous-générique des articles. D'autres éléments de rapprochement et de divergence ont ensuite émergé de l'étude de la construction discursive de l'événement et de la perspectivisation privilégiée, ainsi que, en conséquence, des positionnements discursifs auxquels les journaux sembleraient s'aligner. D'une part, la présence, dans les articles d'information, de propos rapportés entre guillemets peut comporter des éléments linguistiques plus subjectivés. Les caractéristiques lexicales et énonciatives des segments ainsi marqués s'apparentent alors à celles des articles d'opinion. Cela peut donc contribuer à orienter la mise en perspective de l'événement, tout en s'appuyant sur un style discursif qui donne une impression d'objectivité. D'autre part, le choix des mots à reproduire de manière autonymique, la sélection des instances énonciatives, ainsi que les modes de construction

discursive des deux objets de discours analysés représentent des procédés qui n'affichent pas ouvertement la subjectivité énonciative, mais la véhiculent de manière indirecte à travers leur perspective de construction de l'événement. En d'autres mots, ces procédés participent à la « dimension argumentative » (Amossy 2016) du discours. Il faut toutefois souligner qu'une analyse contrastive quantifiée favorise l'identification de ces divergences, qui seraient plus difficilement repérables par une étude exclusivement qualitative des articles d'un seul journal ou de plusieurs.

Enfin, l'analyse contrastive sous-générique a permis de confirmer les hypothèses concernant les variations métaénonciatrices formulées à partir de l'étude des articles d'information, qui ont constitué le point de départ de l'analyse en raison de leur nature plus factuelle et déclarative. L'identification des positionnements discursifs dominants a ainsi suivi une démarche allant des discours journalistiques dont l'énonciateur est plus « objectivisé » à ceux dont il est plus « subjectivisé », suivant la distinction proposée par Moirand (2007). Seule l'articulation de plusieurs niveaux d'analyse du corpus – du global au local pour revenir enfin au global – a permis d'identifier des éléments de divergence et de convergence qui ont trait à l'hétérogénéité aussi bien sous-générique que métaénonciatrice. Grâce à cette démarche, il a été finalement possible d'associer les quatre journaux deux à deux selon la polarisation de leurs positionnements discursifs, encore qu'avec des degrés différents.

Chapitre 5

Analyse du corpus de tweets « simples »

La constitution d'un corpus technodiscursif ne comportant que des tweets simples (TWS) favorise la comparaison intergénérique avec le corpus numérique de nature journalistique (PN). En effet, sur la base de la distinction précédemment proposée entre tweets simples et tweets augmentés (cf. 3.2.2.2), on peut considérer que les tweets simples sont plus autonomes, d'un point de vue à la fois sémantique et sémiotique, que les tweets augmentés.

Si trois types de tweets (originaux, réponses et threads) peuvent être produits sous forme de tweets simples selon la définition donnée plus haut, ceux-ci impliquent toutefois des statuts discursifs différents, déterminant à leur tour divers degrés d'autonomie sémantique (cf. 3.2.2.3). L'analyse comparative entre les trois pratiques vise donc à identifier les traits distinctifs de chacune dans le but de dégager aussi les possibles variations intragénériques dans la construction discursive de l'événement.

Cependant, il faut souligner que tout tweet, même composé d'un texte simple sans aucune forme technolangagière ou d'autres éléments d'augmentation, constitue un technodiscours (Paveau 2017), qui gagne à être appréhendé au sein de l'espace interconnecté dont il est natif. Comme précisé plus haut (3.3), en suivant Longhi (2020b, 2021), nous nous attacherons à observer la complexité technodiscursive des tweets à travers une démarche combinant une approche extractive de nature outillée, focalisée sur la matérialité langagière, et une autre plus contextualisante d'ordre qualitatif. Les interprétations des analyses textométriques impliqueront ainsi un retour constant à l'environnement natif interconnecté de ces productions, en articulant les dimensions micro- et macro-structurelle.

5.1. Aborder le niveau global du corpus : hétérogénéité pluri-technosémiotique et sous-technogénérique

Grâce à la classification des trois types de tweets simples par l'ajout de métadonnées, on peut avoir un aperçu de la constitution interne du corpus TWS. Comparée au corpus TW dans son ensemble (cf. 3.2.4), la distribution des trois pratiques de tweets simples confirme la présence marginale de threads (5 %), alors que la proportion de tweets-réponses est largement supérieure dans TWS (48 % contre 21 % dans TW). Par voie de conséquence, la proportion de tweets originaux simples (47 %) est en revanche inférieure à celle du corpus TW (68 %).

Deux remarques s'imposent à ce stade. D'une part, la plus grande proportion de tweets-réponses simples par rapport au corpus TW témoigne d'une tendance de textualisation significative propre à cette pratique technodiscursive. Le statut discursif de cette pratique à orientation réactive et interlocutive semble de fait influencer sa composition technodiscursive, qui implique rarement des éléments d'augmentation (OSA) (cf. aussi 3.2.4.2). De l'autre, on constate que peu de tweets simples ont été publiés par des comptes appartenant à des médias ou bien relaient des contenus journalistiques, comme on peut le remarquer en raison de la présence limitée de mentions « serviles » (Bigey & Simon 2018) de ces comptes. On peut ainsi supposer que la plupart d'entre eux constituent des discours de commentaire, visant à donner un avis sur l'affaire ou à réagir à un autre tweet, en s'alignant ou en se désalignant. Dans les deux cas, une prise de position est censée se dégager par rapport à l'objet de discours concerné, qu'il s'agisse d'un événement ou d'un autre tweet. Pour ces raisons, le corpus TWS peut représenter un échantillon approprié pour analyser différentes constructions discursives d'un même objet de discours.

5.1.1. Aperçus sémantique, énonciatif et technodiscursif

5.1.1.1. Thématiques principales

Suivant la démarche combinée déjà adoptée pour l'étude du corpus journalistique, les tweets simples regroupés dans TWS ont été tout d'abord soumis à une classification hiérarchique descendante sous *Iramuteq* (Ratinaud 2014, 2018). L'unité de segmentation sur laquelle repose la classification a été pourtant modifiée dans le but de prendre en compte les spécificités génériques du corpus et d'adapter l'analyse en conséquence. En effet, la CHD

a été réalisée non plus sur des segments de texte, comme c'était le cas pour les articles de journal (cf. 4.1.1.1), mais sur le tweet dans son entier en tant qu'unité textuelle à ne pas segmenter davantage, en raison de sa taille limitée. Ces choix différents de paramétrage de l'analyse sont donc subordonnés aux spécificités des deux macro-genres discursifs envisagés, qui restent néanmoins comparables dans une perspective intergénérique.

À la différence du corpus PN, la CHD dans TWS met en évidence une articulation et une diversification plus marquées des mondes lexicaux (Reinert 1993), comme l'illustre le dendrogramme ci-dessous (figure 25). Deux facteurs sont susceptibles d'influencer cette riche diversification thématique : la multiplicité des instances énonciatives qui se confrontent dans cet espace discursif, et la nature discursive plus proche du commentaire, qui comporte une pluralité de perspectives et points de vue, comme c'était le cas pour les articles d'opinion par rapport à ceux d'information (cf. 4.1.2).

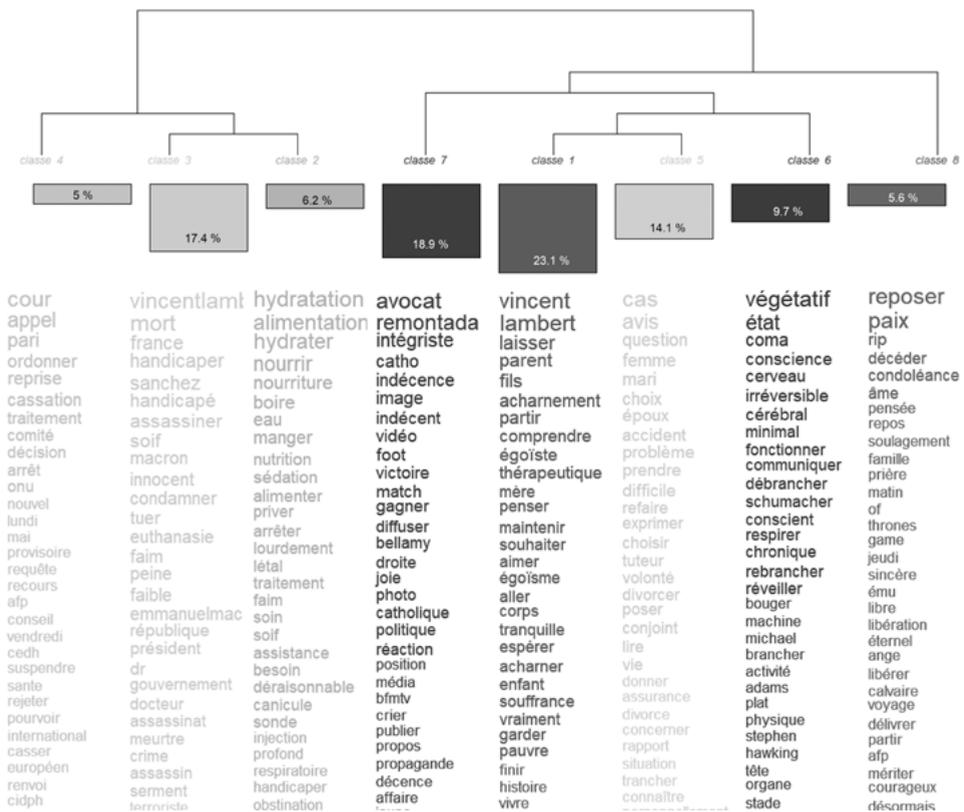


Figure 25. CHD dans le corpus TWS.

La CHD, ayant classé 90,52 % des tweets du corpus, se caractérise par deux ramifications principales. La ramification de gauche se compose de trois classes : elles concernent, respectivement, les décisions médicales et judiciaires qui ont scandé l'affaire (classe 4 – 5,01 %), ainsi que la mort de VL (classe 3 – 17,04 %), abordée en relation avec la question du statut de l'hydratation et de la nutrition artificielles (classe 2 – 6,2 %), comme en témoigne leur position dans le dendrogramme.

Cinq classes sont articulées au contraire dans la ramification droite. Les classes 7 (18,9 %) et 8 (5,6 %) rassemblent des tweets liés à deux événements spécifiques. D'une part, on repère des discours de commentaire relatifs aux déclarations des avocats des parents de VL, qui avaient comparé la reprise des traitements ordonnée par la Cour d'appel de Paris à une « remontada » ; de l'autre, des messages de condoléances publiés à la suite de la mort de VL.

L'affaire dans sa globalité ou dans ses aspects généraux est en revanche abordée dans les discours des classes 1, 5 et 6. Le conflit familial (classe 1 – 23,1 %) représente le monde lexical le plus présent, condensant la divergence polarisée des positions non seulement vis-à-vis de ce cas spécifique, mais aussi plus en général sur la question de la fin de vie. En l'occurrence, c'est la position des parents qui semble faire principalement l'objet des tweets de cette classe. Enfin, les enjeux médicaux, éthiques et affectifs posés par le cas de VL, les raisons de la prise de position de sa femme, ainsi que les implications sociales entraînées par la question de la fin de vie sont traités dans la classe 5 (14,1 %). Les tweets de la classe 6 (9,7 %) portent en revanche sur les conditions et l'état de santé de VL.

5.1.1.2. Répartition des thèmes en fonction des traits pragmatiques et énonciatifs des pratiques technodiscursives

Les différentes thématiques identifiées par la CHD ne se répartissent toutefois pas de manière uniforme selon les divers types de tweets, comme en témoigne le graphique suivant (figure 26).

Seules deux thématiques sont traitées de manière assez équilibrée dans les trois pratiques technodiscursives : le statut de l'alimentation et de l'hydratation artificielles (classe 2), ainsi que la déclaration de la « remontada » (classe 7), bien que cette dernière ait été abordée un peu plus dans des tweets originaux. En revanche, la thématique la plus largement débattue dans le corpus TWS, à savoir le conflit familial et la position des parents (classe 1), a plutôt fait l'objet de tweets originaux et de réponses. Des divergences plus marquées concernent les messages de condoléances (classe 8), qui ont

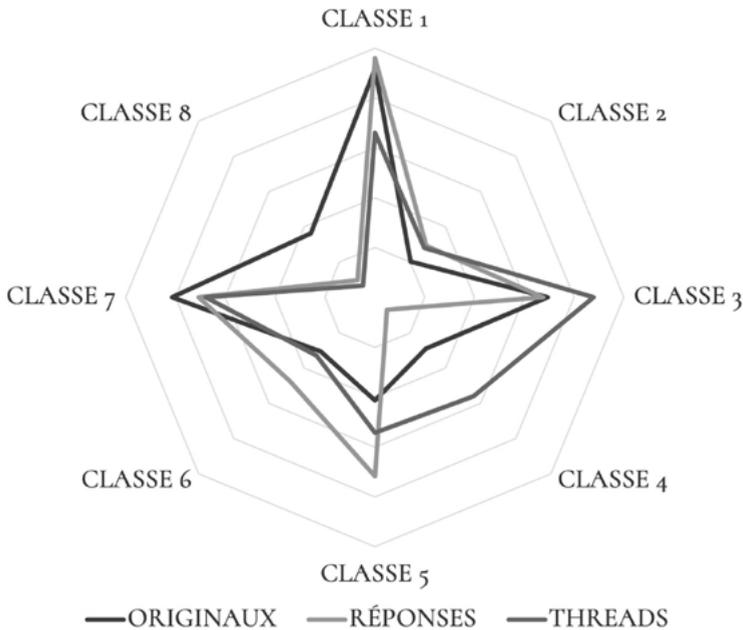


Figure 26. Répartition des divers types de tweets en fonction des classes de la CHD.

été exprimés surtout sous forme de tweets originaux. Par contre, les tweets de la classe 5 abordant la position de la femme de VL et les enjeux sociaux soulevés par ce cas d'actualité, ainsi que les tweets de la classe 6 portant sur la définition de l'état de VL ont davantage donné lieu à des tweets-réponses.

Ces résultats sont pourtant loin d'être surprenants. Nous estimons en effet que la répartition de ces thématiques reflète bien les spécificités énonciatives et pragmatiques incarnées par chaque pratique technodiscursive. D'une part, la nature interdiscursive des tweets originaux permet de répondre à la nécessité de partager l'expression des condoléances avec tous les internautes de la plateforme, sans engager une interaction interlocutive ciblée. De l'autre, il va de soi que des sujets aux contours plus polémiques tels que ceux abordés par les classes 5 et 6 peuvent susciter plus fréquemment une confrontation et aussi un conflit d'opinions, faisant ainsi l'objet d'une pratique interlocutive et réactive telle que les réponses.

Quant aux threads, on constate que l'événement de la mort de VL (classe 3) ainsi que les aspects les plus factuels de la question (classe 4), y ont été traités plus largement. En particulier, les caractéristiques propres à cette pratique peuvent favoriser un discours narratif, susceptible de reparcourir les différents moments et événements de l'affaire. On peut ainsi supposer

que les discours à orientation informative se concentrent notamment dans les threads et les tweets originaux, comme le suggère d'ailleurs la proportion des trois pratiques dans la classe 7 relative aux aspects factuels de ce cas d'actualité.

En bref, il ressort que les spécificités des différentes pratiques, notamment celles d'ordre énonciatif-pragmatique, déterminent au moins partiellement la nature discursive, les sujets thématiques abordés et les choix lexicaux.

5.1.2. *Variations intragénériques d'ordre lexico-sémantique et technolangagier*

La réalisation d'une AFC sur les formes des trois pratiques (figure 27) permet tout d'abord de vérifier, sur le plan lexico-sémantique, les similitudes et les divergences intragénériques ressorties lors de la CHD (figure 25). Étant donné que le corpus TWS implique une tripartition interne en fonction des types de tweets et que l'AFC peut être lancée sur un corpus composé d'au moins quatre sous-parties, la distinction interne à deux pratiques a permis de satisfaire cette contrainte technique. Ainsi, les tweets qui ouvrent le thread (threadSTART) ont été séparés de ceux qui en constituent la suite (threadFWD), et on a fait de même pour les réponses (reply) en distinguant les prolongements d'une réponse (replyFWD).

Au niveau lexical, les threads sont la pratique qui s'éloigne le plus du centre des axes, lequel représente les aspects les plus partagés par les sous-parties du corpus. Cela pourrait donc confirmer, d'un point de vue inductif-quantifié, la nature profondément distincte des threads par rapport aux deux autres pratiques, en conséquence de leur constitution séquentielle et de leurs visées discursives, ainsi que le souligne la caractérisation avancée par Longhi (2022).

Néanmoins, les divergences majeures concernent la pratique des réponses, qui s'opposent sur le plan cartésien à la fois aux tweets originaux et aux threads. On peut se demander si cette opposition principale ne relève pas de leurs statuts discursifs différents, apparentant les tweets originaux et les threads en tant que pratiques initiatives et interdiscursives, et les opposant en même temps aux réponses en vertu de la nature réactive-interlocutive de ces dernières.

Afin d'identifier les ressources linguistiques qui caractérisent le mieux les différentes pratiques technodiscursives, notre attention s'est portée sur les formes les plus spécifiques de chacune, analysées dans leur usage

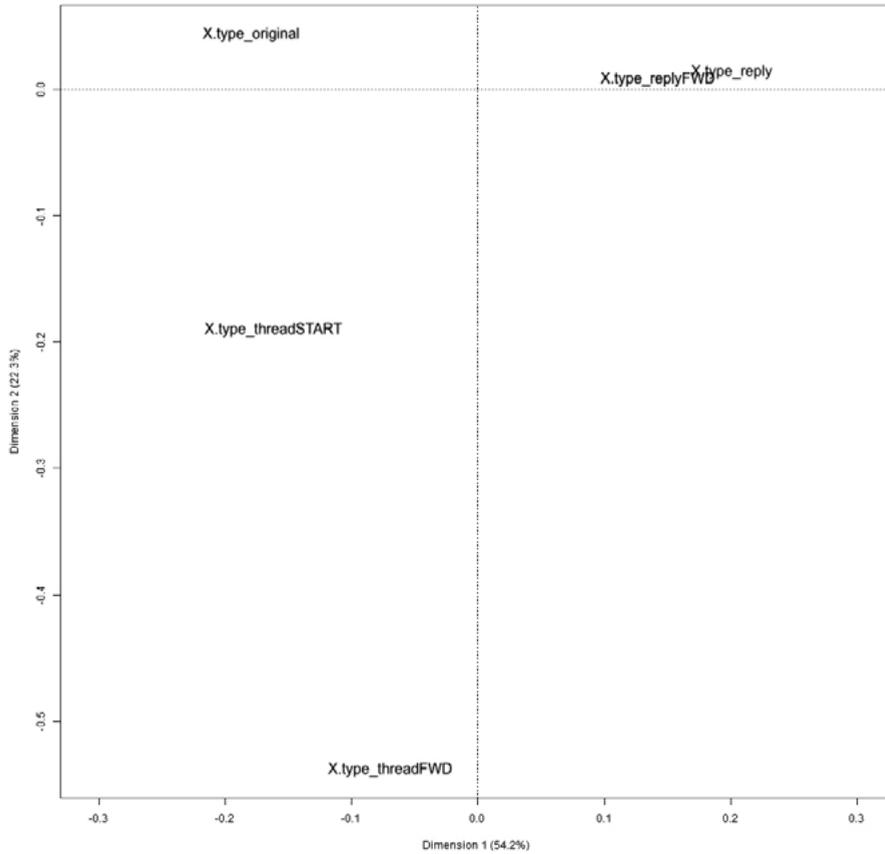


Figure 27. AFC sur les formes dans TWS.

en contexte. Dans le cas des tweets-réponses, l'une des formes les plus spécifiques correspond au pronom de la deuxième personne du pluriel, qui signale la nature distinctement interlocutive de la pratique. La dimension énonciative-pragmatique joue donc un rôle de premier plan dans sa caractérisation sous-générique.

L'analyse des cooccurrences et des concordances de *vous* sous TXM (Heiden *et al.* 2010) clarifie mieux les usages discursifs du pronom ainsi que les structures syntaxiques les plus récurrentes. Son emploi fréquent en relation avec des verbes déclaratifs (*dire, parler, affirmer*) et épistémiques (*savoir, connaître, penser, croire*) sert souvent à marquer le dissensus par rapport au tweet auquel la réponse est associée, ou encore à le réfuter, en contestant le bien-fondé des contenus énoncés et/ou son énonciateur. La mise en dis-

cussion du tweet auquel on répond se réalise en outre par le recours à la structure interrogative au moyen de l'inversion verbe-sujet ou d'éléments interrogatifs tels que *si*, *comment*, *que* et *qui*. On repère également l'emploi du pronom *vous* en cooccurrence avec la conjonction adversative *mais* et avec le connecteur *donc*, les deux étant exploités dans une visée argumentative. L'expression du dissensus et des procédés d'enchaînement logique avec fonction argumentative sont en effet fréquents au sein des réponses. Les exemples¹ suivants illustrent des extraits de tweets-réponses qui incluent quelques-uns des traits distinctifs mentionnés ci-dessus.

- (1) [mention du compte] Vous rendez-vous compte de l'indécence de votre question ? Vous ne savez pas si Vincent Lambert mérite de vivre ? [...] (Tweet-réponse, 21/05/2019).
- (2) [plusieurs mentions de comptes] Mais vous plaisantez ? ce sont les machines qui le tiennent en vie, pas son « refus de mourir » [...] (Tweet-réponse, 01/07/2019).

Comme on peut le remarquer dans ces exemples aussi, la négation constitue un autre élément récurrent qui est spécifique de la pratique des réponses. Par ailleurs, elle assume souvent une valeur polémique (cf. Ducrot 1984) car elle convoque l'énoncé que l'on veut réfuter à travers la forme affirmative présumée dans la négation. Les tweets-réponses impliquent ainsi une nature à la fois dialogale et dialogique. La dimension dialogale résulte de la relation interlocutive qui s'établit entre le tweet-réponse et le tweet ciblé, grâce aussi à la présence des mentions. La dimension dialogique découle en l'occurrence de la négation, qui convoque dialogiquement le contre-discours tout en le réfutant, comme le montre l'exemple ci-dessous.

- (3) @ADMDFRANCE @JeanLucRomero @BFMTV Primo : vincent Lambert n'est pas malade et n'est pas soigné, il est handicapé et était seulement nourri. Secundo : il n'est pas en fin de vie. Tertio : il n'a pas exprimé de choix et ne le peut pas (Tweet-réponse, 09/07/2019).

1. Dans la mesure où il s'agit de tweets « simples » qui, selon la définition proposée, ne se composent que de matériaux langagiers et technolangagiers, une transcription de la partie textuelle a été privilégiée, sans pour autant trahir les objectifs de l'étude. Des éléments de contextualisation sont ainsi fournis entre parenthèses, tels que la pratique technodiscursive et la date de publication du tweet. Les tweets et les extraits de tweets ont été anonymisés, ainsi que les mentions de comptes n'appartenant pas à des institutions, à des personnalités ou à d'autres entités publiques.

La négation représente de fait l'une des ressources privilégiées dans les réponses pour exprimer un dissensus total ou partiel. Elle témoigne de leur caractère polémique en tant que pratique d'échange entre opinions diverses, ce qui avait été déjà remarqué par Julliard (2016 : 154-155). Les analyses inductives-quantifiées et déductives-qualitatives sur le corpus TWS confirment cette observation, en attestant la surreprésentation de divers marqueurs responsables de la mise en discours du contraste d'opinions par rapport aux autres types de tweets (cf. figure 28). Bien que l'on repère aussi des tweets-réponses qui s'alignent sur le tweet auquel ils réagissent, l'expression de l'accord reste toutefois limitée dans cette pratique.

Le suremploi des adverbes d'affirmation et de négation *oui* et *non* contribue à son tour à la nature polémique et dissensuelle des réponses. Si les deux sont attestés parfois en cooccurrence, dans bon nombre de cas, l'adverbe *oui* est toutefois utilisé au sein de constructions concessives, ainsi que le signale la cooccurrence privilégiée avec la conjonction *mais*. La structure prototypique de la concession prend en effet corps dans l'articulation de la réponse avec le tweet auquel on réagit, manifestant la dimension dialogique² marquée de ces technodiscours. En effet, la réponse s'aligne partiellement sur le tweet auquel on réagit avant d'introduire une discordance, comme on peut le voir dans l'exemple 4.

- (4) [nom du compte] @Nagui Oui, je sais tout ça.....mais il est depuis repris par @Nagui pour les candidats en difficulté.....De toute façon, rien à voir avec le drame de #vincentlambert !!! (Tweet-réponse, 21/05/2019).

Pour résumer, trois catégories d'éléments – le pronom *vous*, la négation et les adverbes *oui* et *non* – sont souvent suremployées au sein de tweets-réponses, comme l'illustre la figure 28. Contribuant à la nature polémique (au sens de Amossy 2014) de ces technodiscours, ces ressources jouent également un rôle significatif dans la construction de la position interdiscursive du positionnement de L₁/E₁, qui s'en sert pour marquer son accord ou, plus fréquemment, son désaccord avec le tweet auquel il réagit.

En ce qui concerne en revanche les tweets originaux, le calcul des spécificités permet de dégager une tendance technodiscursive distinctive de cette pratique : le suremploi du hashtag et, notamment, de la forme #VincentLam-

2. Pour plus de détails sur les relations entre concession et dialogisme, voir l'article de Bres (2019).

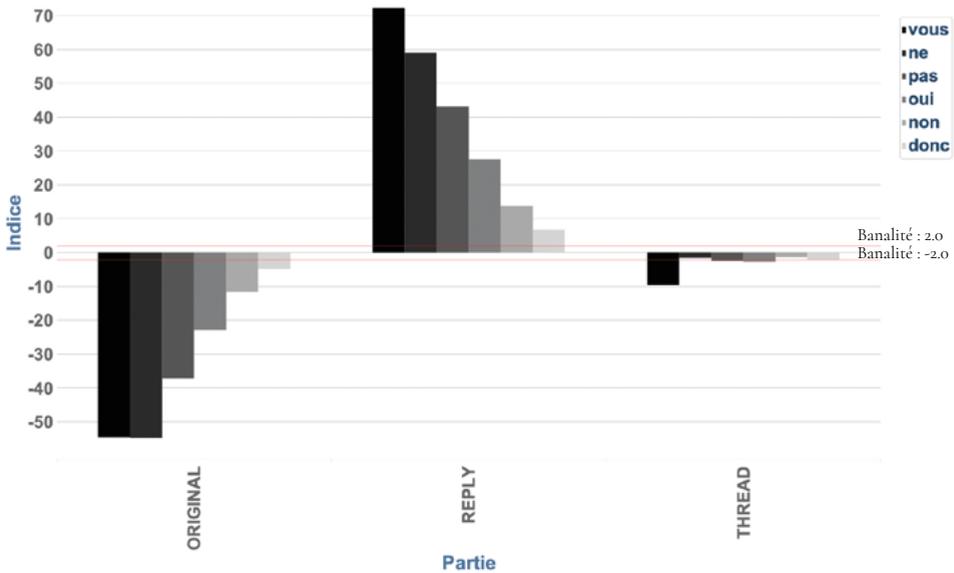


Figure 28. *Spécificités sur les formes vous, ne, pas, oui, non et donc dans TWS.*

bert ou de ses variantes typographiques³ (cf. plus loin figure 30). De fait, plus de la moitié des tweets originaux simples (62 %) incluent au moins un hashtag. Celui-ci les inscrit matériellement dans un fil discursif déterminé et les met en relation interdiscursive avec les autres tweets qui l'incluent à leur tour. Dans le cas spécifique du hashtag *#VincentLambert* et de ses variantes, qu'il soit en position de thématisation, intégré (ex. 5) ou non à la syntaxe de la phrase, il remplit notamment une fonction d'indexation et d'affiliation (Zappavigna 2011 ; Paveau 2017 ; Mercier 2013, 2018 ; Jackiewicz & Vidak 2014 ; Julliard 2016), dans la mesure où il matérialise un réseau interdiscursif thématique, impliquant la confrontation d'opinions et de positions différentes.

(5) L'ONU sur requête des parents de *#VincentLambert* via son comité de protection des handicapés enjoint à la France de ne pas débrancher Vincent. Mais *#VincentLambert* n'est pas un handicapé [...] (Tweet original, 05/05/2019).

De fait, une pluralité de hashtags d'indexation et de thématisation sont fréquents dans les tweets originaux, sans leur être nécessairement spéci-

3. Il est en effet possible de repérer des hashtags constitués de la même chaîne de caractères, qui diffèrent toutefois quant à l'emploi des minuscules et majuscules, comme c'est le cas pour *#VincentLAMBERT*, *#vincentlambert*, etc.

fiques. Dans le premier cas, on repère des hashtags qui renvoient notamment à des instances journalistiques ou à des chaînes et à des programmes télévisés (*#BourdinDirect*, *#cnews*, *#AFP*, *#cestdanslair*), de même qu'à des institutions impliquées dans l'affaire (*#ONU*, *#CEDH*). Dans le deuxième, on constate la présence de hashtags qui favorisent l'encadrement thématique du tweet (*#euthanasie*, *#FinDeVie*, *#santé*), alors que d'autres mettent en évidence des aspects ou des événements particuliers associés à la question par rapport auxquels on prend position, à l'instar des hashtags *#Remontada* et *#DirectivesAnticipées*. Des hashtags de positionnement, qui véhiculent de manière explicite une prise de position, sont également présents, comme c'est le cas pour les hashtags *#PeineDeMort* et *#JesuisVincentLambert*.

À propos de l'emploi des hashtags dans les deux autres pratiques technodiscursives, on remarque, encore une fois, des aspects communs aux tweets originaux et aux threads, alors que les tweets-réponses se distinguent d'un point de vue aussi bien quantitatif que qualitatif. Environ deux tiers (69 %) des tweets appartenant à un thread incluent au moins un hashtag et surtout la forme *#VincentLambert* (610 sur 688 hashtags dans les threads), reflétant ainsi la tendance relevée également dans les tweets originaux.

Les tweets-réponses se distinguent en revanche des autres types de tweets dans la mesure où moins de la moitié (40 %) d'entre eux contiennent au minimum un hashtag. Bien que les mêmes catégories de hashtags repérées dans les tweets originaux soient présentes dans les tweets-réponses, dans ce dernier cas, ils sont à la fois moins nombreux et moins diversifiés. Au-delà de ceux qui renvoient à VL, dans les réponses, les hashtags sont employés pour signaler des personnalités publiques liées à l'affaire, telles que *#Bellamy*, *#Macron*, *#Leonetti*, pour encadrer la question notamment en termes d'*#euthanasie*, ainsi que pour prendre position, comme c'est le cas pour *#SoutienVincentLambert* et *#VincentVivra*, quoique les deux soient utilisés par peu de comptes. Tout en relevant parfois d'usages individuels mais très fréquents, les hashtags de positionnement sont quantitativement plus présents dans les tweets-réponses.

On remarque en effet que la structure formelle des hashtags récurrents dans les réponses inclut des éléments explicites qui marquent une position favorable ou contraire à l'objet de discours thématisé, conformément à la dimension polémique plus accentuée de cette pratique. D'ailleurs, il est courant que des hashtags de positionnement soient créés et utilisés lors de sujets controversés, puisqu'ils permettent d'exprimer l'orientation argumentative du message de manière condensée, en s'alignant en même temps avec les autres discours partisans de la même position. On peut citer à titre

d'exemple le hashtag *#JesuisVincentLambert*⁴, constituant un « écho dialogique » (Vion 2011) à *#jesuischarlie*. Ce dernier, devenu très populaire lors de l'attaque du 7 janvier 2015 à la rédaction de *Charlie Hebdo*, représente désormais une formule (au sens de Krieg-Planque 2009) qui est adaptée à divers types d'événements et mobilisée également hors de la plateforme. Comme le remarque Badouard (2016 : 195), les formules « Je suis X » ou encore « Nous sommes X » « permet[tent] aux membres du public de s'approprier l'événement et de dire leur émotion à la première personne ». Cela participe en outre à « la mise en commun d'une émotion ou d'une indignation [qui] relève toujours d'un processus d'agrégation de prises de parole isolées » (Badouard 2016 : 196).

Les hashtags formulés selon la structure *Je suis X* évoquent la charge émotionnelle et notamment le sentiment de solidarité impliqué dans la formule originale, visant ainsi l'agrégation des internautes autour de ce ressenti. Dans le cas du hashtag *#JesuisVincentLambert*, le fond dialogique activé par sa structure formelle contribue à l'expression aussi condensée qu'explicite d'une prise de position à la fois intradiscursive et interdiscursive. On remarque ainsi que les deux hashtags ne se distinguent pas seulement par une distribution différente dans les deux pratiques (figure 29), mais qu'ils remplissent aussi des fonctions diverses.

Il faut toutefois signaler que l'utilisation du hashtag *#JesuisVincentLambert* se concentre surtout dans les tweets-réponses de deux comptes, ce qui amène à s'interroger sur la question de la représentativité des discours du web 2.0. Des tendances significatives d'un point de vue quantitatif pourraient de fait être déterminées par la production massive de technodiscours de la part d'un nombre limité d'énonciateurs ou encore par l'intervention de bots. L'analyse comparative des hashtags les plus fréquents dans les deux pratiques semble néanmoins confirmer les emplois distinctifs de chacune.

À côté du hashtag de thématisation *#VincentLambert*, d'autres formes sur-employées dans les tweets originaux (figure 30) correspondent à des éléments déterminés de l'affaire en relation avec les mondes lexicaux principalement véhiculés par cette pratique. On repère ainsi des formes liées à l'expression des condoléances (*paix, enfin, repose, reposer*), d'autres qui pourraient relever d'une position favorable à l'arrêt des traitements (*laissez et*

4. Dans ce cas aussi, le hashtag subit des variations typographiques. Les différentes variantes ont été toutefois considérées comme relevant d'un seul hashtag, puisqu'elles n'en modifient pas la portée sémantique.

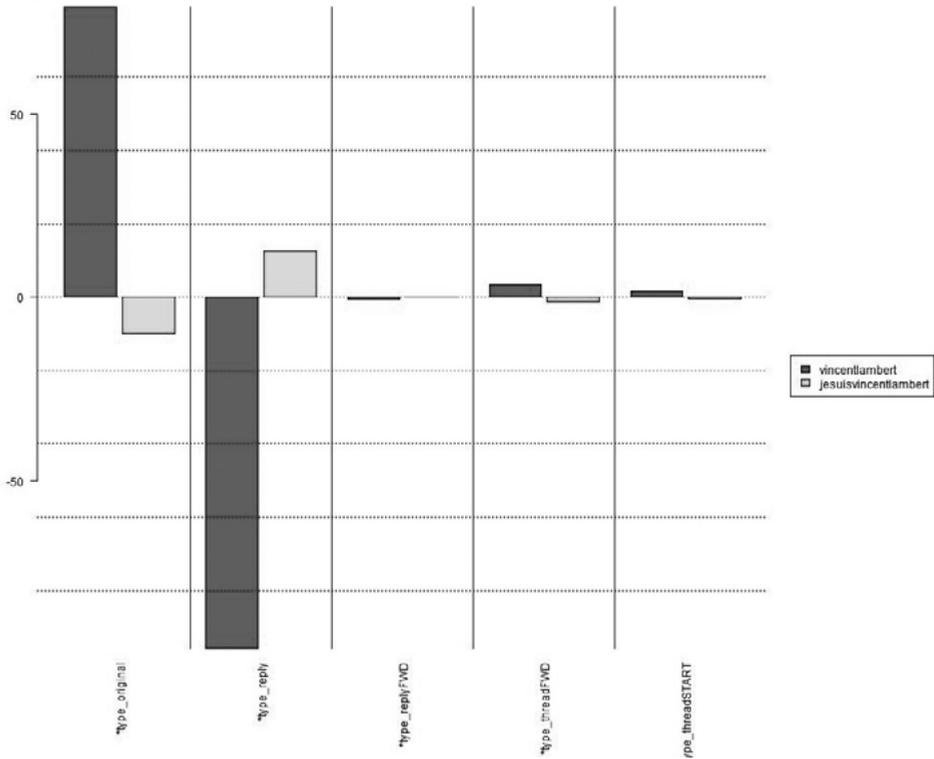


Figure 29. Spécificités sur #VincentLambert et #JesuisVincentLambert dans TWS.

partir, employées aussi en cooccurrence), tout comme des aspects factuels de l'événement (*matin, reprise, cours, paris, arrêt*). Parmi les formes les plus spécifiques, on remarque également beaucoup d'occurrences correspondant aux hashtags les plus fréquents dans cette pratique (*vincentlambert, cnews, bfmtv*), comme l'atteste le graphique de la figure 30.

Le calcul des spécificités des tweets originaux met en relief la centralité du recours au hashtag au sein de cette pratique, qui se caractérise également par la prédominance de certains mondes lexicaux, compte tenu de son statut (techno)discursif.

Par ailleurs, des rapprochements intragénériques ont été repérés entre les tweets originaux et les threads quant à l'usage des hashtags et notamment du hashtag de thématisation #*VincentLambert*, même si cela se produit dans des proportions différentes. Quoi qu'il en soit, les deux pratiques s'opposent nettement à celle du tweet-réponse, où le recours au hashtag en général et l'emploi spécifique de #*VincentLambert* sont largement sous-représentés (cf. figure 31).

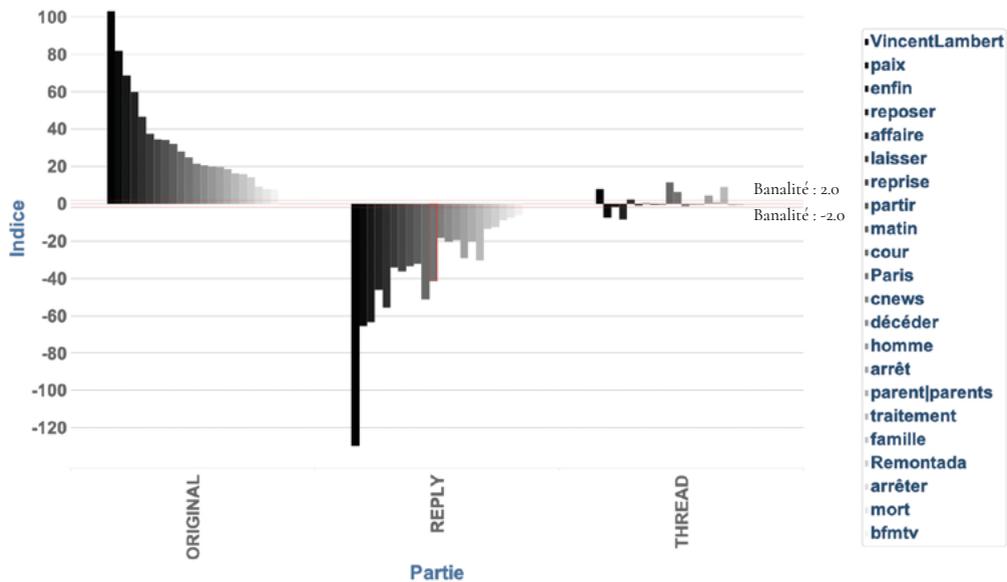


Figure 30. Spécificités sur les formes dans TWS.

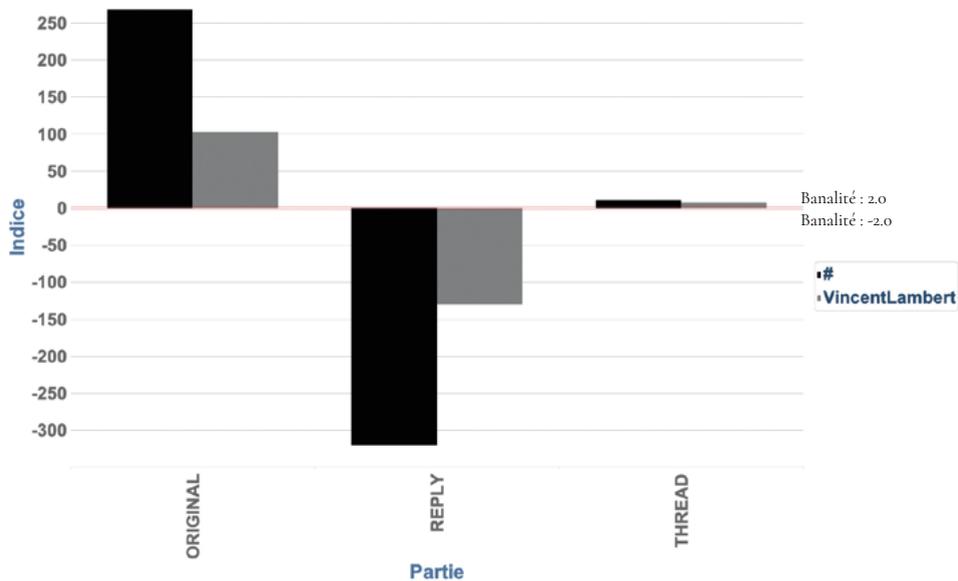


Figure 31. Spécificités sur l'emploi du hashtag en général (#) et de #VincentLambert.

Les formes surreprésentées dans les threads (figure 32) mettent enfin en évidence le statut singulier de cette pratique, où les formes les plus spécifiques correspondent à des signes typographiques, tout particulièrement les guillemets, et à des données numériques (identifiées par @card@ via TXM).

Le suremploi des guillemets est indicatif du ton narratif propre à cette pratique, qui se caractérise souvent par une visée explicative-informative, comme le montre Longhi (2022). Dans le fil discursif matérialisé par le thread, les guillemets sont employés pour marquer des discours rapportés, participant en outre au « processus d'événementialisation » (Charaudeau 2011) qui est typique du discours médiatique. En revanche, le suremploi de données numériques témoigne de la tendance à numéroter les différents tweets d'un même thread pour signaler leur enchaînement progressif. Le recours à la numérotation relèverait ainsi d'une planification du discours (Longhi 2022).

Au-delà des guillemets et des données numériques, les formes lexicales surreprésentées dans les threads (figure 32) rendent compte d'éléments factuels concernant différentes étapes de l'affaire et les acteurs engagés. Ces caractéristiques lexicales confirment donc à leur tour la prédominance de la dimension narrative-événementielle, qui est déjà ressortie de l'étude de la distribution des mondes lexicaux dans cette pratique.

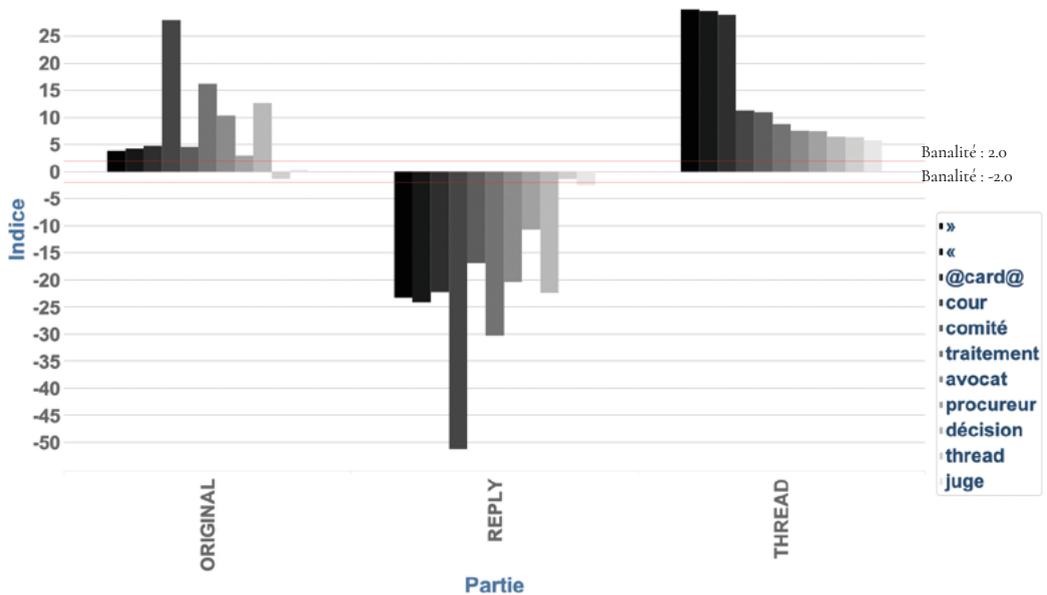


Figure 32. Spécificités sur les formes dans le corpus TWS.

Le calcul des spécificités effectué sur les trois pratiques a mis en évidence des traits distinctifs de nature lexico-sémantique (mots pleins et mondes lexicaux) et technodiscursive (l'emploi des hashtags) qui permettent de les différencier du point de vue intragénérique. Si chacune présente des caractéristiques spécifiques, des rapprochements ressortent toutefois entre les tweets originaux et les threads, qui s'opposent nettement aux réponses. Par ailleurs, des différences significatives au niveau énonciatif-pragmatique ont déjà été relevées, attestant l'orientation interlocutive et polémique marquée propre aux tweets-réponses. Des explorations supplémentaires s'avèrent néanmoins nécessaires pour mieux cerner les traits énonciatifs-pragmatiques des deux autres pratiques.

5.1.3. Variations intragénériques d'ordre énonciatif-pragmatique

Les rapprochements et les écarts entre les trois pratiques technodiscursives relevés lors de l'analyse des formes (5.1.2) sont confirmés par la distribution des catégories grammaticales. Comme on peut le voir dans la figure 33, les tweets originaux et les threads emploient davantage les prépositions (PRP), la ponctuation (PUN) et, notamment, les guillemets (PUN:CIT), mais dans des proportions différentes. En particulier, le recours à des modes autonymiques pour rapporter d'autres propos constitue un élément significatif de la construction dialogique des deux pratiques, contrairement aux tweets-réponses. En effet, ces derniers établissent matériellement une relation dialogique interlocutive avec le tweet auquel ils réagissent et par rapport auquel ils prennent souvent position. Les tweets originaux et les threads se servent en revanche de discours entre guillemets pour se positionner plutôt de manière intradiscursive ou interdiscursive, au-delà de l'emploi de ressources technolangagières spécifiques telles que les hashtags et les mentions.

Cependant, quelques différences se dégagent aussi des deux pratiques plus interdiscursives, à savoir les tweets originaux et les threads. Alors que le recours à des données chiffrées (NUM) se confirme parmi les éléments suremployés dans la pratique des threads, les substantifs (NOM) sont surreprésentés surtout dans les tweets originaux. Ce dernier élément peut être indicatif d'une richesse lexicale plus ample qui caractériserait les tweets originaux. De toute façon, il nous semble qu'aussi bien le recours aux chiffres que la multiplicité des substantifs employés sont déterminés par la nature technodiscursive des pratiques respectives. Si, comme on l'a déjà souligné, la numérotation des tweets appartenant à un thread clarifie leur enchaînement, le suremploi de substantifs dans les tweets originaux procède de leur caractère initiatif,

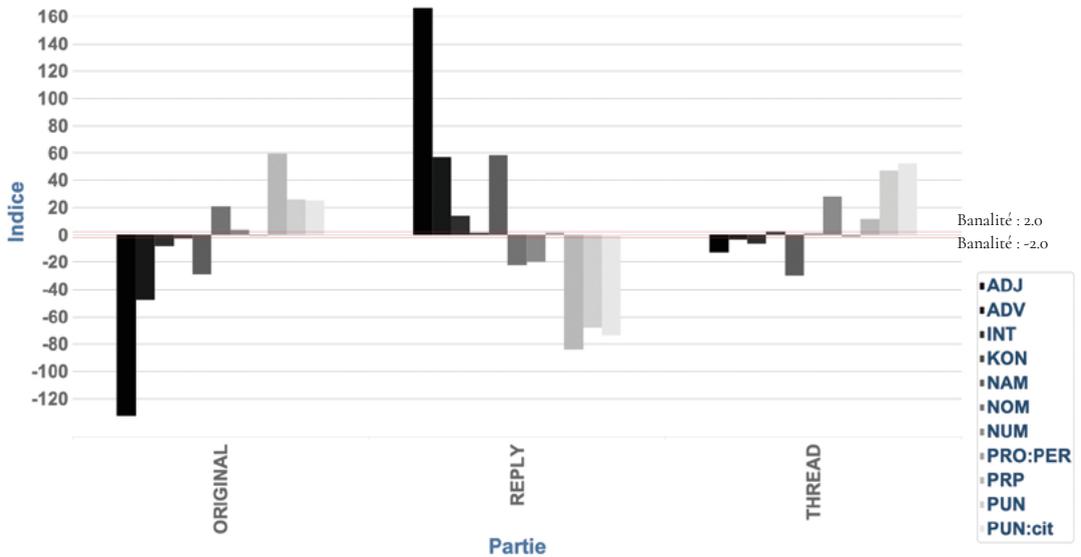


Figure 33. Spécificités sur les types des trois pratiques technodiscursives dans TWS.

qui est censé impliquer une construction discursive plus articulée que dans le cas des pratiques secondaires. Celles-ci peuvent en effet reprendre ou référer de manière elliptique à des éléments présents dans le technodiscours premier dont elles découlent, limitant ainsi l'utilisation de substantifs.

En ce qui concerne en revanche les tweets-réponses, la surreprésentation des adverbes (ADV) et notamment des adverbes de négation atteste le ton souvent polémique de ces technodiscours à orientation interlocutive, comme il est ressorti aussi des études de la section précédente. La dimension interlocutive de la pratique semble influencer également la présence d'éléments interactionnels tels que les interjections (INT), ainsi que celle de noms propres (NAM), souvent dans le cadre de mentions, soulignant ainsi leur nature réactive ciblée. Le calcul des spécificités montre enfin le suremploi de la catégorie des adjectifs (ADJ), qui peuvent rendre compte d'une dimension axiologique plus fréquente.

De manière analogue au corpus PN (cf. 4.1.1.2), l'étude de la distribution des pronoms personnels (figure 34) représente toutefois l'angle d'approche privilégié pour aborder les divergences intragénériques au niveau énonciatif-pragmatique. Comme anticipé plus haut (5.1.2), le large suremploi du pronom de la deuxième personne du pluriel dans les tweets-réponses, utilisé aussi bien dans la fonction de sujet que de complément, témoigne de leur

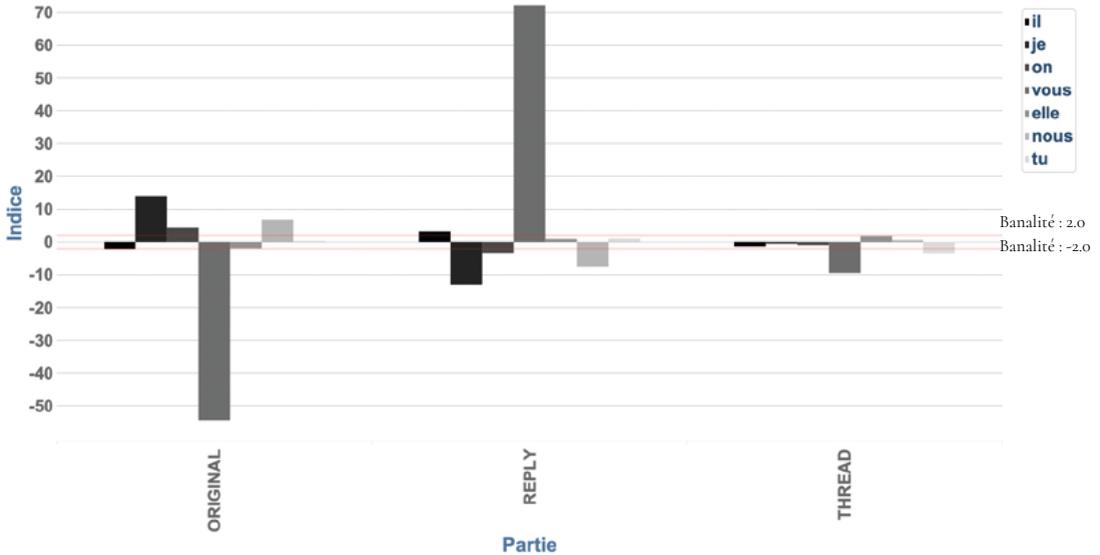


Figure 34. Spécificités sur les lemmes des pronoms personnels dans le corpus TWS.

dimension interlocutive très marquée. Au contraire, les tweets originaux se caractérisent par la prédominance d'une prise en charge énonciative à la première personne du singulier.

Le suremploi du pronom de la première personne du singulier constitue ainsi un trait énonciatif-pragmatique distinctif de la pratique des tweets originaux, de même que le suremploi des hashtags. Par ailleurs, ceux-ci sont utilisés en cooccurrence avec *je* dans 60 % des tweets qui y ont recours. On peut ainsi supposer que la prise en charge énonciative réalisée par le pronom *je* signale une prise de position voire une expérience individuelle au sein du réseau interdiscursif dans lequel le tweet est inscrit par le hashtag. Cela a donc des répercussions sur la construction du positionnement discursif dans les tweets originaux. Ils impliquent un positionnement individuel centré sur le *je*, qui met en avant la subjectivité de l'énonciateur en exprimant une prise de position individuelle. Celle-ci est toutefois mise en relation interdiscursive avec d'autres positionnements mentionnés dans le message, comme l'illustrent les exemples suivants.

- (6) Je ne comprend pas comment on peut être une mère et laisser son fils souffrir encore depuis plus de 10 ans, [...] #vincentlambert #Remontada (Tweet original, 21/05/2019).

- (7) J'espère de tout mon coeur et je prie de mon âme pour #vincentlambert se réveille et sorte de son état végétatif #VincentLambert #Soutien-VincentLambert (Tweet original, 21/05/2019).

Dans l'exemple 6, l'énonciateur prend de la distance par rapport à l'un des deux camps en opposition, en inscrivant son message dans un réseau thématique déterminé par le hashtag de thématisation #Remontada. Il s'agit en l'occurrence d'une forme de désalignement qui souligne la dimension interdiscursive de cette prise de position. En revanche, dans l'exemple 7, la prise en charge énonciative par le *je* met en avant la subjectivité même de l'énonciateur, sans impliquer un contre-discours par rapport auquel prendre position, mais recherchant plutôt un alignement interdiscursif. En l'occurrence, aussi bien la construction discursive que le hashtag de positionnement #SoutienVincentLambert semblent en effet se rattacher à un positionnement contraire à l'arrêt des traitements. Par contre, peu de tweets originaux incluant un *je* présentent également une mention pour s'adresser de manière ciblée à un compte et pour se positionner de manière interlocutive.

Les tweets originaux impliquent également un suremploi des pronoms *on* et *nous*, quoique moins élevé par rapport au pronom *je*. Cela signale le recours privilégié à une prise en charge énonciative qui met en avant la subjectivité de l'énonciateur de manière soit individuelle (*je*) soit plurielle (*on*, *nous*). On peut donc remarquer que les tweets originaux et les tweets-réponses attirent l'attention de manière complètement opposée sur les deux pôles de l'acte énonciatif.

À la différence des deux autres pratiques technodiscursives, les threads ne présentent pas de pronoms personnels largement suremployés, mais ils sous-emploient au contraire de manière significative les deux pronoms interlocutifs *vous* et *tu*. En général, les usages spécifiques des pronoms personnels dans les threads révèlent un style énonciatif qui tend à gommer les instances énonciatives de production. Cela va d'ailleurs de pair avec la visée explicative-informative (Longhi 2022) propre à la pratique. Moins surprenant est en revanche le sous-emploi des pronoms interlocutifs étant donné l'orientation principalement interdiscursive de la pratique.

Divers traits caractéristiques des trois pratiques technodiscursives sont ressortis des analyses inductives-quantifiées réalisées sur deux plans du niveau global du corpus – le plan lexico-sémantique et le plan énonciatif-pragmatique – qui sont également significatifs pour la construction du positionnement discursif (Cagninelli 2022). Néanmoins, des explorations

contextualisées sur des niveaux plus locaux ont été également nécessaires pour vérifier et approfondir les analyses globales. Ces résultats montrent que les trois types de tweets constituent des sous-genres technodiscursifs différents d'un point de vue lexico-sémantique, technolangagier et énonciatif-pragmatique. En d'autres termes, chacun des trois types de tweets simples mis en évidence peut être envisagé comme un sous-(techno)genre discursif spécifique puisqu'il procède d'une pratique technodiscursive déterminée et distincte des autres.

L'identification de ces spécificités sous-génériques est d'ailleurs fondamentale pour passer à l'analyse qualitative des ressources exploitées dans la construction des deux objets de discours envisagés, en adoptant une perspective, tout d'abord, intragénérique et, ensuite, intergénérique. L'étude doit être en effet réalisée en tenant compte des spécificités de chaque pratique. Cependant, certains éléments participant à la construction des positionnements dominants ont déjà émergé des analyses globales effectuées jusqu'à présent sur les différentes pratiques technodiscursives.

5.2. Analyse déductive et contextualisée des niveaux locaux : hétérogénéité sous-technogénérique

L'analyse inductive et quantifiée a permis de repérer divers traits intragénériques qui peuvent caractériser les trois pratiques indépendamment du sujet thématique abordé. L'analyse déductive et contextualisée vise en revanche à approfondir, confirmer ou infirmer ces observations, mais elle apporte également des précisions supplémentaires liées à la thématique qui favorise en l'occurrence la cohésion et cohérence du corpus pluri-générique. Afin d'assurer une comparaison contrastive intergénérique, la démarche suivie pour étudier en contexte les deux objets de discours sélectionnés (la désignation de VL et de son état de santé ainsi que la représentation de la mort) dans le corpus PN est adoptée aussi pour l'exploration déductive-contextualisée du corpus TWS, tout en prenant en compte ses propres spécificités génériques. En raison de leur plus grande autosuffisance sémantique, les tweets originaux constitueront le point de départ de notre analyse déductive-contextualisée. Nous aborderons ensuite les deux autres types de tweets simples de manière contrastive et identifierons ainsi les variations intragénériques.

Au-delà des divergences d'ordre numérique (Paveau 2015), la nature concise des tweets et le volume du corpus TWS représentent deux différences majeures à envisager dans la comparaison avec les articles de

journaux rassemblés dans le corpus PN. Si ces derniers impliquent une textualisation plus développée, qui se déploie sur plusieurs paragraphes et peut, en conséquence, mettre en relation divers objets de discours, la contrainte matérielle des tweets favorise en revanche une sorte de spécialisation thématique. Ainsi, chaque tweet porte plutôt sur un objet de discours déterminé.

En suivant une démarche analogue à celle adoptée pour l'analyse du corpus PN, l'étude déductive-contextualisée se penche, dans un premier temps, sur les formes de référenciation principalement employées pour l'identification des objets de discours sélectionnés et, dans un deuxième temps, sur les procédés respectifs de mise en discours. Dans le cadre limité du présent ouvrage, l'analyse se concentre uniquement sur les procédés de construction et de mise en discours des objets de discours sélectionnés à partir des classes thématiques qui rassemblent la plupart des tweets les abordant, identifiées par des CHD.

5.2.1. Construction des objets de discours : formes de référenciation principales

Dans la mesure où ils constituent souvent des (techno)discours de commentaire, les tweets simples du corpus TWS présentent des formes de référenciation dont la fonction principale est de mettre en perspective les objets de discours ainsi identifiés afin d'exprimer des considérations à leur propos. En particulier, pour référer au protagoniste de l'affaire, le recours à des expressions différant de son nom, même sous forme de hashtag, n'est pas aussi nécessaire que dans le discours journalistique où ces formes sont indispensables pour en fournir une description présentative ou bien pour effectuer des renvois anaphoriques et cataphoriques (cf. 4.2.1). On peut ainsi s'attendre que ces formes de référenciation, d'une part, soient quantitativement moins présentes et, de l'autre, qu'elles participent plus ouvertement à une mise en discours orientée de la question, en s'alignant sur l'un des deux positionnements discursifs polarisés.

L'analyse des spécificités effectuée dans le corpus TWS sur les formes principales de référenciation de VL repérées dans le corpus PN (*[jeune] homme, infirmier, patient*) (4.2.1.2) met en évidence une différence significative entre les deux sous-corpus génériques (PN et TWS), relevant de leurs finalités discursives diverses. Alors que le discours journalistique emploie largement – au-delà des différences métaénonciatrices – la forme *patient* en tant que terme spécialisé du domaine juridico-médical pour référer à VL,

c'est la forme *homme* qui est en revanche la plus fréquente dans les tweets. Quant à sa distribution dans les trois pratiques technodiscursives (figure 35), elle est suremployée dans les tweets originaux, de même que sa variante *jeune homme*, cette dernière étant en tout cas beaucoup moins fréquente dans le corpus.

À travers l'étude des cooccurrents du substantif *homme*, on observe avant tout l'usage de deux déterminants principaux, l'article indéfini et l'adjectif démonstratif, qui favorisent des mises en discours différentes. L'article indéfini permet aussi bien d'attirer l'attention sur la dimension humaine de la question et sur ses enjeux éthiques – de manière similaire à la tendance repérée dans « La Croix » – que d'opérer une généralisation du cas. Les technodiscours comportant cette cooccurrence proposent ainsi des commentaires à partir de l'affaire VL, tout en tendant à mettre en discours un positionnement relatif à la question plus générale de la fin de vie, comme l'illustre l'exemple ci-dessous.

- (8) Ce matin à Reims, un homme très gravement handicapé est mort de faim et de soif. Pourquoi? Vincent est mort mais l'affaire Lambert ne fait que commencer tant Les questions qu'elle pose sont au cœur de notre Humanité #VincentLambert (Tweet original, 11/07/2019).

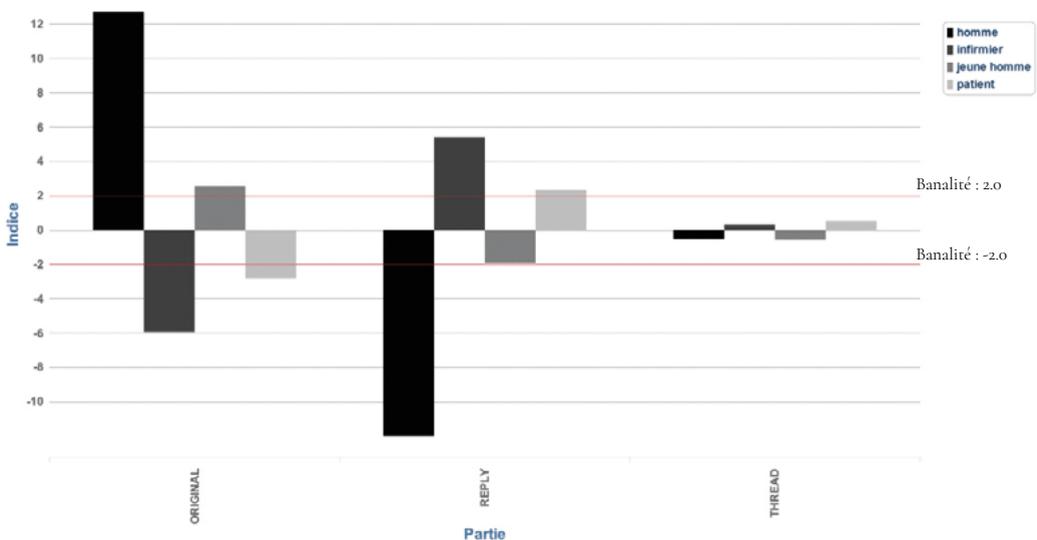


Figure 35. Spécificités sur les formes *homme*, *infirmier*, *jeune homme* et *patient* dans le corpus TWS.

Dans d'autres cas, les dimensions éthique et humaine de la question mises en relief par le substantif *homme* précédé de l'article indéfini sont exploitées pour engager le débat sur la manière dont il faudrait considérer l'existence d'un *homme* qui se trouve dans un état végétatif, en se demandant si cet état correspond encore à une forme de vie digne, comme le montre l'extrait du tweet-réponse suivant (ex. 9). De fait, ce n'est pas un hasard si cette mise en question fait l'objet d'un tweet-réponse, matérialisant une pratique techno-discursive caractérisée souvent par un ton polémique (cf. 5.1.2 et 5.1.3).

- (9) [...] Mais dans cette terrible et triste histoire, un homme (mais l'est-il encore ?) reste seul à subir ce que tout le monde décide pour lui : M. Vincent Lambert ! Mes pensées vont vers lui ! (Tweet-réponse, 21/05/2019).

Par contre, l'emploi de l'adjectif démonstratif, même s'il est moins fréquent que dans le cas d'usage précédent, signale des aspects spécifiques de cette situation particulière, faisant encore une fois écho à une autre tendance relevée dans les articles de « La Croix », qui soulignent l'importance d'envisager chaque situation de fin de vie dans sa singularité (TWR – [...] *Cet homme doit pouvoir aller dans un centre adapté à sa pathologie. Ne pas lui accorder ce droit est discriminant.* [...], 16/05/2019)⁵.

Parmi les autres cooccurrents les plus fréquents avec *homme*, on remarque également l'adjectif *pauvre* ainsi que d'autres formes participant à la mise en relation avec deux questions mobilisées par les deux positionnements : d'un côté, l'appel à le *laisser partir en paix*, en respectant sa *volonté* et en lui évitant de *souffrir*, de l'autre, l'appel à le *maintenir en vie* et à ne pas le laisser *mourir de faim* et de *soif*. En l'occurrence, le choix référentiel favorise la mise en relief de la dimension humaine de ce cas d'actualité, accompagnée aussi de l'expression d'un sentiment d'empathie et de solidarité de la part de l'énonciateur, comme en témoigne notamment l'emploi de l'adjectif *pauvre*. Néanmoins, c'est sa mise en discours contextuelle qui contribue à la construction d'une prise de position polarisée, comme on peut le voir dans l'exemple suivant :

- (10) La fin de l'affaire Vincent Lambert ? Ce Pauvre homme qui est resté tant d'années maintenu en vie contre sa volonté. Le laisser mourir de faim

5. Entre parenthèses, des extraits de tweets sont indiqués à titre d'exemple. Les sigles suivants TWO, TWR, TWTH renvoient à la distinction des trois pratiques de tweets illustrées plus haut : tweet original, tweet-réponse, tweet appartenant à un thread.

et de soif semble cruel mais l'euthanasie est illégale en France en même temps... Qu'il puisse enfin partir en paix (Tweet original, 03/07/2019).

Quant à la forme *infirmier*, elle est surreprésentée dans les tweets-réponses en particulier pour mettre en avant le statut professionnel de VL comme un argument pour invoquer le respect de sa volonté. Il est ainsi considéré comme compétent et éclairé compte tenu de sa profession :

- (11) @HeleneLaporteRN Vincent Lambert n'est pas condamné a mort, il est mort depuis 12 ans, infirmier de métier, il était contre l'acharnement thérapeutique ; je ne suis pas d'accord avec vous [...] (Tweet-réponse, 19/05/2019).

La forme *patient* est enfin suremployée notamment dans les tweets-réponses et, dans une moindre mesure, dans les threads, avec toutefois des mises en discours différentes. Les tweets-réponses y recourent surtout pour aborder le côté et les enjeux médicaux liés à cette situation spécifique. Par ailleurs, cette approche de la question implique souvent un ton polémique qui met en discussion le déroulement de la prise en charge du patient et les diverses décisions de l'institution médicale, en revendiquant en outre les *droits* du patient, le respect de sa *volonté* et de ses *intérêts*. On assiste ainsi à des contestations portant sur la gestion médicale de ce cas ou encore sur la catégorisation *du patient*, tout en élargissant également le commentaire à des questions plus générales, comme c'est le cas pour l'exemple suivant.

- (12) [...] La loi Leonetti, y compris dans sa révision de 2016, s'applique aux patients MALADES et/ou EN FIN DE VIE. Vincent Lambert n'est ni malade (il n'a aucun traitement thérapeutique), ni en fin de vie (Tweet-réponse, 20/05/2019).

Par contre, dans les threads, la forme *patient* peut être utilisée dans une acception générale (*droits du patient*) ou bien faire l'objet de syntagmes définitionnels ayant souvent soit une fonction présentative à travers l'apposition (TWTH – 1 / 2 – *Le corps de Vincent Lambert, patient en état végétatif depuis presque onze ans, [...], 12/07/2019*), soit une fonction de reprise anaphorique (TWTH – [...] *Acte 2 : appliquer la loi Leonetti même si elle ne concerne pas le patient en question [...], 11/07/2019*). Dans les deux cas, ces usages s'apparentent à des procédés relevés dans le discours journalistique et tout particulièrement dans les articles d'information.

En ce qui concerne en revanche les formes de référencement de l'état de santé de VL, on remarque que la qualification en termes de *handicapé* est la plus suremployée dans les tweets-réponses (figure 36) et qu'elle est utilisée aussi en occurrence avec la forme *pauci-relationnel*. Cela signale la présence de l'un des arguments principaux soutenus par le positionnement qui s'oppose à l'arrêt des traitements, c'est-à-dire le refus de considérer VL comme une personne en fin de vie, soutenant au contraire qu'il s'agit d'une situation de handicap (cf. aussi 4.2.1).

Les formes *conscience* et *handicapés* sont également spécifiques des tweets-réponses, alors qu'elles sont sous-employées aussi bien dans les tweets originaux que dans les threads. Le substantif *conscience* est utilisé notamment pour discuter du degré de conscience de VL (TWR – *Un état végétatif c'est bien plus qu'un handicap. Une vie sans conscience c'est quasiment un état de mort cérébrale. [...] (09/07/2019) ; TWR – Vincent Lambert n'est pas malade, il est paralysé et dans un état de conscience très faible. Mais vivant. [...], 05/05/2019*). Il est également employé pour introduire un appel à la réflexion sur la question de la fin de vie, sur les enjeux notamment éthiques qu'elle entraîne ainsi que sur les décisions individuelles à envi-

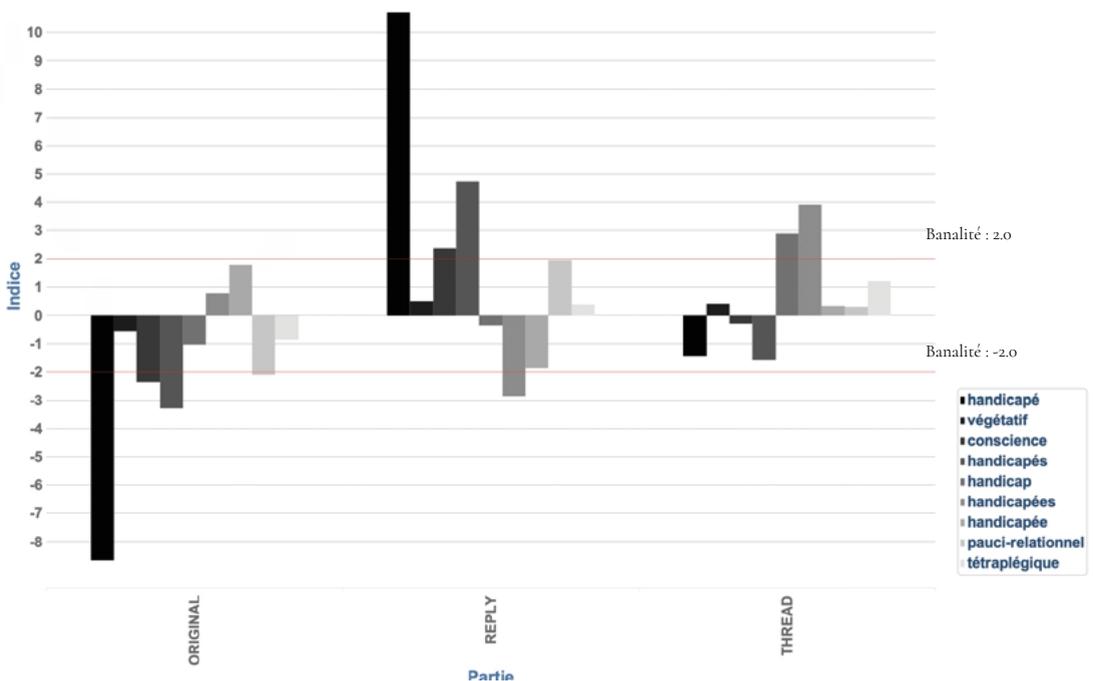


Figure 36. Spécificités sur les formes de référencement principales liées à l'état de santé.

sager (TWR – *Et le cas de #VincentLambert doit faire prendre conscience aux gens de l'importance d'exprimer clairement leur volonté dans ce genre de cas [...], 21/05/2019*).

En revanche, la forme plurielle *handicapés* est utilisée au sein de procédés qui introduisent une perspective plus générale sur la question, à côté des syntagmes dénominatifs qui incluent ce terme tels que *centre des handicapés*, *droits des handicapés* ou *comitat des handicapés*, ceux-ci étant toutefois minoritaires. Les deux formes (*conscience* et *handicapés*) tendent en outre à être mobilisées au sein de technodiscours se rapprochant de l'un des deux positionnements polarisés. Si la forme *conscience* est employée dans les deux cas en termes respectivement de présence (minimale) ou d'absence, le substantif pluriel est au contraire employé davantage dans des tweets s'alignant sur un positionnement pro-maintien. De toute façon, les trois formes (*handicapé*, *conscience* et *handicapés*) font l'objet de mises en discours aux contours polémiques, impliquant une perspective polarisée sur la question.

Quant aux tweets originaux et aux threads, ils présentent des similitudes dans ce cas aussi, comme en témoigne le recours privilégié à la forme *handicapées* (cf. figure 36). On signale des occurrences au sein de la dénomination *Comité des droits des personnes handicapées*, ce qui révèle, encore une fois, que les éléments les plus factuels se concentrent dans les deux pratiques les plus interdiscursives. La forme *handicapées* est en outre utilisée comme qualificatif du substantif *personnes* pour mettre en relation le cas de VL avec la question plus générale du handicap et des répercussions que cet événement pourrait entraîner chez ces personnes. Les formes *handicap* et *(personne) handicapée*, suremployées respectivement dans les threads et dans les tweets originaux, vont également dans la même direction.

On souligne enfin que les expressions *état végétatif* et *état pauci-relationnel* sont récurrentes en particulier dans les threads et encore plus dans les tweets-réponses. Si l'expression *état végétatif* est de fait suremployée dans les threads et dans les tweets-réponses, et sous-employée dans les tweets originaux, elle n'atteint pas dans les trois cas le seuil de banalité fixé conventionnellement à ± 2 sous TXM (Heiden *et al.* 2010). En revanche, l'expression *état pauci-relationnel* s'avère spécifique à la fois des tweets-réponses et des tweets originaux, où elle est respectivement sur- et sous-employée.

5.2.2. Construction des objets de discours : procédés discursifs

5.2.2.1. Procédés de référenciation et de mise en discours de l'état de santé

Dans les tweets originaux, l'orientation de la construction de l'objet de discours – l'état de santé de VL – n'est déterminée pas seulement par différents choix de référenciation (*état végétatif, état pauci-relationnel, handicapé*), mais elle résulte aussi des procédés de mise en discours utilisés. Plus précisément, son état de santé est surtout commenté en relation avec trois aspects : l'opposition des parents à l'arrêt des traitements considéré comme une forme de non-respect de sa volonté, l'assimilation de l'état végétatif de VL à une situation de fin de vie et d'acharnement thérapeutique prolongeant ses souffrances, ainsi que l'indignation face à la gestion contradictoire de ce cas de la part des institutions au fil des années.

En premier lieu, l'état de santé est commenté en relation avec la position des parents, qui est contestée en l'assimilant à un acte d'*acharnement* (ex. 13), ce qui l'encadre de manière négative. La durée de l'affaire est un deuxième argument avancé pour manifester un dissensus, en mettant en avant la non-évolution de la situation médicale du patient (ex. 14). Dans ces discours, la position des parents est présentée comme de l'obstination, et le maintien en vie prolongé d'une personne dans un état végétatif est abordé en relation avec les souffrances qu'il impliquerait, comme le montrent les exemples suivants.

(13) Mais l'affaire Vincent Lambert je trouve que c'est de l'acharnement. Il est en état végétatif ++ alors pourquoi s'obstiner. Et rien à voir avec le validisme (Tweet original, 22/05/2019).

(14) Mais l'affaire de Vincent Lambert, désolé mais les parents attendent quoi pour respecter le souhait de leur fils?? Le pauvre ça fait 10ans il est dans un état végétatif. [...] (Tweet original, 20/05/2019).

En deuxième lieu, des réactions de nature émotionnelle accompagnent la critique de l'assimilation de l'arrêt des traitements à un acte criminel et du manque de respect de la volonté de VL (ex. TWO – *Je trouve ça triste que des personnes assimilent l'arrêt des soins de Vincent Lambert à une peine de mort* (20/05/2019) ; TWO – [...] *je trouve ça honteux de lui faire subir ça*, 20/05/2019). En particulier, il ressort de ces tweets un sentiment de pitié pour les souff-

frances que VL est supposé éprouver (*leur fils le pauvre*) ou bien de l'indignation vis-à-vis de l'*acharnement* dont il serait victime (TWO – [...] *cet acharnement est cruel* [...], 11/05/2019). Dans d'autres tweets, l'arrêt des traitements est en revanche présenté comme ce qui met fin à ses souffrances (TWO – [...] *Heureux pour lui que ce soit fini, il a suffisamment souffert*, 08/07/2019), se chargeant ainsi d'une connotation positive. C'est donc un positionnement pro-arrêt qui apparaît dominant dans cette mise en discours de l'état de santé de VL.

Enfin, bien qu'elle soit moins fréquente que dans les tweets-réponses, la mise en discours de l'état de santé de VL en termes d'*état végétatif* est exploitée dans les tweets originaux pour mettre en évidence sa chronicité et l'impossibilité d'amélioration, représentant un argument majeur pour étayer un positionnement pro-arrêt. En d'autres termes, l'état végétatif est assimilé à une situation de non-vie en tant qu'existence purement biologique et artificielle, qui serait contraire à sa dignité. Dans ce contexte, contrairement aux exemples repérés dans le discours journalistique, le possible degré de conscience du patient en état végétatif ne représente pas un élément en faveur du maintien des traitements. Il devient plutôt un contre-argument, puisque le patient pourrait avoir du mal à supporter cette situation et les éventuelles souffrances associées (TWO – [...] *il est sûrement conscient et être emprisonné dans son corps ça doit être la pire chose possible...* [...], 11/07/2019).

La mise en discours de l'état de santé de VL et notamment la prise de position conséquente de l'énonciateur à cet égard s'accompagnent souvent d'une prise en charge énonciative à la première personne du singulier. Cette tendance est liée à deux fonctions principales : exprimer sa propre réaction émotionnelle ou encore s'identifier à cette expérience. On repère en effet des exemples dans lesquels l'énonciateur imagine se mettre dans la peau de VL et exprime sa propre volonté, qui est très souvent de ne pas être maintenu en vie dans ces conditions. Les verbes *tuer*, *supplier* et *débrancher* sont de fait les plus utilisés dans ce contexte. On assiste ainsi à une personnalisation hypothétique de l'expérience de VL qui permet à l'énonciateur de remettre au web ses propres volontés dans l'éventualité d'une situation future similaire, ainsi qu'en témoigne l'exemple suivant.

- (15) Si un jour je me retrouve dans un état de dépendance comme Vincent Lambert, svp je veux qu'on me débranche. Je n'ai jamais aimé la dépendance (Tweet original, 22/05/2019).

À l'inverse, et même s'ils s'avèrent moins fréquents, des exemples de personnalisation de la situation de VL sont également présents dans des tweets s'apparentant cette fois à un positionnement pro-maintien. En l'occurrence, la mise en avant d'autres situations similaires à celle de VL n'implique pas l'énonciateur lui-même, mais plutôt d'autres expériences, connues de manière directe ou via les médias. La référence à des personnes qui se sont réveillées après de longs comas représente un argument visant l'opposition à l'arrêt des traitements de VL. Ces exemples sont toutefois minoritaires.

Dans les tweets-réponses, l'état de santé de VL est beaucoup plus souvent désigné en termes d'*état végétatif*, parfois en cooccurrence avec l'adjectif *chronique* ou *irréversible*, et cela en particulier dans les tweets se rapprochant d'un positionnement pro-arrêt. L'immutabilité de cet état constitue en effet un argument majeur de cette prise de position, en association aussi avec le manque de conscience qui est supposé le caractériser. Le même concept est exprimé également à travers la négation de l'adjectif *pauci-relationnel* (cf. ex. 16). Ces mises en discours de l'état de santé soulignent en outre une existence purement physique (ex. 17) qui lui serait propre.

(16) [plusieurs mentions de comptes] Les experts mandatés par la justice viennent de conclure que l'"état végétatif chronique irréversible" de Vincent Lambert ne lui laisse plus "d'accès possible à la conscience". Cet état n'est pas l'état de conscience minimale (ou pauci-relationnel) auquel vous faites allusion (Tweet-réponse, 05/07/2019).

(17) [...] Vincent Lambert est dans un état végétatif donc il ne ressent rien et sans les appareils et les soins, il serait mort. Il est juste un corps avec un cœur qui bat, sans aucune conscience (Tweet-réponse, 21/05/2019).

Dans ce contexte, l'arrêt des traitements est donc envisagé comme la seule solution possible pour se délivrer des souffrances d'une vie exclusivement physique. Le discours sur l'état de santé se rapproche ici ouvertement d'une mise en perspective spécifique de la question, visant en même temps à réfuter la construction discursive-argumentative de la position adverse.

Comme on l'a déjà souligné dans la section précédente, l'expression *état végétatif* n'est pourtant pas exclusive d'un positionnement pro-arrêt. Elle est également utilisée dans une mise en discours en faveur du maintien des traitements qui, en l'occurrence, fait l'objet d'une négation polémique (Ducrot 1984). La négation de l'*état végétatif* a en effet pour objectif de refuser cette forme de référenciation qui ferait allusion à une absence de conscience. Par

ailleurs, comme c'est souvent le cas dans les réponses, la structure négative permet de s'opposer au tweet auquel on réplique, donnant lieu à une prise de position dialogique qui marque le désalignement du discours second. Parfois, une catégorisation différente de l'état de santé est également fournie lors de cette mise en discours, avec fonction de contre-argument :

- (18) [...] Mr # VincentLambert n'est pas en fin de Vie, n'est pas en état végétatif, il est handicapé !!!!!!! (Tweet-réponse, 21/05/2019).

La surreprésentation de la forme *légume*, employée avec valeur métaphorique dans les tweets-réponses abordant l'état de santé, est particulièrement significative de la conception de la vie qui y est associée. Le transfert métaphorique activé par l'usage de cette forme pointe la passivité de l'état de santé en question. Malgré son caractère indélicat, cette forme est largement utilisée dans les échanges au sein de la plateforme, témoignant ainsi du registre discursif même violent qui peut caractériser les discours des RSN.

L'état de passivité évoqué par *légume* est aussi bien attribué par le positionnement pro-arrêt (TWR – [...] *Vincent Lambert avait une volonté, celle de ne pas être maintenu comme un légume, il faut la respecter*, 21/05/2019), que réfuté par le positionnement pro-maintien (TWR – [...] *Il respire par lui-même ! Ce n'est pas un légume ! Il est juste trop handicapé pour se nourrir seul ! [...]*, 20/05/2019). Le fait de se trouver dans un état végétatif semble ainsi être considéré, par le positionnement pro-arrêt, comme une situation de vie qui perd sa spécificité. C'est cette condition qui est tenue pour inhumaine et indigne, contrairement à l'interruption des traitements, présentée comme la solution aux souffrances du patient. Ainsi, le maintien en vie de VL est envisagé comme un acharnement thérapeutique, contrastant non seulement avec sa volonté, mais aussi avec le déroulement naturel de la vie, qui est à l'antithèse du caractère artificiel de l'alimentation et de l'hydratation administrées. Au contraire, cette représentation discursive de l'état de santé comme passif et artificiel est réfutée par le positionnement pro-maintien, qui souligne les actes vitaux propres à cet état, comme le fait de respirer de manière autonome, et en propose une contre-catégorisation en termes de handicap. En l'occurrence, l'alimentation et l'hydratation artificielles sont présentées comme nécessaires pour répondre aux besoins essentiels de la personne, celle-ci étant dans l'incapacité de s'alimenter et de s'hydrater en autonomie en raison de son état de handicap.

En comparant les tendances relevées dans les tweets originaux et dans les tweets-réponses, on constate finalement l'emploi constant des formes de

référenciation selon les deux positions polarisées. Ce qui les différencie en revanche, c'est leur mise en discours et en conséquence la structuration du positionnement ainsi que sa polarisation.

Enfin, dans les threads, la mise en discours de l'état de santé de VL tend à être associée avec divers aspects de la polémique traités en revanche séparément dans les autres pratiques. En effet, la qualification de l'état de VL s'articule aussi bien avec la représentation de la mort qu'avec la catégorisation de l'alimentation et de l'hydratation artificielles. Dans cette pratique technodiscursive, l'état de VL est relié à des situations de handicap, s'apparentant donc à une mise en perspective de la question propre à un positionnement pro-maintien. Des représentations métaphoriques de son état et de sa situation sont en outre plutôt surreprésentées dans les tweets appartenant à des threads, au-delà des tweets-réponses. Les substantifs *prisonnier* ou encore *otage* sont en l'occurrence utilisés pour souligner aussi bien son état que son impossibilité d'action, avec, dans ce cas aussi, un alignement sur les constructions discursives principales du positionnement pro-maintien. Il arrive que VL soit présenté comme étant *prisonnier* non pas de son corps – comme c'était le cas dans la mise en perspective proposée par le positionnement pro-arrêt – mais de l'institution juridique (la justice) et médicale (incarnée à la fois par l'équipe médicale et le CHU) ainsi que comme *otage* des décisions d'autrui (TWTW – [...] *retenir prisonnier #VincentLambert dans ce CHU, comme un otage qu'on ne veut pas libérer...* [...], 16/05/2019).

5.2.2.2. Procédés de mise en discours de la mort

La mort représente l'une des thématiques principales relevées par la CHD réalisée sur le corpus TWS (cf. *supra* figure 25). Comme l'analyse du corpus PN l'a déjà montré, la mise en discours de la mort matérialise la polarisation des positionnements discursifs : à une représentation active comme acte provoqué s'oppose une représentation passive comme événement naturel. En premier lieu, l'étude de la répartition de quatre formes parmi les plus communes pour aborder ce sujet – *mort*, *décès*, *mourir* et *décéder* (figure 37) – dans le corpus TWS révèle leur suremploi dans les tweets originaux. En particulier, les deux formes propres au vocabulaire médical, administratif et juridique (*décès* et *décéder*) sont les plus surreprésentées. Cela pourrait dépendre de la présence de tweets de nature informative-factuelle ou bien de tweets qui commentent la nouvelle de la mort en reprenant les termes mobilisés par le discours d'information. Deux autres éléments confirment d'ailleurs cette interprétation. D'une part, la plupart des occurrences de ces

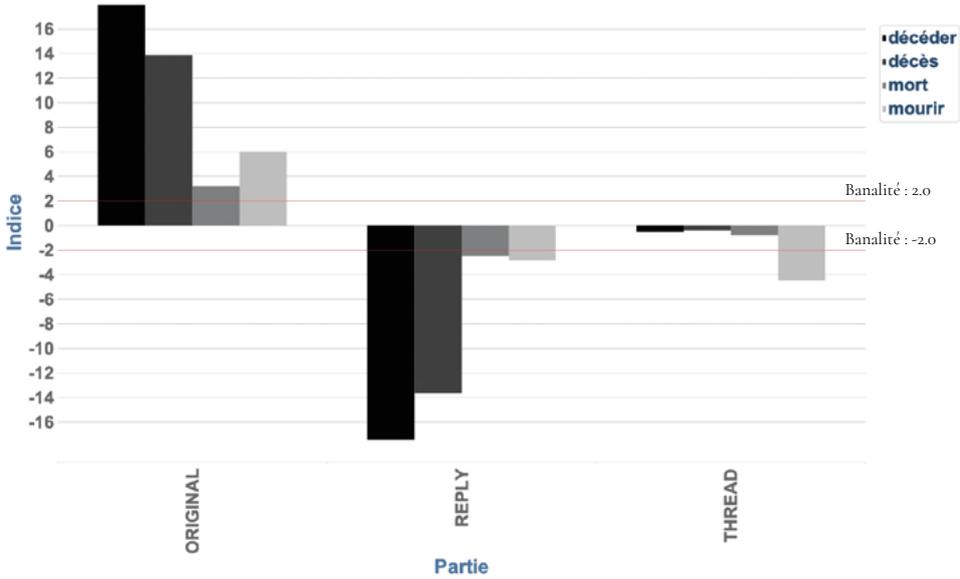


Figure 37. Spécificités sur les lemmes dans le corpus TWS.

formes (80 %) apparaissent au sein de tweets publiés le jour même de la mort de VL. De l'autre, les mots qui cooccurrent le plus avec ces termes concernent l'encadrement déictique (*matin, jeudi, ce, 8h24*) et discursif (*annonce*) de l'événement caractérisant le discours journalistique, comme en témoigne aussi la présence de références explicites à des médias (*AFP*).

On remarque en outre que les formes *décéder* et *décès* sont moins sous-employées dans les threads que dans les tweets-réponses. Ce résultat corrobore donc la présence d'une dimension informative plus marquée dans ces fils technodiscursifs, tout comme dans les tweets originaux. Une distribution plus équilibrée caractérise en revanche les lemmes *mort* et *mourir*, confirmant toutefois les tendances opposées déjà relevées pour les termes plus spécialisés.

L'analyse des lemmes qui cooccurrent le plus avec la forme *mort* dans les trois pratiques met ensuite en évidence la prédominance d'une représentation active de l'événement comme étant le résultat d'un acte non seulement provoqué mais aussi imposé. Cette mise en discours est présente surtout dans les tweets originaux et dans les tweets-réponses. Les deux formes les plus cooccurrentes sont les substantifs *peine* et *mise* qui établissent respectivement un parallélisme entre l'arrêt des traitements de VL et l'exécution d'une *peine de mort*, ainsi qu'une assimilation de l'acte de le débrancher à une *mise à mort*. Connotées de manière fortement négative, les deux expressions non seulement contribuent à la construction d'un positionnement

pro-maintien marqué d'éléments axiologiques et émotionnels, mais elles attirent également l'attention sur les responsables de cette décision en raison de la nature agentive de l'acte ainsi nommé. Les verbes aussi participent à cette mise en discours orientée et polarisée. Parmi les plus utilisés, *condamner*, *programmer*, *décider*, *conduire*, *provoquer*, *donner* et *mettre* partagent tous le trait sémantique de la dimension active de l'acte exprimé.

La connotation axiologique négative de cette mise en discours de la mort s'intensifie en outre lors de la représentation de VL comme un *innocent* (notamment dans les tweets originaux et les threads) condamné à la peine capitale, ou encore comme un *martyr* (threads et réponses) qui subit une mort imposée, sans avoir commis aucun crime. C'est son état de *handicapé* qui est considéré comme la cause de sa *mise à mort*, ce qui est censé provoquer une réaction d'indignation du lecteur. La comparaison du cas de VL avec différentes catégories de criminels est également récurrente dans les tweets incluant ces formes. Cette mise en perspective contribue à son tour à souligner l'injustice dont il est considéré être *victime*, ce terme étant suremployé dans les tweets-réponses, comme on peut le voir dans le graphique de la figure 38.

En bref, la représentation active de la mort comme acte provoqué, qui se réalise à travers des constructions discursives impliquant également des choix lexicaux fortement axiologiques, donne lieu non seulement à une mise en discours orientée, mais elle constitue aussi l'un des arguments principaux du positionnement pro-maintien. Ces constructions de l'objet de

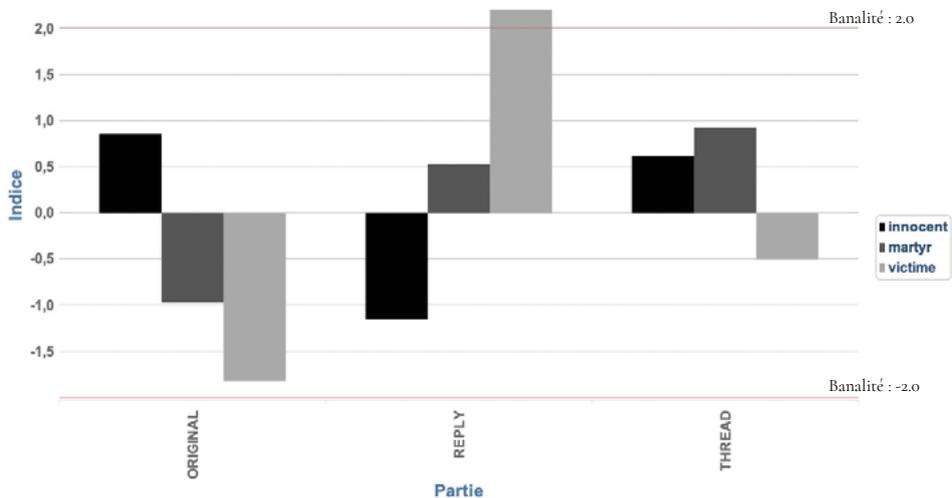


Figure 38. Calcul des spécificités sur les lemmes *innocent*, *martyr* et *victime* dans le corpus TWS.

discours permettent ainsi à l'énonciateur de manifester son dissensus de manière ouverte, en mettant en avant une validation subjective par le recours à des ressources linguistiques axiologiques.

Au-delà des différentes contextualisations repérées de la forme *mort*, la dimension active de cet événement est également véhiculée par le recours aux verbes *tuer*, *assassiner* et *euthanasier* utilisés en relation avec l'acte d'arrêter les traitements. Il en résulte une mise en discours orientée et connotée axiologiquement, qui marque une prise de position ouverte, pouvant inclure aussi des éléments de violence verbale. Ces verbes connotés négativement sont suremployés notamment dans les tweets-réponses (figure 39), qui se caractérisent, dans ce cas encore, par un discours plus polémique.

L'étude des concordances de ces verbes montre que l'assimilation de la mort de VL à un crime – notamment par le suremploi de *tuer* et de *assassiner* dans les tweets-réponses – implique une axiologisation négative intrinsèque aux verbes mêmes, de nature aussi bien sociale que juridique. En particulier, le sémantisme de ces verbes – dans leur acception la plus fréquente et générale recensée dans le *TLFi*⁶ – attire l'attention sur deux aspects parfois explicités également dans leurs contextes d'utilisation ou, plus en général, dans les mises en discours de la mort en tant qu'acte provoqué. D'une part, le verbe *tuer*, souligne la nature violente et intentionnelle de cet acte. De l'autre, la préméditation et l'injustice sont convoquées par le verbe *assassiner*. Ces verbes sont exploités aussi pour contester ceux qui sont tenus pour responsables – directs ou indirects – de la décision d'arrêter les traitements, déclenchant un processus de « montée en tension » (Moïse 2012), manifestation de la dimension émotionnelle du discours. Des sentiments de colère, d'injustice ou encore d'indignation face à cet événement traversent ces discours, marquant un ton plus violent dans la polémique telle qu'elle se déclenche en particulier dans les tweets-réponses.

Le verbe *mourir* est le seul non connoté et il est surreprésenté surtout dans les tweets originaux. L'analyse des cooccurrences de *mourir* dégage toutefois l'emploi contextuel de cinq formes : les substantifs *faim* et *soif*, ainsi que les verbes *faire*, *laisser* et *vouloir*. Ces formes participent de manière significative à deux mises en discours divergentes de la question. De fait, bien qu'il n'implique aucune connotation axiologique, le verbe *mourir* est utilisé souvent en relation avec ce qui est considéré être la cause de la mort, c'est-à-dire la privation de l'alimentation et de l'hydratation artificielles. Par ail-

6. TLFi, *Trésor de la Langue Française informatisé* : <http://atilf.atilf.fr> (dernière consultation le 21/06/2023).

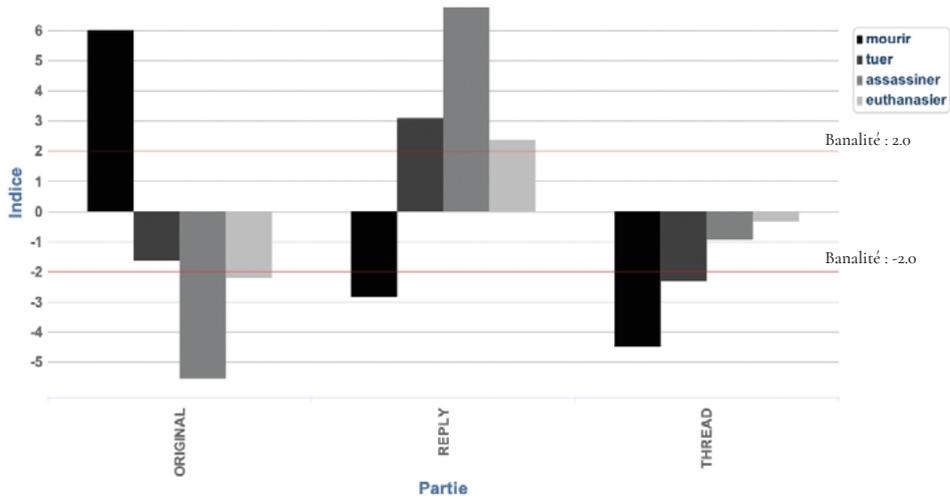


Figure 39. Spécificités sur les lemmes mourir, tuer, assassiner et euthanasier.

leurs, cette contextualisation de *mourir* est fréquemment attestée dans les tweets, en dehors du positionnement dont ils se rapprochent. Cela entraîne une construction polarisée et opposée d'un même objet de discours, tout en recourant à des ressources similaires. Cette construction discursive témoigne en outre de la dimension intrinsèquement dialogique de la mise en discours polémique : le discours adverse est souvent convoqué explicitement ou évoqué implicitement afin de le contester.

En ce qui concerne les verbes qui cooccurrent le plus avec *mourir*, le verbe *laisser* véhicule lui aussi des valeurs sémantiques et axiologiques différentes en fonction des positionnements. Dans les discours s'alignant sur un positionnement pro-arrêt, il fait souvent l'objet de l'exhortation à ne pas empêcher le départ de VL (TWO – *Laissez ce pauvre homme mourir dignement et en paix [...]*, 20/05/2019). Les traits sémantiques de la liberté et de la délivrance sont donc mis en avant dans ce contexte. Par contre, les discours se rattachant à un positionnement pro-maintien emploient le verbe *laisser* surtout de manière polémique pour dénoncer et contester la décision de l'abandonner à une mort de faim et de soif (TWR – [...] *Laisser mourir de faim et de soif un homme en France en 2019 est tout simplement inadmissible [...]*, 11/07/2019). Dans ce contexte, le recours aux adverbes *dignement* et, moins fréquemment, *naturellement* marque également une position pro-arrêt visant à défendre une conception de la vie qui s'oppose à la survie purement biologique du corps.

Le verbe *vouloir* aussi est différemment employé en cooccurrence avec *mourir* selon les deux positionnements. Dans les discours pro-arrêt, il fait référence à la volonté de VL pour demander qu'elle soit respectée. Dans les discours pro-maintien, il est plutôt utilisé pour dénoncer la volonté de provoquer sa mort en arrêtant son alimentation et son hydratation artificielles.

À la représentation de la mort comme conséquence d'un choix délibéré, qui est propre au positionnement pro-maintien, les tweets se rapprochant d'un positionnement pro-arrêt opposent la mise en discours de la mort comme un événement qui survient de façon naturelle. Elle est en effet envisagée comme une délivrance des souffrances entraînées par l'état de santé de VL. En l'occurrence, la connotation positive dont se charge l'événement influence le choix d'utiliser moins de termes usuels tels que *mort* et *mourir*, qui impliquent une charge émotionnelle négative. La mise en évidence de cette dimension positive s'appuie en particulier sur l'utilisation des expressions *laisser partir (tranquille)* et *laisser tranquille* en association éventuelle aussi avec le syntagme adverbial *en paix* (TWO – [...] *laissez le partir enfin en paix*. #VincentLambert, 20/05/2019 ; TWO – [...] *laissons #VincentLambert s'en aller en paix!* [...], 21/05/2019).

En bref, du point de vue intragénérique, on relève que des formes de référenciation connotées et des contextualisations axiologisées caractérisent plus souvent la mise en discours de l'événement de la mort dans les tweets-réponses. Les tweets originaux recourent également à des formes neutres – à l'instar du verbe *mourir* – dont les procédés de mise en discours déterminent la polarisation de la construction discursive. Aucun des quatre verbes analysés (*mourir*, *tuer*, *assassiner*, *euthanasier* – cf. figure 39) n'est en revanche surreprésenté dans les threads, dans lesquels la thématique de la mort est moins traitée. De fait, s'ils se caractérisent par moins de mots-clés liés à ce sujet, cela est dû également à leur nombre plus limité au sein du corpus. On constate toutefois que plusieurs aspects des constructions discursives repérées dans les tweets originaux et dans les tweets-réponses se manifestent dans cette pratique aussi. C'est le cas notamment pour la représentation active de la mort, abordée en particulier en relation avec ce qui est censé en être la cause. Dans cette contextualisation, elle est en effet représentée comme la conséquence de la privation de l'alimentation et de l'hydratation artificielles (TWTH – [...] *il va cesser d'être alimenté et hydraté : cet homme va donc mourir de faim et de soif* [...]. 24/04/2019 ; TWTH – [...] *décider sa mise à mort en le privant d'alimentation et d'hydratation* [...], 12/05/2019). Des choix verbaux très fréquents dans les autres pratiques caractérisent les threads aussi, à l'instar de la cooccurrence *faire mourir* et du

recours au verbe *provoquer* en association avec le substantif *mort*. Quant au verbe *mourir*, même s'il est sous-représenté par rapport au reste du corpus, son emploi en cooccurrence avec *faire* et *laisser* est exploité dans ce cas aussi dans des contextes différents s'alignant sur l'un des deux positionnements.

Pour résumer, on peut remarquer que la construction discursive de la mort présente des caractéristiques assez stables dans la polarisation des positionnements selon les trois pratiques. Un ton plus polémique et parfois violent se manifeste néanmoins dans les tweets originaux et encore plus dans les tweets-réponses en conséquence de leur statut technodiscursif. Au sein de ces deux pratiques, l'étude des mots les plus spécifiques dans les tweets relatifs à ce sujet montre la fréquence de la représentation de la mort non seulement comme un acte provoqué, mais aussi comme un acte criminel ou encore comme l'exécution d'une peine. La distribution des verbes utilisés pour représenter la mort va également dans le sens des autres résultats obtenus jusqu'ici : la dimension interlocutive des réponses favorise la mise en contraste de positions divergentes et, en conséquence, un discours plus polarisé et parfois violent. Une dimension plus factuelle est en revanche plus présente dans les autres pratiques, sans que cela n'empêche une mise en discours orientée de la question et la mise en avant de la subjectivité de l'énonciateur, notamment dans les tweets originaux, par l'emploi du pronom *je*.

5.3. Synthèse comparative et articulation du local au global

L'analyse inductive-quantifiée du corpus TWS a fait ressortir plusieurs éléments de rapprochement et de divergence intragénérique aux niveaux lexi-co-sémantique, technolangagier et énonciatif-pragmatique. Elle a montré que les différents sujets thématiques abordés ont été traités selon des proportions différentes dans les trois types de tweets simples et cela en fonction de leurs statuts technodiscursifs. Par cette approche du niveau global du corpus, plusieurs spécificités énonciatives ont également été identifiées concernant aussi bien l'usage des pronoms personnels que le recours aux différents éléments technolangagiers et discursifs tels que les hashtags, les catégories grammaticales, les signes typographiques et encore les données chiffrées.

L'identification de ces traits distinctifs intragénériques s'est avérée essentielle pour se pencher ensuite sur les niveaux locaux du corpus eu égard aux spécificités technogénériques de chaque pratique. L'étude dé-

ductive-contextualisée de la construction discursive de deux des objets de discours principalement abordés lors du débat public envisagé a permis d'approfondir la caractérisation des trois pratiques technodiscursives, en observant en contexte les fonctionnements des caractéristiques repérées lors de la première phase de l'analyse. En particulier, l'étude contextualisée des spécificités énonciatives-pragmatiques – y compris les traits technolangagiers – a révélé des fonctionnements et des mécanismes que nous supposons être moins susceptibles de varier en fonction des sujets abordés, dans la mesure où ils sont déterminés par le statut technodiscursif même de chaque pratique. Si le statut technodiscursif des différents types de tweets influence également les ressources et les procédés lexico-sémantiques, ces derniers sont néanmoins plus étroitement dépendants du sujet thématique.

L'articulation des résultats inductifs-quantifiés avec les observations déductives-contextualisées rend donc compte de la variation intragénérique à deux niveaux. Les traits repérés au niveau énonciatif-pragmatique sont supposés être plus constants d'un point de vue technodiscursif et transversaux d'un point de vue thématique. Les caractéristiques identifiées au niveau lexico-sémantique sont en revanche liées plutôt à la thématique caractéristique du corpus, mais relèvent aussi de l'influence des spécificités technodiscursives. C'est notamment sur le niveau lexico-sémantique que repose donc l'étude de la variation intergénérique, à savoir la manière dont un même fait d'actualité et ses objets de discours principaux sont diversement mis en discours dans des genres discursifs différents. L'étude donne ainsi accès aux variations dans la construction discursive de l'événement et des positionnements discursifs (Cagninelli 2022, 2024b), déterminées par les spécificités génériques et sous-génériques. En même temps, comme nous l'observerons ci-dessous, la comparaison inter- et intragénérique des constructions discursives principales des positionnements discursifs dominants permettra à son tour de mettre en évidence les différents rôles exercés par les traits spécifiques propres à chaque genre et sous-genre (techno)discursif.

5.3.1. *Polarisations principales, rapprochements et divergences intragénériques*

Suivant une démarche comparable à celle adoptée pour le corpus PN, l'analyse du corpus TWS s'est focalisée sur les formes de référencement, sur leurs contextualisations et sur les procédés discursifs employés pour construire deux des objets de discours principaux de la controverse : les références au

protagoniste de l'affaire, en relation avec son état de santé, et la représentation de la mort. D'une part, la qualification de l'état de santé de VL a constitué un élément central dans l'encadrement à la fois médical et juridique de ce cas spécifique, justifiant en conséquence les différentes prises de position à son égard. De l'autre, la représentation linguistico-discursive de la mort – aussi bien en tant qu'événement particulier que dans sa dimension générale – a été indicative de la conception de la vie à laquelle se rattache l'énonciateur, tendant à s'aligner sur un positionnement plutôt que sur l'autre. La mise en discours de ces deux objets de discours se charge ainsi d'une valeur argumentative au sein du débat public et contribue de manière significative à la construction des positionnements discursifs dominants.

5.3.1.1. Formes de référencement et procédés de mise en discours de l'individu au centre du débat

En ce qui concerne les formes de référencement utilisées pour désigner et se rapporter au protagoniste de l'événement, on constate que seuls les tweets originaux se caractérisent par un suremploi de la forme *homme* et de sa variante, moins fréquente, *jeune homme*, qui contribuent à souligner la dimension humaine de la question ainsi que ses enjeux éthico-sociaux. L'attention est donc attirée, selon les cas, sur les spécificités de cette situation particulière, mais aussi sur la question plus générale de la fin de vie. Ces formes font pourtant l'objet de mises en discours différentes en fonction du positionnement discursif auquel elles se rattachent : le respect de la *volonté* de VL, les *souffrances* que son état peut entraîner et l'appel à le *laisser partir en paix* s'opposent à la dénonciation de le *laisser mourir de faim et de soif*.

La question du respect de sa volonté est abordée également par le recours à la forme *infirmier*, qui est en revanche la plus spécifique des tweets-réponses. En l'occurrence, le statut professionnel de VL est convoqué en tant qu'argument pour témoigner de ses compétences par rapport à cette question éthico-médicale, visant à démontrer le bien-fondé conscient de sa volonté. La compétence impliquée par la profession infirmière est exploitée par ailleurs pour revendiquer l'autorité du dire des énonciateurs qui l'exercent.

Une dernière forme de référencement utilisée pour désigner VL correspond au terme *patient* qui est suremployé notamment dans les tweets-réponses et, dans une moindre mesure, dans les threads. Dans les tweets-réponses, on recourt à ce terme plus technique pour traiter des aspects médicaux de la question et surtout pour contester la gestion médicale de ce cas. Les threads se rapprochent en revanche des usages typiques du discours journalistique,

dès lors qu'ils utilisent *patient* dans des procédés d'apposition et de reprise anaphorique, au-delà de l'emploi du terme dans sa portée générale.

5.3.1.2. Formes de référencement et procédés de mise en discours de l'état de santé

Quant aux formes de référencement relatives à l'état de santé, l'expression *état végétatif* est assez récurrente dans les trois pratiques technodiscursives, tout en n'étant spécifique d'aucune en particulier. Les trois types de tweets comportent pourtant des mises en discours différentes de l'état de santé en termes d'*état végétatif*. Dans les tweets originaux, cette question est abordée notamment en relation avec l'opposition à l'arrêt des traitements de la part des parents, interprétée également comme un non-respect de la volonté de leur fils. L'encadrement de la situation en tant qu'*acharnement* prend appui sur la mise en avant du caractère irréversible de cet état, supposé entraîner aussi des souffrances. Cette articulation thématique participe en outre à un étayage de l'émotion (Micheli 2014), susceptible de provoquer aussi bien de la pitié que de l'indignation lors de la réception-interprétation de ces discours.

Le manque de conscience et le caractère *chronique* et *irréversible* de l'état végétatif sont soulignés plus fréquemment dans les tweets-réponses, qui se caractérisent plus que les autres pratiques par un ton polémique, vif et parfois agressif. Une articulation thématique plus large est propre en revanche aux threads, dans lesquels la mise en discours de l'état de santé va souvent de pair avec la question de l'alimentation et de l'hydratation artificielles ainsi qu'avec le thème de la mort.

Dans tous les cas, les références à *état végétatif* impliquent principalement une passivité découlant d'une existence artificielle – du fait qu'elle dépend de traitements médicaux – et purement biologique, à la suite d'un degré de conscience jugé comme minimal, voire absent. Il en ressort la représentation discursive d'une existence privée des traits distinctifs de la vie pleinement vécue. Il s'ensuit que l'expression *état végétatif* est largement employée dans les discours s'alignant sur un positionnement pro-arrêt.

D'autres formes de référencement à l'état de santé sont en revanche sur- et sous-employées différemment dans les trois pratiques. Un cas particulièrement significatif est celui des formes *handicapé* et *pauci-relationnel*, qui sont mobilisées de manière privilégiée dans les tweets-réponses, au sein de discours relevant principalement d'un positionnement pro-maintien. Comme on l'a déjà précisé plus haut, la qualification de l'état du patient comme une forme de handicap permet en effet de s'opposer à son as-

similation à une situation de fin de vie. L'état de handicap présuppose ainsi un cadrage médical différent, qui n'implique pas les conditions nécessaires pour interrompre les traitements et mettre en œuvre une sédation profonde et continue. Pour ces raisons, il s'agit d'une mise en discours de l'état de santé qui se rapproche d'un positionnement pro-maintien.

Au contraire, le degré de *conscience* constitue un élément à valeur argumentative utilisé de manière opposée dans la polémique pour étayer l'une des deux positions par rapport à cet objet de discours. Cet aspect est abordé en particulier dans les tweets-réponses, qui révèlent encore une fois une nature polémique plus marquée.

5.3.1.3. Représentations discursives de la mort

Quant à la construction de l'objet de discours de la mort, l'analyse du corpus TWS a montré que la représentation de la dimension active de cet événement y est prédominante. Cette représentation est axiologisée négativement par une mise en discours de la mort comme acte provoqué et criminel, à travers un suremploi de diverses expressions qui mettent l'accent sur ces traits sémantiques. Il s'agit ainsi d'une construction discursive qui se rattache principalement à un positionnement pro-maintien contestant la décision d'arrêter l'alimentation et l'hydratation artificielles à une personne considérée comme n'étant pas en fin de vie.

Les discours se rapprochant d'un positionnement pro-arrêt proposent en revanche une représentation positive de la mort en tant qu'acte de soulagement, qui permet de mettre fin à un état de souffrance et de trouver ainsi la paix et la tranquillité. La cooccurrence des verbes *laisser* et *partir* est particulièrement significative de cette mise en discours. D'une part, elle atténue la portée dramatique de l'événement par la métaphore du départ, qui constitue également un euphémisme. De l'autre, elle fait écho et s'oppose à l'acharnement dénoncé par le discours pro-arrêt.

D'un point de vue intragénérique, la confrontation polarisée entre les deux prises de position face à cet objet de discours – et, en conséquence, entre les deux positionnements – transparait clairement dans les tweets originaux et encore plus dans les tweets-réponses, caractérisés par une « dimension argumentative » (Amossy 2016) très marquée. Cela concerne non seulement le niveau énonciatif-pragmatique, ainsi que l'analyse inductive-quantifiée l'a mis en évidence (cf. 5.1.2 et 5.1.3), mais aussi le niveau lexi-co-sémantique. En effet, les tweets-réponses abordent les aspects les plus polémiques de l'affaire et tendent ainsi à se rattacher ouvertement à un po-

sitionnement. Ce n'est pas toujours le cas en revanche des tweets originaux et des threads, qui comportent aussi des discours non polémiques visant l'expression des condoléances (tweets originaux) et des aspects plus factuels (tweets originaux et threads) (cf. 5.1.1).

De plus, les tweets-réponses présentent un degré de violence verbale⁷ plutôt marqué, comme en témoigne la surreprésentation d'expressions fortes telles que les verbes *tuer* et *assassiner* ainsi que les substantifs correspondants. L'analyse des mots les plus spécifiques a d'ailleurs montré la présence dominante d'une mise en discours de la mort comme acte criminel, qui peut également se charger d'une valeur symbolique et politique. Dans cette pratique technodiscursive, la présence distinctive des mentions est en outre exploitée pour avancer des critiques ciblées.

Des similitudes dans la structuration des thématiques et des objets de discours ont enfin été repérées entre les tweets originaux et les threads. Les deux pratiques ont en effet abordé aussi bien les aspects les plus factuels de l'affaire que les questions polémiques. Cependant, dans les threads, la variété des objets de discours et des procédés discursifs est plus restreinte que celle repérée dans les autres pratiques et en particulier dans les tweets originaux. D'ailleurs, cela pourrait résulter de la présence plus limitée de ce type de tweets dans le corpus.

5.3.2. *Variations et constantes sous-génériques d'ordre technodiscursif*

Grace à la variation des démarches et des niveaux d'analyse du corpus TWS en fonction de son hétérogénéité technodiscursive, notre étude a relevé des constantes notamment lexico-sémantiques et des divergences significatives de nature énonciative-pragmatique. Il faut toutefois souligner que les tweets originaux ont représenté un point de repère essentiel pour l'analyse de l'hétérogénéité intragénérique, en vertu du fait qu'ils sont plus autonomes d'un point de vue sémantique que les autres types de tweets du corpus. Constituant respectivement des discours seconds et des éléments interreliés au sein d'un même fil discursif, les tweets-réponses et les threads ont entraîné des défis méthodologiques supplémentaires, dans la mesure où un tweet-réponse tout comme un tweet appartenant à un thread, chacun pris individuellement, est privé de son contexte. Par conséquent, si dans certains cas

7. Pour un cadre général sur les phénomènes de violence verbale dans les réseaux sociaux numériques, on renvoie à Monnier *et al.* (2021).

le technodiscours garde son autonomie sémantique, dans d'autres la prise en compte des discours auxquels il est relié est indispensable pour bien interpréter son sens. D'où le rôle des tweets originaux comme paramètre comparatif par rapport auquel identifier les similarités et les divergences notamment dans les procédés de mise en discours des objets de discours.

La comparaison intragénérique a montré que les choix de référenciation et les constructions discursives des deux objets de discours analysés sont assez stables selon la polarisation des positionnements à travers les trois pratiques technodiscursives. De fait, cela semble rendre compte de deux appréhensions dichotomiques de la question. Un même objet de discours est ainsi construit différemment et de manière polarisée en fonction du positionnement discursif auquel se rattache le technodiscours. Les polarisations principales dans les choix de référenciation et dans les procédés de construction des objets de discours restent en effet plutôt constantes à travers les différentes pratiques technodiscursives, au-delà de quelques variations principalement lexicales. Cela signale l'usage transversal de certaines ressources qui peuvent néanmoins faire l'objet de mises en discours opposées. Des formes sur- et sous-représentées dans les trois pratiques montrent enfin des traits distinctifs d'ordre technodiscursif, liés au statut technodiscursif et aux visées principales de chacune, qui impliquent en conséquence différentes caractéristiques au niveau du registre discursif et du ton.

Il est donc possible d'identifier des ressources linguistiques et des procédés (techno)discursifs quantitativement dominants au niveau non seulement sous-technogénérique – procédant des différentes pratiques technodiscursives – mais aussi au niveau technogénérique et pluritechnosémiotique, c'est-à-dire des ressources et des procédés distinctifs propres au technogénre des tweets simples. Il s'agit en conséquence d'éléments d'ordre interdiscursif, pris dans un interdiscours (cf. 1.3.2) qui se façonne à l'intérieur de l'environnement technodiscursif de la plateforme. Cependant, cet interdiscours technodiscursif peut également s'alimenter de et contribuer à l'interdiscours du débat public sur ce cas d'actualité et, plus en général, sur la fin de vie, de manière transversale aux différents genres de discours.

Si des formes de référenciation et des procédés discursifs plutôt stables ont été repérés selon la polarisation des positionnements, tout en impliquant des différences entre les trois pratiques, des divergences encore plus remarquables ont été identifiées au niveau énonciatif-pragmatique. D'ailleurs, ce dernier dépend davantage des spécificités et des statuts technodiscursifs respectifs. L'articulation des résultats de l'analyse déductive-contextualisée des objets de discours selon les trois pratiques technodiscursives

(5.2) avec les spécificités lexico-sémantiques, technolangagières et énonciatives-pragmatiques mises en évidence grâce à l'analyse inductive-quantifiée (5.1) montre que des variations significatives dans la construction des positionnements discursifs concernent le niveau énonciatif-pragmatique. En effet, les tweets originaux se caractérisent par la présence plus fréquente d'une prise en charge énonciative à la première personne du singulier, par laquelle l'énonciateur adresse son discours – revendiqué donc comme subjectif – aux internautes naviguant sur la plateforme. Cela s'accompagne en outre du recours aux hashtags qui remplissent souvent une fonction d'indexation interdiscursive. En l'occurrence, c'est la mise en avant d'une prise de parole subjective qui est distinctive de cette pratique, à travers une énonciation à la première personne du singulier qui contribue à mettre en évidence la subjectivité de l'instance énonciative.

Comme nous l'avons souligné lors de l'analyse inductive-quantifiée, les tweets-réponses se caractérisent en revanche par une orientation dialogique interlocutive explicite, véhiculée notamment par un suremploi du pronom *vous* et par l'usage des mentions, du fait de la nature même de cette pratique. C'est justement leur dimension interlocutive intrinsèque qui influence et différencie la construction énonciative-pragmatique de ces technodiscours ainsi que des positionnements discursifs qui y sont dominants. Le large recours à la négation à valeur surtout polémique (Ducrot 1984), au pronom de la deuxième personne du pluriel ainsi qu'à la concession constituent les traits spécifiques des tweets-réponses, qui sont employés pour signaler l'alignement ou, plus fréquemment, le désalignement du sujet énonciateur sur le technodiscours – et sur le positionnement de l'interlocuteur – auquel il réagit. Il en résulte la construction dominante d'un positionnement discursif dissensuel.

Une dimension argumentative plus marquée distingue également les tweets-réponses des deux autres pratiques. Leur nature réactive rend compte souvent de la confrontation entre un discours et un contre-discours – ce qui correspond à la définition même de l'argumentation selon Douiry (2016). Dans le cadre de ce technodiscours réactif, le sujet énonciateur tend en outre à enchaîner sur le discours ciblé, en apportant des arguments en soutien de ses affirmations, ainsi qu'à recourir aux propos d'autrui soit pour les réfuter soit pour s'aligner à ceux-ci et consolider son propre dire. Bien que, dans les tweets originaux aussi, l'emploi de hashtags et de mentions contribue à expliciter les dimensions dialogiques interdiscursive et interlocutive des technodiscours, une prise de position dialogique interlocutive plus marquée est propre aux tweets-réponses en raison de leur statut intrinsèquement interlocutif.

Enfin, pour ce qui concerne les threads, il faut avant tout rappeler deux précisions. En premier lieu, il s'agit de la pratique technodiscursive la moins représentée à l'intérieur du corpus TWS, comportant ainsi moins de données à soumettre à l'analyse. En deuxième lieu, l'étude n'a pas nécessairement concerné tous les tweets appartenant à un même thread. Eu égard aux critères adoptés pour la constitution du corpus, il se peut en effet que seulement certains des tweets d'un thread aient été sélectionnés. L'approche contextualisée a néanmoins permis d'analyser ces technodiscours dans leurs interrelations dans le fil discursif d'appartenance, faisant ressortir deux conceptions différentes de la pratique sur la base d'usages distincts. D'une part, les internautes n'ont recours à la pratique du thread que pour prolonger un message incomplet compte tenu du nombre limité de caractères disponibles. En l'occurrence, il s'agit donc d'un moyen pour contourner une contrainte technogénérique. De l'autre, les tweets retenus comme appartenant à des threads se caractérisent souvent par des usages plus congruents aux spécificités de la pratique. En particulier, ils résultent de la planification d'un seul discours qui se développe de manière progressive et interreliée au long de plusieurs tweets. Ces tweets se caractérisent ainsi par une structuration discursive, voire même argumentative, plus construite. Comme le montre Longhi (2022 : 132), le thread « engage une dimension explicative, informative, voire militante », et, d'un point de vue textuel, il est « davantage planifié, construit, et s'intègre dans une pratique sociale numérique proche du billet de blog par exemple ».

D'un point de vue énonciatif et pragmatique, les threads se rapprochent davantage des tweets originaux que des tweets-réponses, exception faite pour le recours privilégié à une énonciation à la première personne du singulier caractérisant les tweets originaux. Au contraire, les threads présentent plus fréquemment une prise en charge déclarative qui peut pourtant avoir différentes valeurs pragmatiques en fonction de la portée sémantique et argumentative du technodiscours. Ainsi, alors que les threads qui abordent les aspects factuels de l'affaire impliquent plutôt une visée informative-explicative, ceux qui concernent la polarisation polémique du débat se caractérisent par une dimension argumentative et même militante plus marquée. Dans le cadre de la construction des positionnements discursifs dominants, la prise de position intersubjective de l'instance énonciative joue en effet un rôle de premier plan dans la polarisation. Dans les threads, c'est plutôt le niveau lexical qui véhicule la subjectivité énonciative à travers les formes de référenciation ainsi que leurs contextualisations entraînant surtout des modes de validation modale d'ordre subjectif.

Bilan final

Hétérogénéités intra- et intergénérique : rapprochements et divergences

La constitution d'un corpus impliquant divers niveaux d'hétérogénéité a permis d'analyser autant de niveaux de variation dans les caractéristiques linguistiques, discursives et technodiscursives des positionnements discursifs qui relèvent à la fois des spécificités génériques et sous-génériques des discours sélectionnés. Cette diversification témoigne d'ailleurs de l'importance des phases de conception et de structuration du corpus afin de créer les conditions nécessaires pour avoir accès aux objets d'étude visés. Ensuite, des présupposés théoriques relatifs aux spécificités sous-génériques ont pu être vérifiés sur corpus à travers des retours quantifiés, avant d'être approfondis en contexte de manière qualitative, comme cela a été le cas pour la distinction sous-générique interne au discours journalistique entre articles d'information et articles d'opinion (4.1.1 et 4.1.2). D'autre part, la complémentarité, la dynamicité et l'adaptabilité des démarches d'analyse en fonction des conditions génériques et sous-génériques ont abouti à des résultats qui ont tantôt confirmé des hypothèses de recherche, tantôt révélé des traits caractéristiques moins attendus au départ, notamment quant à la caractérisation des différentes pratiques technodiscursives du corpus TWS (5.1.2 et 5.1.3).

Les traits linguistiques et les procédés discursifs caractérisant les positionnements discursifs dominants dans les deux macro-genres de discours analysés montrent l'influence des contraintes et des spécificités génériques et sous-génériques respectives, tout en révélant également des aspects récurrents et partagés dans leur polarisation.

Dans le cas du discours journalistique, les articles d'information ne sont pas censés exprimer une prise de position explicite par rapport aux événe-

ments qu'ils construisent et médiatisent. Toutefois, l'étude du sous-corpus PN-INFO a mis en évidence que les choix de référencement, leurs fréquences d'emploi, les procédés de mise en discours, la mise en avant quantitative de certains aspects tout comme la sous-représentation d'autres participent à une construction discursive orientée des objets de discours et, plus en général, des événements. Ceux-ci sont en effet construits en discours à travers une perspectivisation spécifique, qui peut s'aligner sur un positionnement plutôt que sur l'autre. L'analyse multidimensionnelle combinée du sous-corpus PN-INFO a permis de distinguer quatre positionnements discursifs dominants selon une opposition polarisée graduelle sur la base de deux critères principaux : les formes de référencement privilégiées par les différents journaux et leurs fréquences (4.1.2), d'un côté, et la sélection des énonciateurs seconds et des passages à représenter de manière autonymique (4.1.3), de l'autre.

L'analyse comparative de l'hétérogénéité métaénonciatrice d'une perspective sous-générique (PN-INFO vs PN-OP ; cf. 4.3.2 et 4.3.3) a servi ensuite à des fins de validation des résultats relatifs aux divers positionnements dont se rapprochent les articles d'information. Dans ce cas aussi, deux paramètres d'analyse se sont avérés révélateurs : les sources énonciatives sélectionnées par les différents journaux pour s'exprimer dans les articles d'opinion, ainsi que les aspects de la controverse principalement mis en relief.

La comparaison intragénérique entre PN-INFO et PN-OP au niveau de la polarisation énonciative-pragmatique a en outre montré la centralité des choix de référencement, de leurs mises en contexte tout comme de leurs fréquences dans la construction des positionnements dominants. Elle a également mis en évidence le rôle premier du plan sémantique-référentiel du positionnement discursif dans ce genre de discours. S'il va de soi que la construction du positionnement discursif résulte de l'interrelation entre le plan sémantique-référentiel et le plan énonciatif-pragmatique (Cagninelli 2022), il faut souligner que l'un des deux peut avoir néanmoins une fonction prééminente. Dans le cas du discours journalistique d'information, ses caractéristiques énonciatives et pragmatiques (3.1.2.2) n'entraînent pas la présence de prises de position explicites de la part de l'énonciateur premier. Cependant, le niveau lexico-sémantique joue un rôle de premier plan dans la mise en discours de l'événement, qui est construit à partir d'une perspective déterminée, orientant en conséquence la réception-interprétation. Ainsi, les positionnements discursifs dominants dans le discours journalistique précèdent, d'une part, de la

perspectivation privilégiée dans la construction des objets de discours et, plus en général, de l'événement, et, de l'autre, des focalisations proposées.

En effet, la dimension dialogique du discours journalistique est également significative dans la construction des positionnements principaux, à la suite notamment des opérations de sélection et des modes de représentation des discours d'énonciateurs seconds (cf. 3.1.2.3 et 4.1.3). Étant donné que les différents points de vue sur la question sont au moins mentionnés, la fréquence de représentation des discours qui se rapprochent d'un certain positionnement plutôt que de l'autre favorise une mise en discours orientée de l'événement, influençant à son tour sa réception-interprétation. Dans les articles d'information, il est ainsi possible de remarquer une prise de position indirecte vis-à-vis d'autres discours circulant dans l'espace public sur la même question. En l'occurrence, l'alignement interdiscursif sur l'un des deux positionnements antagonistes se réalise par la convocation des propos de sources énonciatives secondes qui s'y relie. Ce n'est donc pas le(s) journaliste(s) à l'origine de l'article qui se rapproche ou prend de la distance de l'un des deux positionnements par les constructions discursives employées. C'est plutôt la sélection des propos et des sources énonciatives convoquées par la représentation de leurs dires qui est responsable de l'alignement ou du désalignement du discours sur l'une des constructions discursives polarisées et, en conséquence, des positionnements discursifs dominants. Une approche quantifiée du niveau global d'un corpus impliquant une hétérogénéité métaénonciatrice représente finalement une condition nécessaire pour identifier ces formes de perspectivation et de focalisation de l'événement à travers plusieurs journaux.

Dans le cas des technodiscours du corpus TWS, les positionnements discursifs dominants ont été en revanche analysés principalement dans une perspective contrastive intragénérique d'ordre technodiscursif, bien que les tweets originaux aient représenté un point de repère essentiel en raison de leur autonomie sémantique.

Les spécificités technodiscursives et la nature pragmatique des technodiscours de la plateforme Twitter ont toutefois diverses répercussions sur les ressources linguistiques et technolangagières ainsi que sur les procédés (techno)discursifs participant à la construction des positionnements principaux. Alors que l'analyse inductive-quantifiée a permis de repérer les rapprochements et les divergences les plus marquées, déterminées par les spécificités sous-technogénériques des trois pratiques, l'analyse déductive-contextualisée a révélé aussi bien les aspects partagés que des différences plus fines dans les usages de ressources similaires. La brièveté de

ces technodiscours et leur tendance fréquente à commenter l'actualité et à exprimer des opinions, des émotions et des expériences subjectives déterminent un large emploi d'éléments linguistiques connotés ainsi que des mises en discours polarisées, donnant ainsi corps à des prises de position assez explicites et tranchées. De fait, dans les trois pratiques, les positionnements discursifs dominants reposent principalement sur le choix de formes de référenciation dont le sémantisme et/ou la mise en discours se caractérisent par une axiologisation polarisée. De même que dans le discours journalistique, certains aspects de la question ou certains objets de discours sont par ailleurs plus souvent abordés que d'autres, comme en témoigne aussi la surreprésentation de certaines expressions et la sous-représentation d'autres.

Cependant, dans les tweets simples, le plan sémantique-référentiel non seulement contribue à l'identification de la construction orientée des objets de discours, mais il permet notamment d'accéder à une explicitation plutôt marquée du positionnement discursif. Le technodiscours se rapproche en effet d'une position plutôt que de l'autre à travers la présence de formes de référenciation et de contextualisation assez clairement polarisées. Dans ce technogène de discours, les ressources axiologiques et connotées sont beaucoup plus présentes et il en est de même pour la dimension émotionnelle qui est plus marquée, donnant ainsi forme à des positionnements plus explicites et, souvent, tranchés. Cela résulte de la portée pragmatique des tweets simples qui constituent généralement des discours de commentaire et donc d'expression des opinions, pouvant se caractériser aussi par des tons conflictuels et même violents, comme c'est le cas en particulier de la pratique réactive-interlocutive des tweets-réponses.

À la différence du corpus PN, la construction des positionnements discursifs principaux dans les tweets simples analysés présente des différences remarquables également au niveau énonciatif-pragmatique, qui entraîne des variations intragénériques significatives liées tout particulièrement aux spécificités sous-technogénériques des trois pratiques (tweets originaux, tweets-réponses et tweets appartenant à un thread). En d'autres termes, si le plan énonciatif-pragmatique ne joue pas un rôle déterminant dans la polarisation des positionnements discursifs, il s'avère néanmoins essentiel dans la caractérisation de leur construction technodiscursive dans les trois pratiques (5.3.2). Ceci nous amène à les considérer comme trois sous-technogènes de discours différents. Il faut d'ailleurs souligner la centralité de la dimension énonciative dans la pratique des tweets-réponses, où elle participe à la construction d'une prise de position dialogique interlocutive

marquant, dans la plupart des cas, le désalignement de l'énonciateur par rapport au discours antérieur.

En bref, nous pouvons affirmer que le niveau lexico-sémantique a une fonction primaire dans la construction polarisée des positionnements discursifs principaux dans le technogénre des tweets simples. Les traits caractérisant la *position référentielle-idéologique* (Cagninelli 2022) sont par ailleurs plutôt constants d'un point de vue intragénérique mais aussi intergénérique. En revanche, les divergences principales des positionnements discursifs dans les trois pratiques concernent plutôt le niveau énonciatif-pragmatique (à la fois la *position intradiscursive* et la *position interdiscursive* – cf. Cagninelli 2022), qui contribue ainsi au repérage des traits distinctifs intragénériques, influencés notamment par les divers statuts technodiscursifs. Le niveau énonciatif-pragmatique du positionnement discursif peut donc contribuer au processus de caractérisation technodiscursive des différents types de tweets simples.

D'une perspective intergénérique, on remarque ainsi deux rapprochements principaux. Des traits communs dans la construction des positionnements discursifs caractérisent les articles d'opinion et les tweets simples, notamment les tweets originaux. Cette similitude s'explique par leur nature énonciative plus subjectivée et une orientation pragmatique plus axée sur l'expression de la subjectivité de l'énonciateur dans le cadre d'un discours de commentaire, bien qu'avec des compétences en la matière et des autorités énonciatives différentes. L'analyse inductive-quantifiée a en effet montré une diversification et une articulation thématiques plus développées dans ces formes discursives. Une richesse lexicale et une prise en charge énonciative explicitement subjective sont en outre ressorties de l'analyse de la construction des positionnements discursifs aux niveaux locaux du corpus.

Des choix lexicaux et des constructions discursives plus ouvertement axiologiques et polarisés rapprochent en revanche les discours autonymiques, représentés au sein des articles d'information, des technodiscours du corpus TWS. En l'occurrence, si bon nombre de tweets originaux proposent une construction orientée de la question et des prises de position ouvertes, cet aspect est largement prédominant dans les tweets-réponses. Ceci est dû à leur caractère réactif et interlocutif, comportant une prise de position dialogique souvent à valeur dissensuelle, à travers un désalignement sur le discours auquel ils réagissent et qui est fréquemment contesté. Quoique minoritaires dans le corpus, les threads présentent des tendances similaires aux tweets originaux : ils abordent tantôt des aspects plus factuels avec un ton informatif qui privilégie une certaine mise en perspective de

la question, tantôt des aspects plus polémiques à travers des mises en discours plus ouvertement partisans. Néanmoins, ils impliquent également des traits distinctifs propres. Ces rapprochements intergénériques mettent en évidence le rôle significatif que les discours autonymiques entre guillemets peuvent jouer au sein des articles d'information, participant ainsi de leur « dimension argumentative » (Amossy 2016) à travers la représentation d'une perspectivation orientée et plus ouvertement axiologisée de la question.

On constate enfin des tendances intergénériques assez stables quant aux formes de référencement les plus fréquentes et les moins connotées. Des usages assez récurrents à travers les genres et les sous-(techno)genres de discours analysés témoignent en effet de la circulation intergénérique des polarisations discursives ainsi que de leur stabilisation interdiscursive.

En guise de conclusion

Dans ce travail, la conception et la constitution d'un corpus discursif pluri-générique, à la fois cohérent et hétérogène, ont créé les conditions nécessaires pour analyser les diverses manifestations des positionnements discursifs dans le débat public sur la fin de vie lié à l'« affaire Vincent Lambert », en tenant compte des spécificités intra- et intergénériques des discours envisagés. Cet ouvrage repose en effet sur une conception du corpus comme un « terrain » d'enquête sociodiscursive (Longhi 2018), qui donne accès à différentes pratiques, formes et stratégies de discursivisation relevant d'une appréhension spécifique du réel. Constitué de données authentiques, le corpus s'avère fondamental pour permettre à l'analyste d'aborder la matérialité linguistique en articulation avec son ancrage sociodiscursif, à partir d'une perspective théorico-méthodologique déterminée.

La structuration du corpus par le biais des métadonnées a notamment matérialisé différents déterminants discursifs et extradiscursifs des textes sélectionnés, en donnant corps à plusieurs niveaux d'hétérogénéité au service de méthodes d'exploration également hétérogènes pour accéder aux dynamiques discursives. La façon dont le corpus est structuré établit donc une matrice exploratoire qui concrétise les niveaux et angles d'approche envisagés dans la phase de sa conception en fonction des objectifs de l'étude. Le corpus ainsi constitué concrétise un emboîtement de contextes, donnant accès aux multiples trajectoires de la construction du sens. Il répond ainsi à plusieurs exigences posées par les fondements théoriques et méthodologiques à l'origine de la conception du langage et de l'approche sur lesquels se fonde cette étude, ainsi qu'en fonction de la nature de l'objet d'étude visé. Finalement, le corpus a constitué à la fois l'objet et le support de l'analyse.

Afin d'illustrer la réflexion théorico-méthodologique développée dans la première partie de l'ouvrage, le corpus a été analysé à travers plusieurs niveaux et degrés d'hétérogénéité générique et sous-générique, avec pour objectif de repérer les spécificités intra- et intergénériques des textes qu'il rassemble, ainsi que leur influence dans la construction des positionnements discursifs dominants. L'hétérogénéité comparable des données est en outre allée de pair avec l'hétérogénéité complémentaire des méthodes adoptées pour les analyser. La démarche multidimensionnelle combinée que nous avons proposée a permis d'explorer les différents niveaux du corpus à partir d'angles d'approche divers, en articulant des explorations outillées avec des analyses qualitatives, avant de mettre en relation les résultats de chacune pour réaliser finalement une étude contrastive intra- et intergénérique. Une démarche comparable a été suivie pour l'étude de chaque (sous-)corpus générique, tout en adaptant les méthodes d'exploration et les éléments à envisager en fonction des contraintes et des spécificités propres à chaque genre et sous-genre de discours. Les traits distinctifs de chacun ont donc eu des répercussions non seulement sur la construction des positionnements discursifs dominants, comme l'a montré l'étude du corpus, mais aussi sur la démarche d'analyse à adopter et sur les outils à privilégier.

Dans les deux cas, l'étude des corpus PN et TWS a été effectuée en suivant une démarche en trois étapes. Le niveau global du corpus et sa structuration interne ont été tout d'abord examinés en recourant à des analyses outillées inductives et quantifiées. Elles ont montré l'apport épistémologique des explorations textométriques à l'analyse de corpus discursifs, en permettant de repérer des traits saillants au niveau aussi bien lexico-sémantique qu'énonciatif-pragmatique. Cette première approche a permis d'avoir accès à un aperçu général des (sous-)corpus (4.1 et 5.1), et de faire ressortir les apparentements et les contrastes intragénériques. À partir de ces résultats, il a également été possible d'identifier les ressources à creuser davantage par l'analyse contextualisée, qui a représenté la deuxième phase de l'analyse.

L'étude déductive et contextualisée des données dans les deux corpus macro-génériques (4.2 et 5.2) a été ainsi ajustée en fonction des caractéristiques principales mises en évidence lors de la première étape. À côté de quelques éléments distinctifs de chaque (sous-)(techno)genre, les procédés de mise en discours de deux des objets de discours significatifs du débat public ont été explorés et comparés en contexte dans une perspective intragénérique. L'analyse a été réalisée sur différents niveaux d'hétérogénéité : l'hétérogénéité métaénonciatrice dans le corpus PN-INFO, comparée en-

suite au niveau sous-générique à celle propre au corpus PN-OP, ainsi que l'hétérogénéité sous-technogénérique au sein du corpus TWS.

Dans la troisième et dernière étape de la démarche d'analyse, les résultats de l'analyse déductive-contextualisée ont été articulés avec ceux de l'analyse inductive-quantifiée (4.3 et 5.3). De fait, les phénomènes discursifs tels que le positionnement discursif gagnent à être analysés en corpus à travers des pratiques herméneutiques et heuristiques qui impliquent une constante déconstruction et reconstruction des textes composant le corpus (cf. Mayaffre 2002b, 2010). La part de subjectivité intrinsèque aux processus d'interprétation est ainsi circonscrite par la réalisation d'une étude ancrée sur des mesures quantifiées et attestées, qui multiplie également les angles d'approche du corpus. Les allers-retours entre les divers niveaux du corpus abordés par des méthodes hétérogènes mais complémentaires favorisent en effet l'objectivation de l'accès aux données. L'hétérogénéité méthodologique concernant aussi bien la diversité complémentaire des méthodes que la multiplication des niveaux d'analyse (cf. Garric 2012) assure en outre une exploration riche et modulable en fonction non seulement des objets et des objectifs de l'étude, mais aussi des résultats partiels obtenus au cours de l'analyse.

Les trois phases de la démarche d'exploration du corpus favorisent finalement l'appréhension de phénomènes discursifs qui résultent de l'interrelation entre divers niveaux de la textualité. Si ces derniers sont inévitablement étudiés de manière séparée, la mise en relation des uns avec les autres ainsi qu'une synthèse contrastive – intra- et intergénérique – sont nécessaires pour saisir les phénomènes tels que le positionnement discursif dans leur globalité. L'articulation de l'analyse en trois étapes donne lieu à une démarche qui va du niveau global du sous-corpus à ses niveaux locaux pour revenir enfin à la dimension globale. L'articulation global-local (cf. aussi Rastier 2011 ; Mayaffre 2010) représente un aspect fondamental de l'approche méthodologique illustrée dans le présent ouvrage. Comme le soutient Rastier (2011), le global détermine le local, mais il faut également tenir compte du fait que le local favorise en retour la compréhension et la caractérisation pragma-discursive des macro-caractéristiques du global. Ainsi, l'étude inductive-quantifiée du niveau global du corpus permet de faire ressortir les caractéristiques génériques et sous-génériques du corpus. En revanche, l'observation contextualisée des niveaux locaux permet de vérifier les hypothèses de recherche et d'approfondir les résultats des analyses inductives-quantifiées, en accédant à des dynamiques discursives attestées. D'où la portée à la fois heuristique et objectivante de cette démarche

d'analyse, dans laquelle l'hétérogénéité comparable des données contribue à mettre en évidence non seulement les spécificités intragénériques des deux sous-corpus macro-génériques principaux (PN et TWS), mais aussi les constantes et notamment les divergences intergénériques.

L'identification des traits distinctifs des divers genres et sous-genres de discours par l'exploration du corpus a joué un rôle important pour saisir les variations intra- et intergénérique des positionnements discursifs principaux. L'analyse a ainsi fait ressortir l'influence des spécificités et des contraintes macro- et sous-génériques sur les ressources et les procédés les plus employés qui participent à la construction de ce phénomène discursif. Ces déterminants influent sur les éléments linguistiques et les niveaux textuels et discursifs privilégiés dans la construction des positionnements discursifs, où chacune des trois formes de prise de position qui y conflue peut avoir un rôle plus ou moins significatif et prédominant selon les différents genres et sous-(techno)genres de discours (cf. aussi Cagninelli 2022). Par ailleurs, les variations intra- et intergénériques repérées montrent non seulement l'impact des spécificités et des contraintes de différents genres et sous-genres de discours sur les ressources langagières, les formes et stratégies de mise en discours, et les procédés discursifs de textualisation, mais aussi les répercussions qu'ils provoquent sur les divers niveaux de la textualité.

Parallèlement aux différences intra- et intergénériques, l'analyse a également mis en évidence certaines constantes surtout d'ordre lexico-sémantique et de nature polarisée, qui ont été employées transversalement aux diverses productions analysées de manière aussi bien intragénérique qu'intergénérique. Ce degré de stabilité plus élevé propre au plan lexico-sémantique des discours – et par voie de conséquence des positionnements discursifs – pourrait en outre représenter une voie d'accès à l'étude interdiscursive de la polarisation des formations discursives lors du débat public envisagé. Si la notion de *formation discursive* a marqué significativement les débuts des recherches en analyse de discours de tradition française (cf. Mazière 2018 ; Charaudeau & Maingueneau 2002), il nous semble qu'elle gagnerait à être mise en relation avec la question des genres de discours. Il serait en effet intéressant d'observer si des divergences lexico-sémantiques mineures sont repérables au sein d'une même formation discursive d'un point de vue intra- et intergénérique. Par ailleurs, on pourrait se demander si les variations intra- et intergénériques repérées au niveau énonciatif-pragmatique sont également constantes dans deux formations discursives distinctes ou si au contraire elles participent à leur tour à leur caractérisation.

L'articulation entre l'hétérogénéité comparable des données, l'hétérogénéité complémentaire des méthodes et la comparaison contrastive entre plusieurs formes d'hétérogénéité matérialisées en corpus a finalement été fondamentale pour mettre en évidence le rôle des spécificités et des contraintes macro-générique et sous-(techno)génériques sur la construction d'un phénomène discursif tel que le positionnement discursif. La réflexion théorico-méthodologique illustrée dans la première partie de l'ouvrage et son application à l'étude du corpus dans la deuxième partie ont montré la centralité des questions à la fois théoriques et méthodologiques relevées par la création du corpus, d'une part, et la portée heuristique et objectivante d'une démarche multidimensionnelle combinée, appliquée à l'analyse des variations intra- et intergénériques, de l'autre. Par le présent ouvrage, nous nous sommes attachée à mettre en avant l'intérêt d'une approche discursive qui prend en compte les divers facteurs sociodiscursifs et notamment ceux d'ordre générique dans les processus de construction du sens. Ce travail a proposé une perspective interprétative outillée dans laquelle la subjectivité interprétante est à la fois circonscrite et corroborée par l'outillage informatique et par la diversification des angles d'approche des données. Des notions traditionnelles de l'analyse du discours telle que celle de *genre de discours* pourraient être testées, voire réactualisées, par l'adoption de cette approche méthodologique où les formes d'hétérogénéité des données et des méthodes constituent un point d'ancrage pour l'analyse des phénomènes discursifs à la lumière de plusieurs formes de contextualisation et à travers plusieurs niveaux et paliers.

Références bibliographiques

- Achard-Bayle Guy & Pešek Ondřej (2022), *Le paragraphe et l'organisation thématico-graphique du texte dans les nouveaux écrits numériques*, « Le français moderne. Revue de linguistique française », 90(1), 75-106.
- Adam Jean-Michel (1997), *Unités rédactionnelles et genres discursifs : Cadre général pour une approche de la presse écrite*, « Pratiques », 94(1), 3-18.
- Adam Jean-Michel (2020a), *La linguistique textuelle : introduction à l'analyse textuelle des discours*, Armand Colin, Malakoff.
- Adam Jean-Michel (2020b), *Postface. Le texte est-il soluble dans le textiel ?*, « Corela », HS-33, doi: 10.4000/corela.11938.
- Amossy Ruth (2014), *Apologie de la polémique*, Presses universitaires de France, Paris.
- Amossy Ruth (2016), *L'argumentation dans le discours*, Armand Colin, Paris.
- Angermüller Johannes (2017), *Renouons avec les enjeux critiques de l'Analyse du Discours. Vers les Études du discours*, « Langage et société », 160-161(2), 145-161.
- Authier-Revuz Jacqueline (1984), *Hétérogénéité(s) énonciative(s)*, « Langages », 73, 98-111.
- Authier-Revuz Jacqueline (2020), *La représentation du discours autre*, De Gruyter, Berlin-Boston.
- Badouard Romain (2016), « Je ne suis pas Charlie ». *Pluralité des prises de parole sur le web et les réseaux sociaux*, in P. Lefébure & C. Sécaïl (dir.), *Le défi Charlie. Les médias à l'épreuve des attentats*, Lemieux Éditeur, Paris, 187-219.
- Badouard Romain, Mabi Clément & Monnoyer-Smith Laurence (2016), *Le débat et ses arènes : À propos de la matérialité des espaces de discussion*, « Questions de communication », 30, 7-23.
- Bakhtine Mikhaïl (1970) [1963], *La poétique de Dostoïevski*, Points Seuil, Paris.

- Bakhtine Mikhaïl (1984), *Les genres du discours*, in *Esthétique de la création verbale*, Gallimard, Paris, 265-308.
- Beacco Jean-Claude (2002), *Corpus*, in P. Charaudeau & D. Maingueneau (dir.), *Dictionnaire d'analyse du discours*, Seuil, Paris, 148-151.
- Benveniste Émile (1966), *Problèmes de linguistique générale. Tome 1* (2006 éd.), Gallimard, Paris.
- Benzécri Jean-Paul (1969), *Statistical analysis as a tool to make patterns emerge from clouds*, in S. Watanabe (dir.), *Methodology of Pattern Recognition*, Academic Press, New York, 35-74.
- Benzécri Jean-Paul (1973), *L'Analyse des Données. Tome 1 : La Taxinomie. Tome 2 : L'Analyse des Correspondances*, Dunod, Paris.
- Biber Douglas (1993), *Representativeness in Corpus Design*, « *Literary and Linguistic Computing* », 8(4), 243-257.
- Biber Douglas, Conrad Susan & Reppen Randi (1998), *Corpus Linguistics: Investigating Language Structure and Use*, Cambridge University Press, New York.
- Bibié Laetitia (2019), *Utilisation de ok sur Twitter, entre (re)présentation de soi et stabilisation énonciative*, « *Lexique* », 25, 57-75.
- Bigey Magali & Simon Justine (2018), *Analyse des discours d'escorte de communication sur Twitter : essai de typologie des tactiques d'accroches et de mentions*, in A. Mercier & N. Pignard-Cheynel (dir.), *#info : Commenter et partager l'actualité sur Twitter et Facebook*, Éditions de la Maison des sciences de l'homme, Paris, 55-86.
- Bonnafous Simone & Tournier Maurice (1995), *Analyse du discours, lexicométrie, communication et politique*, « *Langages* », 117, 67-81.
- Bres Jacques (2019), *Concession*, in J. Bres, A. Nowakowska, & J.-M. Sarale (dir.), *Petite grammaire alphabétique du dialogisme*, Classiques Garnier, Paris, 85-93.
- Bres Jacques, Nowakowska, Aleksandra & Sarale Jean-Marc (dir.) (2019), *Petite grammaire alphabétique du dialogisme*, Classiques Garnier, Paris.
- Brunet Étienne (2016), *Tous comptes faits. Écrits choisis*, Bénédicte Pincemin (éd.), Champion, Paris.
- Bruns Axel (2018), *Gatewatching and News Curation: Journalism, Social Media, and the Public Sphere*, Peter Lang, New York.
- Cagninelli Claudia (2022), *Positionnements discursifs dans le débat public sur la fin de vie : variations génériques entre (inter)subjectivité et interdiscursivité*, Thèse de doctorat, Università di Modena e Reggio Emilia/CY Cergy Paris Université.
- Cagninelli Claudia (2024a), *Discours numériques et approches outillées : quelques réflexions sur les apports des métadonnées*, « *Lingue e Linguaggi* », 65, 385-412.

- Cagninelli Claudia (2024b), *Controverses sociales et genres de discours : constantes et divergences des positionnements discursifs dans le débat public*, « Cahiers de praxématique » [En ligne] 81. doi : <https://doi.org/10.4000/praxematique.9120>.
- Cagninelli Claudia (2024c), *Effets rhétoriques de l'augmentation énonciative des discours numérique. Le cas des "retweets avec citation"*, « Lingue e Linguaggi » 62, 205-226.
- Cagninelli Claudia & Taglioli Massimiliano (2022), *Visaneco* [Computer software].
- Charaudeau Patrick (2001), *Visées discursives, genres situationnels et construction textuelle*, in M. Ballabriga (éd.), *Analyse des discours. Types et genres*, Éd. Universitaires du Sud, Toulouse, consulté sur <http://www.patrick-charaudeau.com/Visees-discursives-genres,83.html>.
- Charaudeau Patrick (2009), *Dis-moi quel est ton corpus, je te dirai quelle est ta problématique*, « Corpus », 8, 37-66.
- Charaudeau Patrick (2011), *Les médias et l'information. L'impossible transparence du discours*, De Boeck Université, Bruxelles.
- Charaudeau Patrick & Maingueneau Dominique (dir.) (2002), *Dictionnaire d'analyse du discours*, Seuil, Paris.
- Condamines Anne (2005), *Linguistique de corpus et terminologie*, « Langages », 157, 36-47.
- Dalbera Jean-Philippe (2002), *Le corpus entre données, analyse et théorie*, « Corpus », 1, doi: 10.4000/corpus.10.
- Demonet Michel, Geffroy Annie, Gouazé Jean, Lafon Pierre, Mouillaud Maurice & Tournier Maurice (1975), *Des tracts en Mai 68. Mesures de vocabulaire et de contenu*, Presses de la FNSP/Armand Colin, Paris.
- Doury Marianne (2016), *Argumentation : Analyser textes et discours observations et analyses, méthodologie pratique, exercices corrigés*, Armand Colin, Paris.
- Ducrot Oswald (1984), *Le dire et le dit*, Minuit, Paris.
- Emerit Laetitia (2016), *La notion de lieu de corpus : un nouvel outil pour l'étude des terrains numériques en linguistique*, « Corela », 14, <https://doi.org/10.4000/corela.4594>.
- Garric Nathalie (2012), *Construire et maîtriser l'hétérogénéité par la variation des données, des corpus et des méthodes*, « Langages », 187, 73-92.
- Garric Nathalie & Longhi Julien (2012), *L'analyse de corpus face à l'hétérogénéité des données : d'une difficulté méthodologique à une nécessité épistémologique*, « Langages », 187, 3-11.
- Garric Nathalie & Longhi Julien (2013), *Atteindre l'interdiscours par la circulation des discours et du sens*, « Langage et société », 144, 65-83.
- Gibson James J. (1979), *The ecological approach to visual perception*, Houghton Mifflin, Boston, MA.

- Graumann Carl F. & Kallmeyer Werner (2002), *Perspective and Perspectivation in Discourse. An Introduction*, in C.F. Graumann & W. Kallmeyer (dir.), *Perspective and Perspectivation in Discourse*, John Benjamins Publishing Company, Amsterdam-Philadelphia, 1-11.
- Guilhaumou Jacques (2002), *Le corpus en analyse de discours : perspective historique*, « Corpus », 1, <http://journals.openedition.org/corpus/8>.
- Guilhaumou Jacques, Maldidier Denise & Robin Régine (1994), *Discours et archive : expérimentations en analyse du discours*, Mardaga, Liège.
- Habert Benoît, Nazarenko Adeline & Salem André (1997), *Les linguistiques de corpus*, Armand Colin, Paris.
- Heiden Serge, Magué Jean-Philippe & Pincemin Bénédicte (2010), *TXM : Une plateforme logicielle open-source pour la textométrie – conception et développement*, in S. Bolasco, I. Chiari & L. Giuliano (dir.), *Proc. of 10th International Conference on the Statistical Analysis of Textual Data – JADT 2010*, LED Edizioni Universitarie, Milano, vol. 2, 1021-1032.
- Jackiewicz Agata & Vidak Marko (2014), *Étude sur les mots-dièse*, in F. Neveu, P. Blumenthal, L. Hriba, A. Gerstenberg, J. Meinschaefter, & S. Prévost (dir.), *SHS Web of Conferences*, 8, 2033-2050.
- Julliard Virginie (2016), *#Theoriedugenre : comment débat-on du genre sur Twitter ?*, « Questions de communication », 30, 135-157.
- Krieg-Planque Alice (2009), *La notion de « formule » en analyse du discours. Cadre théorique et méthodologique*, Presses universitaires de Franche-Comté, Besançon.
- Kwak Haewoon, Lee Changhyun, Park Hosung & Moon Sue (2010), *What is Twitter, a social network or a news media?*, in *Proceedings of the 19th International Conference on World Wide Web – WWW '10*, 591-600.
- Lafon Pierre (1980), *Sur la variabilité de la fréquence des formes dans un corpus*, « Mots. Les langages du politique », 1, 127-165.
- Lebart Ludovic & Salem André (1994), *Statistique textuelle*, Dunod, Paris.
- Lebart Ludovic, Pincemin Bénédicte & Poudat Céline (2019), *Analyse des données textuelles*, Presses de l'Université du Québec, Québec.
- Leblanc Jean-Marc (2017), *Approfondissement – En savoir plus sur l'analyse factorielle des correspondances*, in É. Née (dir.), *Méthodes et outils informatiques pour l'analyse des discours*, Presses universitaires de Rennes, Rennes, 218-228.
- Leblanc Jean-Marc, Fleury Serge & Née Émilie (2017), *Quels outils logiciels et pour quoi faire ?*, in É. Née (dir.), *Méthodes et outils informatiques pour l'analyse des discours*, Presses universitaires de Rennes, Rennes, 123-161.
- Longhi Julien (2011), *Visées discursives et dynamiques du sens commun*, L'Harmattan, Paris.

- Longhi Julien (2012), *Discours, style, format : contraintes et niveaux de structuration de la textualité des Tweets de Mouloud*, in *Actes du 3^e Congrès Mondial de Linguistique Française*, SHS Web of Conferences, 1127-1141.
- Longhi Julien (2015), *L'acte de nommer comme constitution de formes : discursivité de la production du sens*, « *Langue française* », 188, 121-136.
- Longhi Julien (2018), *Du discours comme champ au corpus comme terrain : contribution méthodologique à l'analyse sémantique du discours*, L'Harmattan, Paris.
- Longhi Julien (2019), *Contours, perspectives et tensions des "humanités numériques"*, « *Sens-Dessous* », 24, 123-135.
- Longhi Julien (2020a), *Théoriser le dynamique, modéliser la variation, et outiller l'herméneutique : le(s) sens en question(s) – Theorising The Dynamic, Modeling the Variation, and Equipping Hermeneutics: The Meaning(s) in Question*, « *Critical Hermeneutics* », 4, 35-64.
- Longhi Julien (2020b), *Explorer des corpus de tweets : du traitement informatique à l'analyse discursive complexe*, « *Corpus* », 20, <https://doi.org/10.4000/corpus.4567>.
- Longhi Julien (2021), *Du corpus réflexif au corpus réfléchi : la plateforme #Idéo2017 pour extraire contextuellement les pratiques citationnelles et analyser la circulation des discours politiques sur Twitter*, « *Le discours et la langue* », 12(2), 99-114.
- Longhi Julien (2022), *Le thread, un texte cousu de fil numérique ?*, « *Le Français Moderne – Revue de linguistique Française* », 90(1), 107-134.
- Magué Jean-Philippe, Rossi-Gensane Nathalie & Halté Pierre (2020), *De la segmentation dans les tweets : signes de ponctuation, connecteurs, émoticônes et émojis*, « *Corpus* », 20, <https://doi.org/10.4000/corpus.4619>.
- Maingueneau Dominique (2007), *Genres de discours et modes de généricité*, « *Le français aujourd'hui* », 159, 29-35.
- Maingueneau Dominique (2014a), *Discours et analyse du discours : une introduction*, Armand Colin, Paris.
- Maingueneau Dominique (2014b), *Aux limites de la généricité*, in M. Monte & G. Philippe (dir.), *Genres & textes* [En ligne], Presses universitaires de Lyon, <https://doi.org/10.4000/books.pul.3054>.
- Maingueneau Dominique (2021), *Analyser les textes de communication*, Armand Colin, Paris.
- Maingueneau Dominique (2022), *Le texte : une autorité menacée ?*, « *Le français moderne. Revue de linguistique française* », 90, 6174.
- Marchand Pascal & Ratinaud Pierre (2017), *Entre distinctivité et acceptabilité. Les contenus des sites Web de partis politiques*, « *Réseaux* », 204, 71-95.
- Mayaffre Damon (2002a), *Les corpus réflexifs : entre architextualité et hypertextualité*, « *Corpus* », 1, doi : <https://doi.org/10.4000/corpus.11>.

- Mayaffre Damon (2002b), *L'Herméneutique numérique*, « L'Astrolabe. Recherche littéraire et Informatique », numéro spécial, 151-161. <hal-00586512>.
- Mayaffre Damon (2005), *Rôle et place des corpus en linguistique : réflexions introductives*, in *Actes des Journées d'Etude TOUlousaines JETOU 2005*. Toulouse, 5-17.
- Mayaffre Damon (2008), *De l'occurrence à l'isotopie : Les co-occurrences en lexicométrie*, « Syntaxe et sémantique », 9, 53-72.
- Mayaffre Damon (2010), *Vers une herméneutique matérielle numérique. Corpus textuels, Logométrie et Langage politique*, Université Nice Sophia Antipolis, <https://theses.hal.science/tel-00655380>.
- Mayaffre Damon, Pincemin Bénédicte & Poudat Céline (2019), *Explorer, mesurer, contextualiser. Quelques apports de la textométrie à l'analyse de discours*, « Langue française », 203, 101-115.
- Mazière Francine (2018), *L'analyse du discours : histoire et pratiques*, Presses universitaires de France, Paris.
- McEnery Tony & Hardie Andrew (2012), *Corpus Linguistics: Method, Theory and Practice*, Cambridge University Press, Cambridge.
- Mellet Sylvie (2002), *Corpus et recherches linguistiques : Introduction*, « Corpus », 1, doi: 10.4000/corpus.7.
- Mercier Arnaud (2013), *Twitter l'actualité : usages et réseautage chez les journalistes français*, « Recherches en Communication », 39, 111-132.
- Mercier Arnaud (2015), *Twitter, espace politique, espace polémique. L'exemple des tweet-campagnes municipales en France (janvier-mars 2014)*, « Les cahiers du numérique », 11, 145-168.
- Mercier Arnaud (2018), *Hashtags : tactiques de partages et de commentaires d'informations*, in Mercier A. & N. Pignard-Cheynel (dir.), *#Info: commenter et partager l'actualité sur Twitter et Facebook*, Éditions de la Maison des sciences de l'homme, Paris, 87-129.
- Mercier Arnaud & Pignard-Cheynel Nathalie (dir.) (2018), *#Info : commenter et partager l'actualité sur Twitter et Facebook*, Éditions de la Maison des sciences de l'homme, Paris.
- Meunier Jean-Guy (2020), *La rencontre du sémiotique et du "numérique" : Le rôle d'une modélisation conceptuelle*, « Semiotica », 234, 177-198.
- Micheli Raphaël (2014), *Les émotions dans les discours : Modèle d'analyse, perspectives empiriques*, De Boeck Duculot, Louvain-la-Neuve.
- Moirand Sophie (2007), *Les discours de la presse quotidienne : observer, analyser, comprendre*, Presses universitaires de France, Paris.
- Moirand Sophie (2018), *Le discontinu des catégories linguistiques confronté aux catégories et concepts des analyses du discours et au continu du déroulement de la parole "située"*, « Semiotica », 223, 49-70.

- Moirand Sophie (2020), *Retour sur l'analyse du discours française : suivi de quelques réflexions sur une sémantique du discours en construction*, « Pratiques », 185-186, doi: 10.4000/pratiques.8721.
- Moïse Claudine (2012), *Argumentation, confrontation et violence verbale fulgurante*, « Argumentation et analyse du discours », 8, <https://doi.org/10.4000/aad.1260>.
- Mondada Lorenza & Dubois Danielle (1995), *Construction des objets de discours et catégorisation : une approche des processus de référenciation*, « TRANEL. Travaux Neuchâtois de Linguistique », 23, 273-302.
- Monnier Angeliki, Seoane Annabelle, Hubé Nicolas & Leroux Pierre (dir.) (2021), *Discours de haine dans les réseaux socionumériques*, « Mots. Les langages du politique », 125, 9-14.
- Née Émilie (dir.) (2017), *Méthodes et outils informatiques pour l'analyse des discours*, Presses universitaires de Rennes, Rennes.
- Oger Claire & Ollivier-Yaniv Caroline (2003), *Analyse du discours institutionnel et sociologie compréhensive : vers une anthropologie des discours institutionnels*, « Mots. Les langages du politique », 71, 125-145.
- Paveau Marie-Anne (2010), *Interdiscours et intertexte. Généalogie scientifique d'une paire de faux jumeaux*, in D. Ablali & M. Kastberg Sjöblom (dir.), *Linguistique & littérature : Cluny, 40 ans après*, Presses universitaires de Franche-Comté, Besançon, 93-105.
- Paveau Marie-Anne (2013a), *Genre de discours et technologie discursive : tweet, twittécriture et twittérature*, « Pratiques », 157-158, 7-30.
- Paveau Marie-Anne (2013b), *Analyse discursive des réseaux sociaux numériques* [Dictionnaire]. *Technologies discursives*. Consulté le 30/03/2023 à l'adresse <https://doi.org/10.58079/uowwy>.
- Paveau Marie-Anne (2015), *Ce qui s'écrit dans les univers numériques : Matières technolangagières et formes technodiscursives*, « Itinéraires », 2014-1. <https://doi.org/10.4000/itineraires.2313>.
- Paveau Marie-Anne (2017), *L'analyse du discours numérique : dictionnaire des formes et des pratiques*, Hermann, Paris.
- Pêcheux Michel (1969), *Analyse automatique du discours*, Dunod, Paris.
- Pêcheux Michel (1975), *Les Vérités de La Palice*, Maspero, Paris.
- Pincemin Bénédicte (1999), *Construire et utiliser un corpus : le point de vue d'une sémantique textuelle interprétative*, in A. Condamines, M.-P. Péry-Woodley & C. Fabre (éds), *Atelier Corpus et TAL : pour une réflexion méthodologique (TALN 99)*, Cargèse, 26-36.
- Pincemin Bénédicte (2012a), *Hétérogénéité des corpus et textométrie*, « Langages », 187(3), 13-26.
- Pincemin Bénédicte (2012b), *Sémantique interprétative et textométrie*, « Texto! », XVII(3), 1-21, http://www.revue-texto.net/docannexe/file/3049/pincemin_texto11.pdf.

- Pincemin Bénédicte (2020), *La textométrie en question*, « Le Français Moderne – Revue de linguistique Française », 88(1), 26-43. halshs-02902088.
- Pincemin Bénédicte (2022), *Sémantique textométrique*, in A. Biglari et D. Ducard (éds), *La sémantique au pluriel : théories et méthodes*, Presses universitaires de Rennes, Rennes, 373-396.
- Pincemin Bénédicte, Heiden Serge, Lay Marie-Hélène, Leblanc Jean-Marc & Viprey, Jean-Marie (2010), *Fonctionnalités textométriques : Proposition de typologie selon un point de vue utilisateur*, in S. Bolasco, I. Chiari & L. Giuliano, *Proceedings of 10th International Conference JADT 2010 – Sapienza University of Rome*, LED Edizioni Universitarie, Milano, vol. 2, 341-354.
- Poudat Céline & Landragin Frédéric (2017), *Explorer des données textuelles : méthodes – pratiques – outils*, De Boeck supérieur, Paris.
- Rabatel Alain (2012), *Positions, positionnements et postures de l'énonciateur*, « TRANEL, Travaux Neuchâtelois de Linguistique », 56, 23-42.
- Rabatel Alain (2017), *Pour une lecture linguistique et critique des médias : empathie, éthique, point(s) de vue*, Lambert-Lucas, Limoges.
- Rabatel Alain (2021), *La confrontation des points de vue dans la dynamique figurale des discours : Énonciation et interprétation*, Lambert-Lucas, Limoges.
- Rastier François (2001), *Arts et sciences du texte*, Presses Universitaires de France, Paris.
- Rastier François (2011), *La mesure et le grain : sémantique de corpus*, Honoré Champion, Paris.
- Ratinaud Pierre (2014), *IRAMUTEQ : Interface de R pour les Analyses Multidimensionnelles de Textes et de Questionnaires*. [Logiciel] <http://www.iramuteq.org>.
- Ratinaud Pierre (2018), *Amélioration de la précision et de la vitesse de l'algorithme de classification de la méthode Reinert dans IRaMuTeQ*, in D.F. Iezzi, L. Celardo, & M. Misuraca (éds), *JADT 2018, Proceedings of the 14th international conference on statistical analysis of textual data*, Universitalia, Rome, vol. 2, 616-625.
- Reinert Max (1990), *Alceste une méthodologie d'analyse des données textuelles et une application : Aurelia De Gerard De Nerval*, « Bulletin of Sociological Methodology/ Bulletin de Méthodologie Sociologique », 26(1), 24-54.
- Reinert Max (1993), *Les "mondes lexicaux" et leur "logique" à travers l'analyse statistique d'un corpus de récits de cauchemars*, « Langage et société », 66(1), 5-39.
- Sarfati Georges-Elia (2008), *Pragmatique linguistique et normativité : Remarques sur les modalités discursives du sens commun*, « Langages », 170(2), 92-108.
- Sarrazin Sophie (2017), *Référenciation*, in C. Détrie, P. Siblot, B. Verine & A. Steuckardt (éds), *Termes et concepts pour l'analyse du discours. Une approche pragmatique*, Honoré Champion éditeur, Paris, 330-331.

- Siblot Paul (2001), *De la dénomination à la nomination : Les dynamiques de la signification nominale et le propre du nom*, « Cahiers de praxématique », 36, 189-214.
- Sinclair John (1991), *Corpus, Concordance, Collocation*, Oxford University Press, Oxford.
- Sinclair John (1996), *Preliminary Recommendations on Corpus Typology*, in *EAGLES Document EAG-TCWG-CTYP/P*, <http://www.ilc.cnr.it/EAGLES96/corpus typ/corpus typ.html>.
- Sitri Frédérique (1996), *Interdiscours et construction de l'objet de discours*, « Linx », 8, 153-172.
- Tognini-Bonelli Elena (2001), *Corpus Linguistics at Work*, John Benjamins Publishing Company, Amsterdam-Philadelphia.
- Tournier Maurice (1969), *Le centre de recherche de lexicologie politique de l'E.N.S. de Saint-Cloud*, « Langue française », 2, 82-86.
- Veniard Marie (2013), *La nomination des événements dans la presse : essai de sémantique discursive*, Presses universitaires de Franche-Comté, Besançon.
- Vion Robert (2011), *Polyphonie énonciative et dialogisme*, in J. Bres, A. Nowakowska, J.-M. Sarale & S. Sarrazin (dir.), *Actes du colloque international Dialogisme : langue, discours* (8-10 septembre 2010, Montpellier), Laboratoire Praxiling, Montpellier, <http://www.praxiling.fr/dialogisme-langue-discours.html>.
- Williams Geoffrey (2006), *La linguistique et le corpus : Une affaire prépositionnelle*, « Texto, revue de linguistique en ligne », <http://www.revue-texto.net/Parutions/Livres-E/Albi-2006/Williams.pdf>.
- Zappavigna Michele (2011), *Ambient Affiliation: A Linguistic Perspective on Twitter*, « New Media & Society », 13(5), 788-806.
- Zappavigna Michele & Martin James R. (2018), *#Communing affiliation: Social tagging as a resource for aligning around values in social media*, « Discourse, Context & Media », 22, 4-12.

Graphiques

Deux logiciels ont été utilisés pour effectuer les analyses outillées. Nous précisons ci-dessous les graphiques que nous avons réalisés en fonction de leur usage respectif :

- Graphiques issus de Iramuteq (Ratinaud 2014) : figures n° 6, 7, 8, 9, 10, 11, 12, 21, 23, 25, 27 et 29.
- Graphiques issus de TXM (Heiden *et al.* 2010) : figures n° 13, 15, 16, 18, 19, 20, 24, 28, 30, 31, 32, 33, 34, 35, 36, 37, 38 et 39.

Remerciements

Cet ouvrage aborde des questions théoriques et méthodologiques qui peuvent se poser dans le cadre d'une recherche discursive où le corpus joue un rôle de premier plan, tant dans la création des conditions d'accès à l'objet d'étude, que dans les apports épistémologiques qui découlent de son exploration. Issu d'une partie de ma thèse de doctorat, il se nourrit également des réflexions, des discussions et des expériences qui ont marqué mon parcours post-doctoral.

Je tiens ainsi à remercier toutes les personnes avec lesquelles j'ai eu l'opportunité d'échanger sur ce projet et/ou de collaborer tout au long de ce parcours. En particulier, j'exprime tout d'abord ma reconnaissance à mes directeurs de thèse, Chiara Preite et Julien Longhi, qui m'ont encadrée et accompagnée durant mon doctorat ainsi que par la suite. Je les remercie pour leur disponibilité constante à discuter de nouvelles questions qui m'interrogent, et pour tous les conseils qu'ils m'ont dispensés au fil des ans. Je remercie en outre Julien Longhi de m'avoir fait l'honneur de préfacer ce livre.

J'adresse mes remerciements à Chiara Molinari, qui encadre mes recherches actuelles, pour les discussions autour de mes travaux en cours, pour ses conseils et pour les projets dans lesquels elle m'a engagée. Je tiens également à remercier Stefano Vicari, qui a eu la bienveillance de lire mon travail, pour sa disponibilité et ses encouragements.

Je voudrais exprimer ma gratitude à Cécile Desoutter qui, après m'avoir guidée au début de mes études, a renouvelé son soutien lors de la lecture d'une première version de ce travail. J'adresse également mes remerciements aux relecteurs anonymes de cette même version-là pour leurs remarques et suggestions.

Je souhaite en outre remercier Nora Gattiglia pour nos échanges animés concernant l'analyse du discours, ainsi que pour les activités que nous avons eu le plaisir d'organiser sur ce thème. Mes remerciements s'adressent aussi à Silvia Modena pour les diverses occasions de collaboration et de discussion au cours de mes parcours doctoral et post-doctoral.

Mes remerciements vont enfin à toutes les personnes qui m'ont soutenue et encouragée au cours de mes activités de jeune chercheuse.

Index des notions

Des notions parmi les plus mobilisées le long du texte sont listées ici en guise de support à la lecture. Précisons par ailleurs que des variantes ou des notions en relation avec la notion indexée sont parfois regroupées sous une même entrée. Des variantes graphiques ont également été prises en compte.

- Analyse déductive-contextualisée, 36, 103, 116, 133, 137, 168, 202, 203.
- Analyse du discours (française)/ADF, 13, 14, 20, 21, 24, 27, 31, 32, 40, 42, 45, 46, 113, 204, 205.
- Analyse factorielle des correspondances/AFC, 27, 98, 99, 103, 104, 109, 110, 112, 134, 137, 138, 144, 154, 155.
- Analyse inductive-quantifiée, 36, 82, 90, 95, 126, 147, 157, 167, 168, 185, 192, 197, 199, 202, 203.
- Analyse informatisée, méthode/exploration informatisée, 12, 19, 20, 22, 24, 74, 82, 119.
- Approche multidimensionnelle combinée (analyse - ; démarche -), 13, 15, 25, 42, 90, 196, 202, 205.
- Circulation interdiscursive, 26, 40, 46, 48.
- Classification hiérarchique descendante (CHD), 96, 97, 98, 99, 100, 101, 102, 104, 105, 106, 107, 108, 144, 145, 147, 150, 151, 152, 153, 154, 169, 179.
- Comparabilité/comparaison intergénérique, 51, 76, 78, 84, 85, 149, 168, 186.
- Comparaison intragénérique, 86, 186, 191, 196.
- Concordances/Analyse des concordances, 118, 119, 120, 125, 130, 131, 155, 182.
- Cooccurrences/Analyse des cooccurrences, 118, 120, 124, 127, 132, 155, 182.
- Corpus pluri-générique, 41, 168.
- Dialogisme, 32, 62, 72, 157.
- Dimension interdiscursive, 43, 88, 119, 167, 192.
- Discours autonymiques/discours non autonymiques, 63, 66, 67, 112, 113, 114, 115, 116, 135, 138, 147, 200.
- Discours d'information/articles d'information, 34, 47, 56, 58, 59, 60, 61, 66, 68, 103, 110, 111, 113, 114, 115, 116, 117, 120, 132, 134, 136, 137, 138, 140, 141, 142, 143, 144, 145, 146, 147, 148, 151, 172, 179, 195, 196, 197, 199, 200.

- Discours d'opinion/articles d'opinion, 34, 47, 58, 59, 60, 61, 66, 84, 101, 103, 106, 110, 111, 114, 115, 117, 136, 137, 138, 139, 140, 142, 143, 144, 145, 146, 147, 151, 195, 196, 199.
- Discours numériques, 14, 26, 29, 31, 32, 78.
- Énonciation/acte d'énonciation/instance d'énonciation, 12, 40, 56, 57, 58, 59, 60, 61, 66, 76, 100, 106, 112, 113, 115, 116, 118, 119, 135, 192, 193.
- Focalisation, 129, 134, 143, 197.
- Genre de discours/genre discursif, 11, 12, 13, 14, 21, 29, 31, 32, 36, 40, 41, 42, 43, 44, 45, 46, 48, 50, 51, 53, 54, 55, 64, 66, 84, 90, 186, 191, 196, 204, 205.
- Hétérogénéité comparable, 13, 48, 50, 78, 84, 202, 204, 205.
- Hétérogénéité interne et hétérogénéité externe, 44, 48, 50, 64, 76.
- Hétérogénéité intragénérique/intergénérique, 84, 85, 117, 190, 195.
- Hétérogénéité métaénonciatrice, 64, 95, 103, 112, 114, 116, 134, 136, 137, 138, 148, 196, 197, 202.
- Hétérogénéité sous-générique, 15, 65, 85, 86, 95, 103, 104, 112, 134, 147, 148, 202.
- Hétérogénéité sous-technogénérique, 150, 168, 203.
- Hétérogénéité technodiscursive, 76, 190.
- Informatique/exploration informatique/outils informatiques, 14, 22, 24, 25, 29, 81, 98.
- Interdiscours, 32, 42, 191.
- Liens interdiscursifs, 43, 45, 47, 50.
- Logométrie/logométrique, 13, 14, 24, 27.
- Macro-genre/macro-générique, 13, 33, 34, 42, 53, 56, 59, 70, 76, 151, 195, 202, 204, 205.
- Métadonnées, 26, 28, 50, 61, 65, 66, 79, 80, 81, 82, 84, 85, 86, 88, 89, 90, 95, 105, 113, 138, 201.
- Métaénonciateur, 58, 64, 111, 117, 132, 137.
- Modalisation autonymique, 63, 108, 113.
- Modes autonymiques (de représentation)/Formes autonymiques de représentation, 62, 63, 112, 113, 114, 126.
- Monde lexical, 96, 99, 101, 109, 151, 152, 160, 161, 163, 164.
- Niveau énonciatif-pragmatique, 102, 104, 109, 147, 164, 185, 186, 189, 191, 192, 198, 199, 202, 204.
- Niveau global du corpus, 15, 26, 36, 64, 83, 90, 91, 95, 96, 126, 133, 137, 144, 147, 148, 150, 167, 185, 197, 202, 203.
- Niveau lexico-sémantique, 104, 109, 154, 167, 186, 196, 199.
- Niveau local du corpus, 26, 27, 36, 83, 90, 91, 95, 116, 118, 126, 133, 137, 148, 168, 185, 199, 203.
- Objectivisé vs subjectivisé, 59, 60, 66, 101, 102, 106, 111, 112, 116, 133, 143, 147, 148, 199.
- Objet de discours, 12, 13, 42, 51, 57, 90, 116, 117, 118, 119, 123, 125, 129, 134, 136, 140, 148, 150, 159, 168, 169, 175, 183, 186, 187, 189, 190, 191, 196, 197, 198, 202.
- Objet sémiotique d'augmentation (OSA), 72, 73, 76, 81, 84, 85, 86, 87, 88, 150.
- Perspectivation, 56, 57, 63, 125, 128, 133, 134, 136, 141, 143, 144, 147, 196, 197, 200.
- Pluritechnosémiotique/pluritechnosémiotique, 26, 31, 70, 72, 73, 74, 76, 86, 191.
- Polarisation, 13, 42, 45, 57, 58, 59, 62, 66, 96, 100, 101, 104, 109, 116, 117, 118, 122, 123, 125, 127, 131, 134, 137, 141, 143, 148, 179, 184, 185, 186, 191, 193, 195, 196, 198, 204.
- Positionnement discursif, 12, 13, 41, 60, 117, 118, 119, 120, 132, 134, 136, 138, 143, 146, 147, 148, 166, 167, 169, 179, 186, 187, 191, 192, 193, 195, 196, 197, 198, 199, 201, 202, 203, 204, 205.
- Pratiques technodiscursives, 33, 34, 50, 70, 75, 81, 85, 152, 153, 154, 156, 164, 165, 167, 168, 170, 179, 186, 188, 190, 191, 193, 195.

- Référenciation/(choix de - ; forme de - ; procédé de -), 12, 41, 57, 61, 90, 116, 117, 118, 119, 120, 121, 123, 125, 129, 133, 134, 135, 136, 145, 146, 169, 173, 175, 177, 179, 184, 186, 187, 188, 191, 193, 196, 198, 200.
- Relation interdiscursive, 31, 47, 49, 50, 158, 166.
- Réseau interdiscursif, 42, 65, 71, 158, 166.
- Sous-genre/sous-générique, 13, 14, 15, 33, 34, 41, 42, 58, 59, 61, 62, 65, 66, 77, 78, 90, 100, 101, 103, 105, 109, 110, 113, 117, 136, 138, 145, 146, 147, 148, 155, 168, 186, 190, 195, 196, 202, 203, 204.
- Spécificités (analyse/calcul des-), 27, 96, 97, 101, 102, 110, 111, 112, 115, 117, 119, 120, 123, 126, 127, 129, 146, 158, 161, 162, 163, 164, 165, 166, 169, 170, 173, 180, 181, 183.
- Spécificités génériques, spécificités intra- et sous-génériques, spécificités textuelles, spécificités technodiscursives, 12, 13, 14, 15, 29, 30, 34, 36, 40, 42, 45, 48, 49, 50, 51, 53, 67, 68, 69, 78, 79, 80, 84, 88, 92, 150, 151, 153, 168, 185, 186, 191, 192, 193, 195, 197, 198, 201, 202, 204, 205.
- Technodiscours, 30, 31, 33, 34, 55, 67, 69, 72, 73, 74, 75, 77, 78, 79, 80, 81, 82, 84, 85, 86, 88, 149, 157, 165, 170, 174, 191, 192, 193, 197, 198, 199.
- Technogenre/(sous-)technogénérique, 31, 33, 54, 55, 78, 87, 150, 168, 185, 191, 193, 197, 198, 199, 203.
- Textométrie/textométrique, 12, 13, 14, 24, 26, 27, 28, 35, 41, 74, 82, 89, 149, 202.
- Textualisation, 33, 40, 54, 150, 169, 204.
- Textualité numérique, 30, 31, 32, 36, 37, 49, 54, 55, 75, 78.
- Textualité numérikée, 30, 49, 54.
- Textualité numérisée, 30, 49.
- Tweet augmenté, 72, 73, 74, 76, 81, 85, 86, 87, 88, 149.
- Tweet simple, 72, 73, 74, 75, 76, 78, 81, 84, 85, 86, 88, 149, 150, 156, 158, 168, 169, 185, 191, 198, 199.
- Variation intergénérique, 11, 35, 40, 41, 43, 41, 48, 92, 186, 204, 205.
- Variation intragénérique, 11, 15, 43, 76, 78, 79, 86, 92, 104, 111, 112, 147, 154, 164, 168, 186, 198.

Index des noms

Cet index répertorie les noms des auteures et auteurs des travaux scientifiques et des logiciels cités dans le texte.

- Achard-Bayle, Guy, 54, 65.
Adam, Jean-Michel, 29, 31, 59, 66, 91, 102.
Amossy, Ruth, 47, 62, 129, 134, 148, 157, 189, 200.
Angermuller, Johannes, 19, 20, 143.
Authier-Revuz, Jacqueline, 58, 61, 62, 63, 64, 108, 112, 113.
- Badouard, Romain, 45, 46, 160.
Bakhtine, Mikhaïl, 32, 62.
Beacco, Jean-Claude, 19.
Benveniste, Émile, 96, 100, 101.
Benzécri, Jean-Paul, 27, 98, 103.
Biber, Douglas, 20, 23.
Bibié, Laetitia, 77.
Bigey, Magali, 71, 72, 150.
Bonnafous, Simone, 27.
Bres, Jacques, 32, 42, 61, 62, 72, 157.
Brunet, Étienne, 24.
Bruns, Axel, 47.
- Charaudeau, Patrick, 20, 21, 23, 32, 41, 46, 48, 50, 56, 57, 59, 60, 66, 99, 113, 116, 138, 163, 204.
Condamines, Anne, 19.
- Dalbera, Jean-Philippe, 23.
Demonet, Michel, 24.
Doury, Marianne, 192.
Dubois, Danielle, 119.
Ducrot, Oswald, 61, 156, 177, 192.
- Emerit, Laetitia, 78, 79, 83.
- Garric, Nathalie, 32, 36, 37, 41, 42, 43, 44, 48, 60, 89, 203.
Gibson, James J., 69.
Graumann, Carl F., 57.
Guilhaumou, Jacques, 21, 24.
- Habert, Benoît, 19, 20.
Hardie, Andrew, 20.
Heiden, Serge, 110, 121, 155, 174.
- Jackiewicz, Agata, 71, 158.
Julliard, Virginie, 71, 86, 157, 158.
- Kallmeyer, Werner, 57.
Krieg-Planque, Alice, 160.
Kwak, Haewoon, 47.

- Lafon, Pierre, 27, 101.
 Landragin, Frédéric, 20, 22, 28, 36, 85.
 Lebart, Ludovic, 20, 24, 28.
 Leblanc, Jean-Marc, 98, 101.
 Longhi, Julien, 21, 22, 23, 25, 26, 29, 32, 35, 36,
 37, 40, 41, 42, 48, 60, 77, 80, 81, 82, 83, 89,
 90, 116, 117, 119, 149, 154, 163, 167, 193, 201.
 Magué, Jean-Philippe, 74.
 Maingueneau, Dominique, 20, 21, 31, 32, 33,
 40, 41, 45, 48, 50, 54, 57, 58, 62, 64, 113,
 116, 117, 143, 204.
 Marchand, Pascal, 97.
 Martin, James R., 71.
 Mayaffre, Damon, 20, 23, 24, 25, 26, 27, 35,
 36, 47, 65, 88, 89, 91, 118, 203.
 Mazière, Francine, 20, 21, 204.
 McEnery, Tony, 20.
 Mellet, Sylvie, 23.
 Mercier, Arnaud, 47, 68, 71, 158.
 Meunier, Jean-Guy, 25, 36, 90.
 Micheli, Raphaël, 188.
 Moirand, Sophie, 20, 21, 45, 50, 56, 57, 59,
 60, 62, 65, 66, 101, 102, 111, 112, 148.
 Moïse, Claudine, 182.
 Mondada, Lorenza, 119.
 Monnier, Angeliki, 190.

 Née, Émilie, 28, 98.

 Oger, Claire, 46, 48.
 Ollivier-Yaniv, Caroline, 46, 48.

 Paveau, Marie-Anne, 30, 31, 33, 42, 49, 54,
 55, 65, 67, 69, 71, 73, 78, 79, 80, 89, 149,
 158, 168.
 Pêcheux, Michel, 22, 24, 32, 42.
 Pešek, Ondřej, 54, 65.
 Pignard-Cheynel Nathalie, 68.
 Pincemin, Bénédicte, 22, 24, 27, 28, 35, 43,
 44, 79, 91, 96, 101, 103.
 Poudat, Céline, 20, 22, 28, 36, 85.

 Rabatel, Alain, 56, 57, 58, 60, 61, 116, 118.
 Rastier, François, 19, 21, 22, 26, 27, 32, 34, 35,
 39, 40, 64, 65, 67, 89, 91, 203.

 Ratinaud, Pierre, 96, 97, 109, 150.
 Reinert, Max, 96, 151.

 Salem, André, 20, 24.
 Sarfati, Georges-Elia, 48.
 Sarrazin, Sophie, 118.
 Siblot, Paul, 62.
 Simon, Justine, 71, 72, 150.
 Sinclair, John, 20, 23.
 Sitri, Frédérique, 116.

 Taglioli, Massimiliano, 80, 81.
 Tognini-Bonelli, Elena, 20, 43.
 Tournier, Maurice, 24, 27.

 Veniard, Marie, 57, 62.
 Vidak, Marko, 71, 158.
 Vion, Robert, 160.

 Williams, Geoffrey, 20.

 Zappavigna, Michele, 71, 158.

Scienze dell'antichità, filologico-letterarie e storico-artistiche

dello stesso argomento nel catalogo tab

Marie Veniard, *La nominazione degli eventi nella stampa. Saggio di semantica discorsiva*, 978-88-9295-218-8 (ISBN edizione digitale 978-88-9295-420-5)

Marie-Anne Paveau, *Prediscorsi. Senso, memoria, cognizione*, 978-88-9295-569-1 (ISBN edizione digitale 978-88-9295-570-7)

Julien Longhi, *Dal discorso come campo al corpus come terreno. Contributo metodologico all'analisi semantica del discorso*, 978-88-9295-948-4 (ISBN edizione digitale 978-88-9295-949-1)



Corpus discursifs, hétérogénéité des données, hétérogénéité des méthodes.

Une approche méthodologique des variations intra- et intergénériques

di Claudia Cagninelli

prefazione di Julien Longhi

direttore editoriale: Mario Scagnetti

editor: Laura Moudarres

caporedattore: Giuliano Ferrara

redazione: Giulia Ferri

progetto grafico: Sara Pilloni

