



HAL
open science

Terminologie de l'endométriose et représentations de la maladie : regards croisés entre presse généraliste et discours spécialisés

Julie Humbert-Droz

► To cite this version:

Julie Humbert-Droz. Terminologie de l'endométriose et représentations de la maladie : regards croisés entre presse généraliste et discours spécialisés. 9e Congrès Mondial de Linguistique Française, Jul 2024, Lausanne, Suisse. pp.05003, 10.1051/shsconf/202419105003 . hal-04650125

HAL Id: hal-04650125

<https://hal.science/hal-04650125v1>

Submitted on 16 Jul 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Terminologie de l'endométriose et représentations de la maladie : regards croisés entre presse généraliste et discours spécialisés

Julie Humbert-Droz

Centre de recherche en linguistique appliquée (CeRLA) - Université Lumière Lyon 2

julie.humbert-droz@univ-lyon2.fr

Résumé. Cet article propose d'alimenter la réflexion sur la notion de circulation des termes, dans le contexte particulier de l'endométriose. Cette réflexion s'articule autour de l'exploration outillée de deux corpus comparables, l'un composé d'articles de presse généraliste française et l'autre de textes de spécialité. Différents points de variation entre les deux corpus sont mis au jour à l'aide de la combinaison d'indices quantitatifs et qualitatifs et permettent d'identifier et de discuter les représentations de l'endométriose dans la presse, en comparaison avec les textes spécialisés. Les analyses menées sur les corpus montrent ainsi que, dans la presse, ces représentations incluent systématiquement des questions comme les douleurs liées à l'endométriose, les lacunes dans les connaissances actuelles sur cette maladie ou encore l'association entre endométriose et revendications féministes. Dans la presse, ces représentations reflètent également les variations qui existent au sein de la communauté des spécialistes, à propos notamment de la définition même d'*endométriose*. Ces observations participent à nourrir la réflexion sur les différentes voix représentées dans la presse et sur leur rôle dans la circulation des termes.

1 Introduction et objectifs

La circulation des termes médicaux en dehors des sphères spécialisées soulève d'importantes questions relatives à leur réception et leur appropriation par les non-spécialistes et les patients (King et Watson 2005, Delavigne et al. 2022). La nécessité de rendre les connaissances médicales accessibles pour le grand public a été démontrée dans des études portant, entre autres, sur les usages et les reformulations des termes dans différents types de textes (médias de vulgarisation, forums, médias généralistes, etc.) (Delavigne 2019, Rouillard 2016). Parallèlement, les médias contribuent à informer le grand public sur les questions médicales et participent, dans le même temps, à véhiculer certaines représentations qui s'éloignent parfois de celles véhiculées par les spécialistes (Delforge 2016, Dempster et al. 2022, Moirand et Reboul-Touré 2023). Ces représentations, lorsqu'elles sont erronées ou lorsqu'elles simplifient dans une trop large mesure les connaissances des spécialistes, peuvent porter préjudice aux personnes atteintes, en les stigmatisant ou en les discréditant auprès du grand public (Costes et Laval 2019, Balfour 2023). Ce constat s'applique également à l'endométriose, pour laquelle certaines idées reçues perdurent et peuvent avoir un impact négatif sur sa compréhension par le grand public (Young et al. 2015, Galès 2020), malgré les actions de sensibilisation menées par les associations de patientes (EndoFrance 2023).

Dans ce contexte, nous proposons d'explorer les représentations de l'endométriose dans un corpus de presse française, en comparaison avec un corpus de textes spécialisés. Tirant profit des acquis de la terminologie et des travaux sur la circulation des termes (Delavigne 2001, Gaudin 2003, Humbert-Droz 2021, Condamines et Picton 2022), nous cherchons plus précisément à appréhender les différences, lexicales et terminologiques, entre les deux corpus. Nous nous focalisons pour ce faire sur plusieurs termes et unités lexicales, que nous sélectionnons dans un premier temps à partir d'informations quantitatives et, dans un second temps, pour leur centralité. L'exploration du fonctionnement de ces unités dans les corpus permettra de saisir les représentations de l'endométriose dans la presse et d'identifier dans quelle mesure elles divergent de celles de spécialistes, ou convergent avec celles-ci. Cette étude s'inscrit en outre dans un projet plus vaste, qui s'intéresse à la circulation des termes liés à l'endométriose entre spécialistes, patientes et non-spécialistes, ainsi qu'à la réception et à l'appropriation des termes par les femmes qui en sont atteintes et par le grand public¹. L'identification des représentations de la maladie dans un corpus de presse constitue

ainsi la première phase de ce projet. À terme, les résultats serviront à mieux comprendre comment l'usage des termes varie entre ces communautés et à mieux cerner les conséquences de ces variations sur la réception et l'appropriation de ces termes par les non-spécialistes.

Cet article propose ainsi, dans la section 2, de poser les bases théoriques sous-tendant notre étude, en définissant d'abord la notion de circulation des termes et en explicitant certaines spécificités liées au domaine médical, puis en donnant quelques points de repères à propos de l'endométriose et des enjeux de compréhension de cette maladie par le grand public, que nous mettons en lien avec la circulation des termes. Dans la section 3, nous explicitons l'approche méthodologique adoptée ici et décrivons d'abord les deux corpus sur lesquels repose notre analyse puis les différents indices que nous explorons. Nos observations sont présentées et discutées dans la section 4, où nous passons en revue certains indicateurs quantitatifs avant de nous focaliser, d'une part sur les termes désignant les différents types d'endométriose attestés dans les deux corpus et, d'autre part, sur les contextes définitoires du terme *endométriose*. Enfin, nous synthétisons nos observations en conclusion et proposons quelques pistes pour nourrir la réflexion sur les dynamiques de circulation des termes dans le contexte de l'endométriose.

2 Contexte théorique

2.1 Circulation des termes

Notre étude porte de manière générale sur la notion de circulation des termes. Cette notion renvoie au constat que les termes ne sont pas employés par des spécialistes uniquement, dans des situations de communication spécialisées uniquement. Au contraire, les termes sont susceptibles d'apparaître dans de multiples situations, caractérisées par un degré de spécialisation plus ou moins élevés. La notion de circulation des termes permet alors de rendre compte non seulement de la richesse des usages réels des termes, mais également de la diversité des locuteurs qui les emploient. C'est essentiellement à travers le prisme de la socioterminologie que cette notion s'est affirmée, dans des travaux consacrés à la description des fonctionnements des termes en dehors des discours des spécialistes, particulièrement dans le cadre de la vulgarisation scientifique (p. ex. Gaudin 1993, 2003, Delavigne 2001). La grande variabilité des termes (dénominations employées, sens véhiculés et construits, variations entre communautés, évolution dans le temps, etc.) est au cœur de ces travaux (Delavigne et Gaudin 2022). La notion de circulation des termes est alors mobilisée pour interroger des problématiques liées, entre autres, à la réception et à la compréhension des termes par les non-spécialistes, à la négociation du sens des termes par les non-spécialistes, aux dynamiques de changement sémantique à l'œuvre lorsque les termes s'émancipent des discours spécialisés (Gaudin 1993, Ungureanu 2006, Humbert-Droz 2021, Cappeau et Schnedecker 2022, Delavigne et al. 2022).

Parallèlement, il est reconnu que les médias contribuent largement à la circulation des termes (Meyer 2000, Beacco et al. 2002, Gaudin 2003). Nombre de travaux s'intéressent ainsi au traitement, dans la presse, de certains sujets scientifiques, dans la perspective de la terminologie (p. ex. Delavigne 2001, Condamines et Picton 2014), de l'analyse de discours (p. ex. Beacco et al. 2002, Moirand et Reboul-Touré 2023) ou des études journalistiques (p. ex. Costes et Laval 2019, Dempster et al. 2022). Deux éléments en particulier sont à retenir : d'une part, la couverture très large dans les médias de certains événements scientifiques entraîne la diffusion également large de termes liés à ces événements (Gaudin 1993, Pearson 1998, Humbert-Droz 2021) et, d'autre part, cette diffusion large de termes dans les médias peut donner lieu à des glissements sémantiques, tels que de nouvelles connotations qui apparaissent dans les usages des termes ou des changements plus durables (Moirand 2007, Le Draoulec et al. 2014).

2.2 Quelques particularités de la circulation des termes dans le contexte médical

Dans le contexte de la médecine, les médias jouent un rôle particulièrement important dans la diffusion d'informations médicales auprès d'un public non spécialiste (Delforge 2016, Moirand et Reboul-Touré 2023). Plusieurs auteurs affirment même que les médias deviennent, ou sont devenus, la principale source

d'information pour les non-spécialistes (King et Watson 2005, Dempster et al. 2022) Cependant, étant donné que l'usage des termes dans les médias peut entraîner des conséquences sur leur sens, la nature et l'ampleur de ces variations sémantiques peuvent également avoir un impact sur la réception des termes par les non-spécialistes (Rouillard 2016, Delavigne 2019). Aussi, vu que « le traitement médiatique d'un sujet est particulièrement révélateur de nos représentations sociales » (Costes et Laval 2019 : 32), ces variations participent à véhiculer certaines représentations qui s'éloignent des usages des termes par les spécialistes, avec pour conséquence la diffusion de stéréotypes sur certaines maladies, voire la stigmatisation des malades, par exemple dans le cas de la schizophrénie (Balfour 2023).

De la même manière, la diffusion, dans les médias, de certaines représentations erronées ou éloignées de la réalité peut avoir pour conséquence l'installation d'un certain manque de confiance envers la médecine et les médecins (Dempster et al. 2022). Cette perte de confiance peut parfois entraîner des modifications de comportement, par exemple avec l'arrêt spontané d'un médicament suite à la forte médiatisation d'incidents liés à ce médicament (Delforge 2016). L'impact de la médiatisation de la médecine et de la circulation des termes médicaux dans les médias va donc bien au-delà des questions de compréhension des notions et de réception des termes par les non-spécialistes. Dans cet article, cependant, nous restreignons notre propos aux différentes représentations d'une maladie dans les médias, l'endométriose, particulièrement lorsque ces représentations diffèrent de celles véhiculées par les spécialistes.

2.3 L'endométriose, une maladie complexe

L'endométriose est une maladie gynécologique inflammatoire chronique. Elle se définit par « la présence en dehors de la cavité utérine de tissu semblable à la muqueuse utérine » qui subit « l'influence des modifications hormonales » au cours de chaque cycle menstruel (EndoFrance 2023). Cette maladie touche une à deux femmes sur dix, provoque de nombreux symptômes douloureux et reste incurable à ce jour (Young et al. 2015, Ilschner et al. 2022). Elle est qualifiée de multifactorielle et de complexe, en raison des différentes théories élaborées pour expliquer ses causes, dont aucune ne satisfait complètement les chercheurs (Galès 2020), de la combinaison de facteurs génétiques, environnementaux et hormonaux qui la font évoluer et de la multiplicité des formes qu'elle peut prendre (Galès 2020, EndoFrance 2023). Parallèlement, la méconnaissance de l'endométriose, tant chez les professionnels de santé que chez le grand public (Galès 2020), entraîne un retard diagnostique important, entre 7 et 10 ans (Young et al. 2015), ce qui retarde d'autant la mise en place de soins adaptés (EndoFrance 2023).

La nature multifactorielle de l'endométriose et l'étendue possible des lésions à plusieurs organes exigent la collaboration de plusieurs disciplines pour diagnostiquer la maladie et prendre en charge les patientes. De même, la recherche sur l'endométriose s'effectue au sein de plusieurs disciplines médicales, et non uniquement en gynécologie et en chirurgie. Cette forte interdisciplinarité, tant sur le plan des pratiques que sur le plan de la recherche, entraîne une multiplication des points de vue sur la maladie, qui peuvent par ailleurs se traduire par certaines variations (sémantiques, lexicales) dans les données textuelles. C'est précisément ce qui rend une analyse linguistique et terminologique pertinente, selon nous, afin de saisir et de mieux comprendre ces variations et leurs enjeux du point de vue de la circulation des termes.

3 Aspects méthodologiques : une approche outillée en corpus comparable

Notre démarche d'analyse repose sur l'exploration de deux corpus, appuyée par différents outils. Nous adoptons une approche contrastive : les corpus sont comparés afin de faire émerger des phénomènes pertinents pour notre objectif. Nous cherchons ainsi à mettre au jour des différences de fonctionnements des termes et d'autres unités lexicales, qui permettent d'appréhender les représentations de l'endométriose véhiculées dans la presse. Le recours systématique à des outils d'analyse de corpus relève en outre d'une volonté de mener une analyse objective et reproductible ; notre propre subjectivité d'analyste est ainsi « cadrée » par les résultats fournis par les outils (Habert 2004, Jacques 2016, Humbert-Droz 2021, Condamines et Picton 2022). Notre approche s'articule autour de l'observation de plusieurs indices

quantitatifs et qualitatifs repérables en corpus. Leur définition et leur repérage, ainsi que les outils utilisés, font l'objet de la section 3.2 ; la section 3.1 détaille la constitution des corpus.

3.1 Corpus exploités dans cette étude

Les deux corpus comparables (p. ex. Sinclair 1994) sur lesquels nous basons notre étude sont composés d'articles de presse (désormais *EndoPresse*) et de textes spécialisés (désormais *EndoExp*). Ils visent à refléter, d'une part, les discours de la presse écrite française autour de l'endométriose et, d'autre part, les discours des professionnels de santé. La constitution de ces deux corpus repose sur des principes d'échantillonnage² développés dans Humbert-Droz (2021). Ils s'appuient sur des critères de sélection des textes précis qui tiennent compte de l'objectif de l'étude (Condamines et Picton 2022). Par exemple, le choix d'inclure des articles de presse généraliste repose à la fois sur le constat du rôle crucial des médias dans la circulation des termes et sur la place toujours plus grande qu'occupent les médias dans la recherche d'informations médicales par les locuteurs non spécialistes, comme nous l'expliquons *supra*. En outre, les articles ont été échantillonnés dans de la presse aussi bien nationale que régionale, dans des journaux à parution aussi bien hebdomadaire que quotidienne, et sur différents supports (en ligne avec équivalent papier ou en ligne uniquement) afin de représenter aussi largement que possible la presse française. Précisons également que chaque article a été sélectionné individuellement sur la base de sa pertinence vis-à-vis de l'endométriose. Ce corpus compte 528 798 occurrences (voir Tableau 1).

Pour le corpus *EndoExp*, la complexité de l'endométriose et l'interdisciplinarité de la recherche sur cette maladie, de son diagnostic et de sa prise en charge rendent nécessaire l'inclusion de textes spécialisés émanant de plusieurs disciplines médicales, comme l'imagerie médicale, l'urologie ou la pharmacie, en plus de la gynécologie. Les discours sur cette maladie sont « éclatés » et considérer un seul domaine médical ne semble pas satisfaisant pour refléter leur diversité. Pour reprendre les propos de Delavigne, ce corpus n'est donc pas constitué de textes provenant d'un domaine en particulier, comme c'est le plus souvent le cas en terminologie (p. ex. L'Homme 2020 : 133), mais de « discours *sur* » (Delavigne 2022 : 402) l'endométriose. Par ailleurs, les textes sélectionnés ont été rédigés par différents types de spécialistes (médecins, sages-femmes, chirurgiens, etc.) et relèvent de différents genres (articles de recherche, mémoires de fin d'études, etc.). Ce corpus totalise 1 004 108 occurrences (Tableau 1).

La période couverte par les deux corpus va de 2018 à 2022. Ce choix se fonde sur la publication début 2018 des recommandations de bonne pratique pour la prise en charge de l'endométriose par la Haute Autorité de santé (HAS) et le Collège national des gynécologues et obstétriciens français (CNGOF), dont le texte a été approuvé en décembre 2017. Ce texte remplace les recommandations du CNGOF de 2006, tient compte des dernières avancées de la recherche et promeut un accompagnement pluridisciplinaire des patientes (HAS et CNGOF 2017). Il a vocation à être un texte de référence pour le diagnostic et le traitement de l'endométriose par les gynécologues, les chirurgiens, les médecins généralistes, les sages-femmes et les autres professionnels de santé impliqués. Ces recommandations sont également caractérisées par une prise en compte des besoins des patientes, à travers l'association EndoFrance, qui a participé au groupe de travail à l'origine de leur rédaction³. Leur diffusion se veut large et leur impact ainsi que leur retentissement dans les médias importants. Enfin, en tant que premier texte officiel qui promeut une approche de l'endométriose devant faire collaborer plusieurs disciplines médicales, l'année de sa publication, 2018, est une année charnière. Pour toutes ces raisons, nous l'avons choisie comme point de départ pour notre étude en corpus.

Tableau 1. Constitution et taille des corpus

	Corpus <i>EndoExp</i>	Corpus <i>EndoPresse</i>
Genres textuels	Articles de recherche, chapitres d'ouvrages, thèses, mémoires, résumés	Articles de presse généraliste
Sources des textes	Revue spécialisées (p. ex. <i>Gynécologie Obstétrique Fertilité & Sénologie</i>), universités françaises, archives ouvertes (p. ex. HAL)	Presse française nationale et régionale, quotidienne et hebdomadaire, en ligne ou avec équivalent papier (p. ex. <i>Le Monde, Ouest-France, Le HuffPost, Le Figaro</i>)
Nombre de textes	166 textes	785 articles
Taille	1 004 108 occurrences	528 798 occurrences

3.2 Indices explorés et outils utilisés

L'analyse s'articule autour d'indices basés sur la fréquence et la répartition des termes dans les corpus et sur leur distribution, particulièrement les différences de distribution que l'on peut observer entre les corpus. Dans le premier cas, nous nous situons dans une perspective exploratoire et cherchons à obtenir un aperçu des termes et des autres unités significativement plus fréquentes dans chacun des corpus, selon l'hypothèse que ces unités peuvent être particulièrement révélatrices des représentations de l'endométriose dans les corpus (Balfour 2023 : 57sq.). Nous utilisons pour ce faire le système d'extraction semi-automatique de termes TermStat⁴ (Drouin 2003) dont nous complétons les résultats avec les fonctionnalités de repérage de *keywords* (ou mots-clés) et de recherche en CQL (*Corpus Query Language*) de Sketch Engine (Kilgarriff et al. 2004). La complémentarité de ces deux outils nous permet de nous focaliser non seulement sur les similarités, en observant par exemple les candidats-termes extraits à la fois du corpus *EndoExp* et du corpus *EndoPresse*, mais également sur les différences, en tenant compte des *keywords* identifiés par Sketch Engine dans chaque corpus par rapport à l'autre. À travers le repérage des *keywords*, nous cherchons également à identifier les thèmes les plus spécifiques à chacun des corpus (Scott 1997, Balfour 2023). Enfin, dans cette approche quantitative, la prise en compte systématique des contextes des corpus vise à éviter la surinterprétation dans l'analyse.

Dans le second cas, nous nous focalisons plus particulièrement sur la distribution du terme *endométriose*, que nous explorons à partir de contextes riches en connaissances. Il s'agit de contextes qui indiquent « at least one item of domain knowledge that could be useful for conceptual analysis » (Meyer 2001 : 281). Nous nous intéressons surtout ici aux contextes définitoires, autrement dit aux contextes explicitant une définition, entière ou partielle, pour le terme *endométriose*. Cette étape vise à approfondir l'aperçu obtenu avec le premier indice, en tenant compte des différentes manières de définir cette maladie dans *EndoPresse* et en identifiant les points où ces définitions diffèrent de celles observables dans *EndoExp*. Les contextes définitoires sont identifiés à l'aide de marqueurs⁵, listés à partir de différents travaux existants (notamment Rebeyrolle 2000, Sierra et al. 2022) et de la ressource MAR-REL⁶ (Lefevre 2017). Ils sont recherchés dans les corpus avec Sketch Engine et à l'aide de requêtes CQL, ce qui permet de cibler précisément le type de contextes souhaité. Enfin, parmi les résultats retournés avec cette méthode, nous n'avons pas tenu compte des contextes définitoires portant sur un hyponyme d'*endométriose* (par exemple *endométriose thoracique*) et nous n'avons retenu que les contextes définitoires apparaissant dans les limites d'une phrase.

À partir de ces éléments, nous proposons trois types d'observations : d'abord les résultats obtenus quantitativement avec l'extraction de candidats-termes, puis les différents types d'endométriose évoqués dans les corpus à partir des termes complexes formés sur la base du terme *endométriose* et enfin les principaux résultats issus de l'analyse des contextes définitoires.

4 Résultats et discussion : représentations de l'endométriiose dans la presse

4.1 Aperçu quantitatif global

L'extraction effectuée à l'aide de TermoStat a fourni une liste de 816 et 242 candidats-termes avec un score de spécificité plus grand que ou égal à 30, dans les corpus *EndoExp* et *EndoPresse* respectivement. Après suppression du bruit (entre autres, des doublons dus à des analyses morphosyntaxiques différentes, des erreurs de découpage, des unités en anglais qui ne sont pas des emprunts en français), 582 et 227 candidats-termes respectivement sont conservés, ce qui représente 672 unités différentes (certaines ont été extraites dans les deux corpus à la fois). Cet échantillon est composé majoritairement de noms (490, dont 296 noms simples et 194 syntagmes nominaux), mais également d'adjectifs (154), de verbes (26) et d'adverbes (2).

Précisons à ce stade qu'une unité peut ne pas avoir été extraite par TermoStat dans l'un des corpus et être tout de même attestée dans ce corpus, avec une fréquence généralement plutôt faible, de la même manière qu'une unité peut avoir été extraite mais avec un score de spécificité inférieur à 30, qui l'aurait exclue de la première étape. Pour cette raison, nous avons observé la fréquence de chaque candidat-terme dans les deux corpus et avons identifié les unités qui sont communes aux deux corpus et celles qui n'apparaissent réellement que dans l'un ou l'autre des corpus. C'est ce qu'illustre le Tableau 2, qui détaille la répartition des unités dans les corpus en fonction de leur catégorie grammaticale et indépendamment de leur fréquence. Les pourcentages se réfèrent au nombre total d'unités par catégorie, indiqué dans la dernière colonne.

Tableau 2. Répartition des unités dans les corpus *EndoExp* et *EndoPresse*

	Nombre en commun	Nombre dans <i>EndoExp</i> uniquement	Nombre dans <i>EndoPresse</i> uniquement	Nombre total
Noms	385 (78,6 %)	93 (19 %)	12 (2,4 %)	490
Adjectifs	111 (72,1 %)	41 (26,6 %)	2 (1,3 %)	154
Verbes	24 (92,3 %)	2 (7,7 %)	0	26
Adverbes	2 (100 %)	0	0	2
Total	522 (77,7 %)	136 (20,2 %)	14 (2,1 %)	672

Cette répartition amène une première remarque : bien que l'extraction avec TermoStat ait permis d'identifier 227 candidats-termes dans *EndoPresse*, une grande majorité des candidats-termes repérés dans *EndoExp* est également attestée dans la presse. En effet, 522 des 658 unités attestées dans le corpus *EndoExp* apparaissent également dans *EndoPresse*, soit 79,3 %. Cela signifie ainsi qu'une grande partie de la terminologie employée par les spécialistes dans le corpus *EndoExp* « passe » dans la presse.

Il faut cependant nuancer ce constat, puisque plus d'un quart de ces unités (144/522, soit 27,6 %) apparaissent entre 1 et 5 fois seulement dans le corpus *EndoPresse*. Cette faible fréquence suggère que ces unités ne sont pas véritablement diffusées dans l'usage des journalistes. Un autre élément pointe vers cette interprétation : le plus souvent, ces unités sont employées dans des citations directes ou indirectes, autrement dit dans du discours rapporté. Le choix de ces unités n'est donc pas nécessairement celui des journalistes, mais plutôt de leurs sources. En outre, l'observation du discours rapporté montre que les citations proviennent des spécialistes, comme dans l'exemple a, qui illustre *échographie endovaginale* (en gras dans l'exemple), ou des patientes, comme dans l'exemple b, qui illustre *biopsie* (en gras également).

- a. « *L'échographie endovaginale* consiste à explorer l'utérus et les ovaires par l'intermédiaire d'une sonde échographique placée dans le vagin », décrit le Dr Sauvanet. (corpus *EndoPresse*)

- b. « *J'avais 23 ans et un jour de cycle, j'ai eu de fortes douleurs au ventre [...]. Là, on me fait passer une biopsie et on me dit que j'ai une endométriose [...].* » (corpus *EndoPresse*)

Une dernière remarque peut être faite à partir du Tableau 2. En comparaison avec les autres catégories grammaticales, et particulièrement les verbes et les adverbes, les adjectifs tendent davantage à rester réservés à l'usage des spécialistes, du moins semblent-ils se diffuser dans une moindre mesure dans l'usage des journalistes. L'observation des adjectifs permet par ailleurs de remarquer qu'une grande partie de ceux qui sont attestés uniquement dans *EndoExp* servent à décrire les localisations anatomiques des lésions d'endométriose ou des organes impactés ou encore à détailler certains actes chirurgicaux ou examens radiologiques sur le plan anatomique. C'est le cas d'adjectifs comme *urétéral*, *vésico-utérin*, *pariétal*, *pleural*, *postérieur*, *discoïde*, *latéral*, *sagittal*, *sigmoïdien* et *proximal*. En outre, bien que d'autres adjectifs de ce type soient attestés dans les deux corpus (*abdominal*, *iliaque*, *pelvien*, *vaginal*, etc.), l'usage qui en est fait diffère dans une large mesure entre les corpus. En effet, contrairement à leur usage dans *EndoExp*, dans le corpus *EndoPresse*, ces adjectifs servent le plus souvent à caractériser les localisations des douleurs provoquées par l'endométriose (avec des syntagmes comme *douleur pelvienne*), voire, dans certains cas, à évoquer la douleur associée à un examen effectué pour diagnostiquer l'endométriose (par exemple dans *toucher vaginal douloureux*). Il semble alors que les adjectifs plutôt caractéristiques du corpus *EndoExp* représentent un niveau de description qui ne pénètre pas la presse et l'ensemble de ces observations tend ainsi à montrer que la presse reste plus générique. Ce constat est par ailleurs attendu, puisque les textes rassemblés dans le corpus *EndoExp* sont bien plus spécialisés que les articles de presse. Nous développons cependant davantage cette question dans la section suivante, où nous nous intéressons aux différents hyponymes formés sur la base du terme *endométriose* (voir section 4.2, notamment à partir du Tableau 3).

Parallèlement, d'autres thématiques que les actes chirurgicaux dans *EndoExp* et les localisations des douleurs dans *EndoPresse* émergent de l'observation des autres unités extraites automatiquement, par exemple une thématique liée au cycle menstruel, avec des unités comme *menstruation*, *règles* et *syndrome prémenstruel*, une thématique autour de l'infertilité, avec *problème de fertilité* et *stimulation ovarienne* notamment, ou encore une thématique liée à la sexualité, avec *vie sexuelle* ou *fonction sexuelle*. Ces thématiques sont toutes communes aux deux corpus et ces unités y sont toutes attestées. En revanche, d'autres thématiques ne sont évoquées que dans l'un ou l'autre des corpus : les différents types de lésions n'apparaissent que dans *EndoExp*, avec des unités comme *lésion digestive* ou *lésion d'endométriose profonde*. Deux thématiques sont caractéristiques du corpus *EndoPresse* : le sexisme (avec *sexisme* et *sexiste*) et la violence (avec par exemple *violence gynécologique* ou *violence obstétricale*).

Afin d'approfondir ces questions et de cibler plus précisément les différences lexicales les plus significatives entre les deux corpus, nous avons extrait les *keywords* du corpus *EndoExp* avec Sketch Engine, avec le corpus *EndoPresse* comme corpus de référence, et vice versa. Dans les deux cas, les *keywords* confirment les tendances observées : par exemple, dans *EndoExp*, des termes comme *salpingectomie bilatérale*, *chirurgie robotique* et *résection complète* confirment l'importance des aspects chirurgicaux liés au traitement de l'endométriose. En revanche, dans *EndoPresse*, de nombreux *keywords* peuvent être rapprochés de thématiques plutôt féministes et militantes, notamment *combat*, *féministe*, *militant* ou *patriarcal*. Rappelons à ce stade que cette méthode d'exploration de corpus vise à faire ressortir les unités pour lesquelles il existe une différence de fréquence significative dans un corpus en entier par rapport à un autre. Dans le cas de ces dernières unités, le lien avec l'endométriose n'est pas toujours évident de prime abord. Cependant, étant donné que les articles constitutifs du corpus *EndoPresse* traitent tous, de près ou de loin, d'endométriose (cf. section 3.1), alors il est tout à fait probable que les thèmes évoqués par ces *keywords* le soient précisément pour leur rapport avec l'endométriose. C'est d'ailleurs ce que confirment les données : dans les exemples ci-dessous, le terme *endométriose* apparaît dans des énumérations qui renvoient aux différents thèmes défendus par des personnes féministes et militantes. Il semble donc bien que, dans la presse, la question de l'endométriose fasse partie d'un spectre beaucoup plus large de revendications ou de préoccupations féministes et militantes.

- c. *La militante déplore l'absence de traitement médical, d'examen clinique de détection de l'endométriose ou encore du remboursement des protections hygiéniques.* (corpus *EndoPresse*)

- d. *Or, au même moment, une nouvelle génération de féministes a entrepris de réinvestir le corps des femmes, avec un focus important sur la génitalité : le clitoris, les règles, l'endométriose, les violences gynécologiques et obstétricales, et bien sûr les violences sexuelles, avec le paroxysme que constitue le mouvement #MeToo.* (corpus EndoPresse)

Les autres résultats issus de Sketch Engine (unités extraites dans *EndoPresse* par rapport à *EndoExp*) peuvent par ailleurs être regroupés autour :

- des protections menstruelles et de la précarité menstruelle, avec des unités comme *précarité menstruelle, choc toxique, protection périodique, distributeur de protections, culotte menstruelle* ;
- des violences faites aux femmes en général, avec par exemple *violence sexuelle, viol, violence faite aux femmes, agression sexuelle*. À noter que les aspects juridiques ressortent également, avec des *keywords* comme *justice, accuser, judiciaire, enquête judiciaire* ;
- de l'égalité plus généralement, avec notamment *égalité salariale, égalité hommes-femmes*, ou simplement *égalité et inégalité*.

Là encore, le terme *endométriose* apparaît le plus souvent dans des énumérations. La maladie est donc mise sur le même plan que les autres sujets débattus dans un contexte politique et/ou militant (exemples e à g, avec les unités concernées en gras).

- e. *A l'heure où [...] les menstruations débarquent doucement dans l'espace médiatique et politique, le chemin semble encore très long à parcourir pour que ce tabou soit enfin levé : « Malgré une récente forme de libération de la parole vis-à-vis des menstruations, une meilleure reconnaissance de maladies gynécologiques comme l'**endométriose** ou encore le débat public sur la **précarité menstruelle** [...] (corpus EndoPresse)*
- f. *la parole des femmes se libère pour faire reconnaître les **violences obstétricales**, l'**endométriose** ou encore la difficulté à payer chaque mois tampons et serviettes.* (corpus EndoPresse)
- g. *Preuve en est, donc, qu'avec les règles et l'**endométriose**, la ménopause est aussi un enjeu d'**égalité** qui a besoin d'être abordé bien plus qu'il ne l'est aujourd'hui* (corpus EndoPresse)

À la lueur de ces résultats, il semble que les questions liées à l'égalité hommes-femmes et à la place des femmes dans la société soient étroitement associées aux représentations de l'endométriose véhiculées dans *EndoPresse*. Cependant, étant donné que l'un des objectifs des méthodes employées dans cette section était précisément d'identifier les unités qui se distinguent par leur fréquence inhabituelle dans les corpus, il nous semble nécessaire de compléter ces observations par l'exploration d'autres indices, plus qualitatifs.

4.2 Spécificité de la presse et types d'endométriose

Dans cette section, nous cherchons à interroger le caractère plutôt spécifique ou plutôt générique de la presse, en comparaison avec les textes spécialisés, toujours dans une démarche exploratoire. Dans ce but, nous nous focalisons sur un ensemble de termes plus cohérent d'un point de vue conceptuel que les candidats-termes extraits automatiquement. Il s'agit des différents types d'endométriose.

Pour explorer cette dimension, nous observons la répartition des termes complexes formés sur la base du terme *endométriose* dans les deux corpus, suivant l'hypothèse de l'attestation d'un plus grand nombre de ces termes dans *EndoExp* que dans *EndoPresse*. Dans cette perspective, nous dressons un parallèle entre le nombre de termes complexes attestés dans les corpus et la spécificité qui y est représentée. En effet, les termes complexes que nous visons désignent des types d'endométriose, autrement dit il s'agit d'hyponymes de l'hyperonyme *endométriose*. Ils sont donc, par définition, plus spécifiques que leur hyperonyme (L'Homme 2020 : 102). Par ailleurs, étant donné que l'expansion d'un terme précise les caractéristiques spécifiques de l'hyponyme, nous considérons que, plus l'expansion des termes est longue, plus les concepts auxquels ils renvoient sont spécifiques (par exemple, *endométriose profonde digestive* est plus spécifique qu'*endométriose profonde*, qui est plus spécifique qu'*endométriose*).

Nos observations se fondent sur l'ensemble des termes complexes repérés dans les corpus qui ont *endométri*ose comme tête et une expansion composée d'au moins un modifieur. Cette liste est constituée à partir des candidats-termes extraits automatiquement par TermoStat et elle est complétée par la recherche systématique, avec Sketch Engine, des occurrences du terme *endométri*ose suivies par un modifieur. Nous recensons ainsi 210 termes complexes dans le corpus *EndoExp* et 19 dans le corpus *EndoPresse*. Le Tableau 3 illustre la répartition de ces termes dans les corpus, en fonction du nombre de modifieurs observés.

Tableau 3. Nombre de termes complexes sur la base du terme *endométri*ose dans *EndoExp* et *EndoPresse*

	1 modifieur	2 modifieurs	3 modifieurs	4 modifieurs	Total
Corpus <i>EndoExp</i>	97	94	16	3	210
Corpus <i>EndoPresse</i>	19	0	0	0	19

Dans la majorité des cas, les modifieurs sont :

- des adjectifs (p. ex. *endométri*ose *diaphragmatique*, *endométri*ose *superficielle diaphragmatique droite*, *endométri*ose *profonde postérieure recto-vaginale asymptomatique*),
- des verbes au participe passé avec valeur adjectivale (p. ex. *endométri*ose *diagnostiquée*),
- une combinaison des deux (p. ex. *endométri*ose *complexe opérée*),
- des syntagmes prépositionnels (p. ex. *endométri*ose *du diaphragme*, *endométri*ose *des voies urinaires*).

Au vu de la répartition illustrée dans le tableau, notre hypothèse semble se confirmer : le faible nombre de termes complexes avec *endométri*ose comme tête dans *EndoPresse* semble compatible avec l'hypothèse selon laquelle ce corpus traite de concepts moins spécifiques que le corpus *EndoExp*. Bien que nous nous attendions à ce constat, il nous semble tout de même important de souligner le très faible nombre de ces termes complexes dans le corpus *EndoPresse*. Par ailleurs, la plupart de ces termes sont attestés avec une très faible fréquence dans ce corpus : 15 d'entre eux apparaissent à une seule reprise et deux à deux reprises. Seuls les termes *endométri*ose *profonde* et *endométri*ose *sévère* comptabilisent plus de deux occurrences (12 et 34 respectivement).

En outre, il est possible que l'absence de nombreux termes désignant certains types d'*endométri*ose, comme *endométri*ose *extra-pelvi*enne ou *endométri*ose *thoracique*, contribue à minimiser la sévérité de la maladie, ou du moins suggère-t-elle que la presse ne s'intéresse pas à l'ensemble des localisations et des formes possibles de l'*endométri*ose. Autrement dit, si le nombre élevé de termes attestés dans *EndoExp* pour désigner des types d'*endométri*ose et la longueur parfois considérable de ces termes (avec trois ou quatre modifieurs) peuvent être révélateurs de la complexité de cette maladie, alors il semble que cette complexité n'apparaisse pas dans *EndoPresse*, du moins pas à travers l'emploi de ces mêmes termes.

4.3 Définir l'*endométri*ose dans les corpus

La recherche de contextes définitoires avec Sketch Engine a retourné 180 contextes pertinents dans le corpus *EndoExp* et 116 dans le corpus *EndoPresse*. Deux types de définitions caractéristiques du domaine médical peuvent être identifiés à partir de ces contextes : des définitions *symptomatiques*, qui mettent l'accent sur les symptômes associés à une maladie (notamment ici les symptômes douloureux, comme dans l'exemple h), et des définitions *histologiques*, qui s'attachent à la description de la maladie sur les plans anatomique et biologique, c'est-à-dire ici de ce qui compose les lésions d'*endométri*ose (exemple i).

- h. *L'endométri*ose est une maladie féminine complexe qui entraîne des douleurs pelviennes invalidantes. (corpus *EndoExp*)
- i. *L'endométri*ose est une pathologie gynécologique définie histologiquement par la localisation ectopique de glandes ou de stroma endométrial. (corpus *EndoExp*)

Dans cette section, nous abordons la question de la définition en discours de l'endométriase à travers plusieurs angles et discutons trois aspects : d'abord les différentes unités mises en relation d'hyponymie ou de cause avec *endométriase* et les caractéristiques spécifiques ou les aspects plus descriptifs que l'on peut observer dans les contextes, ensuite les symptômes listés dans les définitions de type symptomatique et enfin les descriptions histologiques de l'endométriase.

4.3.1 Contextes définitoires, hyponymie et causalité

L'observation des contextes repérés à l'aide des marqueurs montre que, dans la plupart des cas, le terme *endométriase* est mis en relation avec une autre unité présentée comme un hyperonyme, plus précisément dans 151 contextes (soit 83,9 %) dans le corpus *EndoExp* et 92 (soit 79,3 %) dans le corpus *EndoPresse*. Dans ces contextes, neuf unités fonctionnent comme des hyperonymes potentiels d'*endométriase*, dont huit dans *EndoExp* et cinq dans *EndoPresse*. Nous considérons à ce stade qu'il s'agit d'hyperonymes *potentiels* car, bien qu'apparaissant dans des patrons similaires (par exemple « déterminant + terme spécifique + être un + terme générique », comme dans « l'endométriase est une pathologie ») et bien que les marqueurs employés pour les repérer puissent indiquer une relation d'hyponymie, il n'est pas certain que ces unités soient toutes des hyperonymes « réels » du terme *endométriase*.

Parmi ces unités, deux groupes peuvent être distingués : d'une part, des unités qui montrent que l'endométriase est un type de maladie, avec *maladie*, *affection*, *pathologie* et *atteinte*, et qui sont donc bien des hyperonymes du terme à l'étude (notons également l'unité *entité*, que nous relevons dans deux contextes dans le corpus *EndoExp* et qui renvoie à un concept encore plus générique) ; d'autre part, des unités qui indiquent que l'endométriase est la cause d'un symptôme ou d'un dysfonctionnement, avec *cause*, *facteur*, *facteur de risque* et *source*. Ces dernières n'illustrent pas véritablement une relation d'hyponymie avec *endométriase* mais plutôt une relation de causalité dans la plupart des cas. Dans les deux corpus, le premier groupe représente la majorité des contextes, 131 (ou 86,8 %) dans *EndoExp* et 86 (ou 93,5 %) dans *EndoPresse*, et l'hyperonyme le plus fréquent est *maladie* dans les deux corpus. Notons cependant que *pathologie* apparaît à la même fréquence dans le corpus *EndoExp*, soit 62 fois, contre sept occurrences dans *EndoPresse*.

Au-delà de ces aspects plutôt généraux, plusieurs différences entre les deux corpus sont à remarquer. D'abord, pour les contextes avec les unités classées dans le premier groupe, on observe que les caractéristiques spécifiques sont exprimées avec différents adjectifs, les plus fréquents dans les deux corpus étant *gynécologique*, *chronique* et *fréquent* (exemples j à l). Cependant, l'observation des adjectifs employés dans l'un des deux corpus uniquement montre que, dans *EndoExp*, l'accent est surtout placé sur le caractère complexe et hétérogène de l'endométriase, avec des adjectifs comme *hétérogène*, *pluriel*, *plurifactoriel*, *polygénique* et *polymorphe* (exemples l et m). En revanche, dans *EndoPresse*, on retrouve des adjectifs mettant en avant le vécu des patientes, avec notamment *difficile* (dans *difficile à vivre*, par exemple, en k) et *invisible*. On observe également, et de manière plus flagrante, des adjectifs et des locutions qui soulignent les lacunes dans la connaissance et la prise en charge de l'endométriase, avec par exemple *mal comprise*, *méconnue*, *peu connue*, *mal diagnostiquée* (exemples j et k), ainsi que, dans une certaine mesure, des lacunes dans la recherche sur cette maladie, avec *sous étudiée*. À noter que la frontière entre la nature strictement définitionnelle et le caractère plutôt descriptif de ce type de contextes reste parfois difficile à établir.

- j. *L'endométriase est une maladie gynécologique encore trop peu connue alors qu'on la retrouve en moyenne chez 10 % des femmes.* (corpus *EndoPresse*)
- k. *L'endométriase est une maladie chronique gynécologique, mal connue, difficile à vivre* (corpus *EndoPresse*)
- l. *L'endométriase est une pathologie gynécologique multifactorielle* (corpus *EndoExp*)
- m. *L'endométriase est une affection très polymorphe avec une symptomatologie très variée.* (corpus *EndoExp*)

Ensuite, pour les contextes avec les unités classées dans le deuxième groupe, l'unité *cause* est la plus fréquente dans les deux corpus. Cette similarité révèle cependant des différences importantes lorsque l'on observe les contextes en détail. Dans *EndoPresse*, l'endométriose est présentée comme une cause d'infertilité dans tous les contextes relevés (exemple n), alors que, dans *EndoExp*, la question de l'infertilité est évoquée en association avec d'autres symptômes ou dysfonctionnements, tels que les douleurs ou les dysménorrhées, ou encore la qualité de vie (exemples o et p).

- n. *L'endométriose est la première cause d'infertilité chez les femmes.* (corpus *EndoPresse*)
- o. *L'endométriose est la cause de 25 à 50% des cas d'infertilité, de 25% des cas de douleurs pelviennes et de 20% des cas de dysménorrhées.* (corpus *EndoExp*)
- p. *L'endométriose est une cause fréquente de douleurs, d'hypofertilité, et d'altération de la qualité de vie des patientes* (corpus *EndoExp*)

Cela ne signifie toutefois pas que, dans le corpus *EndoPresse*, l'altération de la qualité de vie des femmes atteintes d'endométriose n'apparaît jamais dans ce type de contextes. Cet aspect est au contraire bien présent, notamment lorsque l'unité *source* le met en relation avec *endométriose*. Il s'observe ainsi par exemple à travers le terme *précarité professionnelle* en q ou de manière plus explicite en r :

- q. *L'endométriose est une source indéniable de **précarité professionnelle*** (corpus *EndoPresse*)
- r. *L'endométriose est une source de handicap invisible **impactant fortement la qualité de vie des femmes atteintes*** (corpus *EndoPresse*)

Ces différences de point de vue sur l'endométriose rappellent les différences dans les thématiques repérées dans la section 4.1, avec notamment l'infertilité, qui semble plus centrale dans *EndoPresse* que dans *EndoExp*. Ces observations permettent par ailleurs de compléter ces thématiques, en particulier avec les lacunes dans les connaissances actuelles sur cette maladie, mises en lumière dans *EndoPresse*. À ce stade, il semble que ces questions soient au cœur des représentations de la maladie dans la presse.

4.3.2 Point de vue symptomatique

Dans le corpus *EndoExp*, 22 contextes illustrant des définitions symptomatiques ont été repérés et 12 dans *EndoPresse*. Dans les définitions symptomatiques, la principale différence entre les corpus concerne le degré de détail dans les manières de décrire les symptômes associés à l'endométriose. En effet, dans la moitié des cas dans *EndoExp*, aucun symptôme n'est évoqué et l'on observe seulement les unités *symptôme*, *symptomatologie* et *symptomatique*. Dans six cas, la notion de variabilité de ces symptômes apparaît également ; les noms *symptôme* et *symptomatologie* sont alors modifiés par les adjectifs *variable* et *varié* (par exemple *symptomatologie très variée*, *symptômes variables*). Dans les autres cas, les symptômes douloureux sont évoqués, mais sans décrire précisément les types de douleurs ou leurs localisations, dans des syntagmes comme *douleur pelvienne* et *symptôme douloureux*, ou avec l'adjectif *douloureux* qui modifie l'hyperonyme, comme dans l'exemple ci-dessous.

- s. *L'endométriose est une maladie chronique, complexe, **douloureuse*** (corpus *EndoExp*)

Au contraire, dans le corpus *EndoPresse*, la douleur est centrale dans ces contextes. Sur les 12 contextes repérés, seul un ne l'évoque pas. Dans les autres cas, l'intensité des douleurs est exprimée avec des adjectifs comme *violent* ou *intense* (exemples t et u) et, dans trois cas, les types de douleurs sont précisés dans des énumérations de symptômes (exemple v).

- t. *maladie [...] qui se manifeste par de **violentes douleurs pelviennes*** (corpus *EndoPresse*)
- u. *L'endométriose est une maladie [...] provoquant d'**intenses douleurs*** (corpus *EndoPresse*)
- v. ***douleurs gynécologiques récurrentes, telles que des règles douloureuses, de douleurs pelviennes, une sexualité douloureuse, des problèmes urinaires, des troubles intestinaux ou des fatigues chroniques*** (corpus *EndoPresse*)

Ces observations ne signifient pas que l'étendue et la diversité des symptômes associés à l'endométriose ne sont pas traitées dans les textes des spécialistes, mais plutôt que les symptômes douloureux ne sont pas systématiquement inclus dans les contextes définissant explicitement l'endométriose. Parallèlement, l'omniprésence, dans *EndoPresse*, de la douleur, de son intensité et des régions du corps qu'elle concerne peut être mise en relation avec les aspects plus militants observés à plusieurs reprises dans ce corpus.

4.3.3 Point de vue histologique

Dans cette section, nos observations se fondent sur les 47 et 44 contextes présentant des définitions histologiques dans *EndoExp* et dans *EndoPresse* respectivement. Ces définitions sont composées de trois caractéristiques. Si nous reprenons la définition de l'endométriose donnée en 2.3, à savoir « la présence en dehors de la cavité utérine de tissu semblable à la muqueuse utérine », ces caractéristiques sont :

- « la présence », c'est-à-dire l'existence dans le corps,
- « en dehors de la cavité utérine », qui précise la localisation,
- « tissu semblable à la muqueuse utérine », qui décrit la composition des lésions d'endométriose.

Dans les contextes définitoires correspondant à ce type de définition, deux différences principales entre les corpus sont observées. D'une part, les manières d'exprimer la première caractéristique varient fortement dans le corpus *EndoPresse*. En effet, si, dans le corpus *EndoExp*, cette caractéristique est majoritairement exprimée avec l'unité *présence* (dans 81 % des contextes repérés), elle ne l'est que dans 57 % des contextes dans *EndoPresse*. Les formulations alternatives observées dans *EndoExp* sont par ailleurs plutôt neutres, avec par exemple les noms *ectopie*, *localisation ectopique* ou encore le verbe *se retrouver*. En revanche, on voit apparaître une idée de mouvement dans *EndoPresse*, en particulier avec *migration anormale*, *reflux*, *remonter (dans le corps)* et *migrer*. Cette idée de mouvement, qui n'est pas présente chez les spécialistes, est renforcée par l'emploi de *coloniser* dans plusieurs contextes, comme dans l'exemple ci-dessous.

- w. « L'endométriose est une maladie gynécologique chronique de la femme en âge de procréer qui se caractérise par le développement d'une muqueuse utérine (l'endomètre) en dehors de l'utérus, **colonisant** d'autres organes avoisinants », définit le ministère de la Santé. (corpus *EndoPresse*)

D'autre part, nous observons des variations importantes à propos de la troisième caractéristique également. Par exemple, dans *EndoExp*, bien que cette caractéristique soit le plus souvent exprimée par *tissu endométrial* et *glandes et stroma endométraux*, elle l'est dans 11 % des contextes par *endomètre*, *cellules de l'endomètre* et *tissu de l'endomètre*, ce qui s'éloigne considérablement de *tissu semblable à la muqueuse utérine*, que nous citons au début de cette section, et semble même contradictoire. L'idée de ressemblance entre ce qui compose l'endométriose et l'endomètre est tout de même évoquée à plusieurs reprises dans ce corpus, dans 6 % des contextes, avec *tissu de type endométrial* et *ressemblant à du tissu endométrial*.

Ces proportions marginales dans *EndoExp* sont exacerbées dans *EndoPresse* : bien que nous retrouvions ces trois tendances, les formulations dénotant un rapprochement avec l'endomètre représentent 34 % des contextes (par exemple *endomètre*, *fragments de l'endomètre*) et celles explicitant une ressemblance avec l'endomètre 23 % (par exemple *tissu semblable à la muqueuse utérine (endomètre)*). En revanche, les descriptions plus consensuelles dans *EndoExp* n'apparaissent qu'une fois pour *tissu endométrial* dans *EndoPresse* ; *glandes et stroma endométraux* n'étant pas attesté dans ce corpus. Enfin, notons une description caractéristique d'*EndoPresse*, qui représente 25 % des contextes : *cellules d'origine utérine*. Cette description apparaît toujours dans la même définition, qui semble être reprise d'associations de patientes, comme l'illustre l'exemple x.

- x. *l'endométriose, une maladie qui se caractérise par la présence de cellules d'origine utérine en dehors de l'utérus [...], selon les associations.* (corpus *EndoPresse*)

En plus des différences, parfois importantes, repérées entre et dans les corpus, ces observations mettent surtout en lumière le rôle d'autres intermédiaires dans la diffusion d'informations médicales. En particulier, la définition illustrée dans l'exemple w est entre guillemets et provient du ministère de la Santé et, dans

l'exemple x, il est précisé que la définition provient d'associations. Dans les deux cas, les définitions apparaissent à plusieurs reprises à l'identique dans le corpus *EndoPresse*, sans les références systématiques au ministère de la Santé ou aux associations, ce qui suggère qu'elles sont bien diffusées dans la presse. Il semble alors que différents intermédiaires, plus ou moins officiels, jouent un rôle important non seulement dans la diffusion des informations médicales, mais également dans les représentations de l'endométriose présentes dans la presse. Nous discuterons davantage ce point dans la conclusion.

Enfin, soulignons encore une fois les divergences observées dans les deux corpus à propos de la caractéristique concernant la composition de l'endométriose. Ces divergences, voire ces contradictions, renvoient en réalité au fait que définir l'endométriose n'est pas une tâche si aisée pour les spécialistes, dont les avis diffèrent manifestement. Il semble alors évident que ces désaccords se retrouvent également dans la presse, notamment par le biais des propos des spécialistes qui y sont rapportés. En revanche, à ce stade de notre réflexion, nous pensons que ce type de décalages peut en réalité nuire à l'efficacité de la diffusion des connaissances, voire constituer une vraie source de confusion pour les non-spécialistes qui consultent des articles de presse.

5 Remarques conclusives

Nous avons cherché dans cet article à explorer les représentations de l'endométriose dans la presse en comparaison avec les textes spécialisés. À partir de l'observation du fonctionnement de plusieurs termes et d'autres unités dans un corpus de presse généraliste en comparaison avec un corpus spécialisé sur l'endométriose, nous avons mis au jour plusieurs éléments permettant d'approcher les représentations de la maladie dans le corpus *EndoPresse*, notamment lorsqu'elles diffèrent du corpus *EndoExp*. L'un des éléments constants, qui est ressorti à chaque étape de l'analyse, concerne le traitement de thématiques différentes dans les deux corpus, et particulièrement l'existence de thématiques propres à la presse. Nous avons ainsi mis en lumière, aussi bien avec des indices plutôt quantitatifs qu'avec des indices plutôt qualitatifs, l'association presque systématique de l'endométriose avec des questions d'infertilité et la récurrence de thématiques féministes et militantes, dont le lien avec l'endométriose n'est d'ailleurs pas toujours évident de prime abord. C'est le cas des thématiques liées à l'égalité hommes-femmes et aux violences faites aux femmes, que nous avons surtout illustrées dans la section 4.1. Dans d'autres cas, la centralité de plusieurs aspects semble mettre les femmes au cœur des préoccupations liées à l'endométriose dans la presse, évoquant en ce sens une orientation féministe également. Parmi ces aspects, on retrouve notamment les douleurs, leur intensité et leurs nombreuses localisations possibles dans le corps, ou encore les lacunes dans les connaissances sur l'endométriose, qui impactent le diagnostic de la maladie et la prise en charge thérapeutique des femmes atteintes, ainsi que leur qualité de vie de manière plus générale. Nous avons abordé ces points principalement dans la section 4.3. À la lueur de ces observations, il semble que les préoccupations et revendications féministes et militantes fassent partie intégrante des représentations de l'endométriose qui sont véhiculées dans la presse, et même qui se distinguent de celles observables dans les discours des spécialistes.

Parallèlement, à travers nos observations, nous avons montré la nécessité de toujours tenir compte des contextes d'apparition des termes et des autres unités dans les corpus, en plus des informations strictement quantitatives. C'est en effet la combinaison de différents indices quantitatifs et qualitatifs qui a permis de mettre en évidence le fait que les unités les moins fréquentes dans le corpus *EndoPresse*, parmi l'ensemble des unités identifiées automatiquement, apparaissent en réalité le plus souvent dans du discours rapporté. L'emploi de certains termes et unités dans les articles de presse ne relève ainsi pas uniquement de la volonté propre des journalistes, mais également des sources qu'ils choisissent de citer. Ce constat pose la question des « discours représentés » (Moirand et Reboul-Touré 2023) dans la presse et interroge leur rôle non seulement dans les processus de diffusion des connaissances auprès du grand public, mais également dans le processus global de circulation des termes. Dans nos données, nous avons en effet constaté la pluralité des voix présentes dans la presse, autrement dit la pluralité des discours qui y sont représentés, avec bien évidemment les voix des spécialistes (comme nous le montrons notamment dans Humbert-Droz et Picton (2022) dans un autre domaine), mais également les voix de plusieurs organismes, comme le ministère de la

Santé, l'Institut national de la santé et de la recherche médicale (Inserm), EndoFrance et d'autres associations de patientes. De la même manière, les voix des patientes sont également représentées à travers certaines citations dans la presse. Ces dernières se distinguent par leur caractère « non officiel », mais gardent toute leur importance pour identifier les représentations de l'endométriose et, de manière plus générale, pour nourrir la réflexion plus globale sur la notion de circulation des termes. Ces remarques, et ce dernier point en particulier, montrent, d'une part, la complexité des dynamiques à l'œuvre dans les processus de circulation des termes et, d'autre part, la nécessité de tenir compte de discours plus variés pour mieux comprendre comment les termes liés à l'endométriose circulent entre les spécialistes, les patientes et le grand public. Nous poursuivrons ainsi notre réflexion à l'aide de données plus diversifiées, qui permettront de refléter les discours des patientes, autrement dit les voix des femmes atteintes d'endométriose, et les discours des associations, qui semblent jouer ici un rôle central.

Pour terminer, avec cet article, nous avons dressé un premier portrait des représentations de l'endométriose dans la presse écrite française, en comparaison avec les discours des spécialistes sur cette maladie. Son orientation exploratoire et les nombreuses pistes qui en ressortent, et qu'il nous faudra approfondir, montrent tout l'intérêt d'une approche linguistique et terminologique des discours sur l'endométriose. La poursuite de notre projet, avec notamment la prise en compte de discours plus diversifiés, nous mènera à mieux comprendre comment les connaissances sur cette maladie se forment et circulent auprès de différents types de locuteurs. Les perspectives nouvelles qui en découleront auront également pour objectif de poser des bases pour initier une future collaboration interdisciplinaire avec des acteurs de la santé, comme le suggèrent par exemple Sanselme-Cardenas (2022) ou Ablali et Wiederspiel (2023). Par ailleurs, d'un point de vue appliqué, les pistes qui émergeront, outre celles qui émergent déjà de cette première étude, pourront, nous l'espérons, servir à la création de ressources plus adaptées au grand public et aux patientes, dans un cadre d'éducation thérapeutique ou non, à la suite par exemple des propositions de Carretier et al. (2009). Nous pensons en particulier qu'il est fondamental aujourd'hui non seulement de tenir compte des divergences que l'on peut observer entre la presse et les discours des spécialistes (et même à l'intérieur de la communauté de spécialistes, où coexistent des points de vue parfois très différents (Humbert-Droz 2023)), mais également de les expliquer et de préciser les enjeux sous-tendant les recherches sur l'endométriose, en replaçant les discours des spécialistes dans leur contexte, voire en les relégitimisant (Lamy 2017). Tout cela dans l'objectif global de contribuer à mieux informer les femmes concernées par l'endométriose et le grand public de manière plus générale.

Références bibliographiques

- Ablali, D., Wiederspiel, B. (2023). La santé publique : Défis et enjeux sociétaux pour les sciences du langage. Présentation. *Langages*, 231(3), « Textes et discours face à des questions de santé publique » : 9-24.
- Balfour, J. (2023). *Representing schizophrenia in the media: A corpus-based approach to UK press coverage*. Londres : Routledge.
- Beacco, J.-C., Claudel, C., Doury, M., Petit, G., Reboul-Touré, S. (2002). Science in media and social discourse: New channels of communication, new linguistic forms. *Discourse Studies*, 4(3) : 277-300.
- Biber, D. (1993). Representativeness in Corpus Design. *Literary and Linguistic Computing*, 8(4): 243-257.
- Cappeau, P., Schnedecker, C. (2022). La circulation des désignations relatives à la pathologie dans les consultations médicales. Dans M. Blasco (Éd.), *Parler à l'hôpital. Écouter ce qui est dit, décrypter ce qui se dit*, 88-104. Münster : Nodus Publikationen.
- Carretier, J., Delavigne, V., Fervers, B. (2009). Du langage expert au langage patient : Vers une prise en compte des préférences des patients dans la démarche informationnelle entre les professionnels de santé et les patients. *Sciences-Croisées*, 6.
- Condamines, A., Picton, A. (2014). Des communiqués de presse du Cnes à la presse généraliste. Vers un observatoire de la diffusion des termes. Dans P. Dury, J. de Hoyos, J. Makri-Morel, F. Maniez, V. Renner, & M. B. Villar Diaz (Éds.), *La néologie en langue de spécialité : Détection, implantation et*

- circulation des nouveaux termes*, 165-188. Lyon : Centre de recherche en terminologie et traduction, Université Lumière Lyon 2.
- Condamines, A., Picton, A. (2022). Textual terminology: Origins, principles and new challenges. Dans P. Faber & M.-C. L'Homme (Éds.), *Theoretical perspectives on terminology: Explaining terms, concepts and specialized knowledge* : 219-236. Amsterdam : John Benjamins.
- Costes, M., Laval, S. (2019). Le vocable de la schizophrénie au prisme de la presse quotidienne nationale française : Un effet de mode ? *Les Cahiers du journalisme - Recherches*, 2(3) : 29-44.
- Delavigne, V. (2001). *Les mots du nucléaire. Contribution socioterminologique à une analyse des discours de vulgarisation*. Thèse de doctorat. Université de Rouen.
- Delavigne, V. (2019). Littératies en santé et forums de patients : Des formes d'ergonomie discursive. *Éla. Études de linguistique appliquée*, 195(3) : 363-381.
- Delavigne, V. (2022). Intégrer les usages dans une ressource terminologique : Quelques pistes méthodologiques. *Ela. Études de linguistique appliquée*, 208(4) : 397-415.
- Delavigne, V., Gaudin, F. (2022). Founding principles of Socioterminology. Dans P. Faber & M.-C. L'Homme (Éds.), *Theoretical perspectives on terminology: Explaining terms, concepts and specialized knowledge* : 177-195. Amsterdam : John Benjamins.
- Delavigne, V., Picton, A., Thibert, E. (2022). Socioterminologie et terminologie textuelle : L'expertise en questions. *Congrès Mondial de Linguistique Française (CMLF 2022)*.
- Delforge, F. (2016). *Place des médias dans la relation médecin-malade en médecine générale : Impact de la médiatisation d'une question de santé publique*. Thèse de doctorat. Université Amiens Jules Verne.
- Dempster, G., Sutherland, G., Keogh, L. (2022). Scientific research in news media: A case study of misrepresentation, sensationalism and harmful recommendations. *Journal of Science Communication*, 21(1), A06.
- Drouin, P. (2003). Term Extraction Using Non-Technical Corpora as a Point of Leverage. *Terminology*, 9(1) : 99-117.
- Drouin, P., Morel, J.-B., L'Homme, M.-C. (2020). Automatic term Extraction from newspaper corpora: Making the most of specificity and common features. *Proceedings of the 6th International Workshop on Computational Terminology (COMPUTERM 2020)*: 1-7.
- EndoFrance. (2023). L'endométriase. <https://www.endofrance.org/> [dernière consultation : 4 octobre 2023].
- Galès, M.-R. (2020). *Endométriase : Ce que les autres pays ont à nous apprendre*. Paris : Éditions Josette Lyon.
- Gaudin, F. (1993). *Pour une socioterminologie : Des problèmes sémantiques aux pratiques institutionnelles*. Rouen : Publications de l'Université de Rouen.
- Gaudin, F. (2003). *Socioterminologie : Une approche sociolinguistique de la terminologie*. Bruxelles : De Boeck-Duculot.
- Habert, B. (2004). Outiller la linguistique : De l'emprunt de techniques aux rencontres de savoirs. *Revue française de linguistique appliquée*, 9(1) : 5-24.
- HAS, CNGOF. (2017). *Prise en charge de l'endométriase – Recommandations de bonne pratique*.
- Humbert-Droz, J. (2021). Définir la détermination : Approche outillée en corpus comparable dans le domaine de la physique des particules. Thèse de doctorat. Université Toulouse - Jean Jaurès, Université de Genève.
- Humbert-Droz, J. (2023). Analyse de la terminologie de l'endométriase en corpus : propositions pour une prise en compte de la pluridisciplinarité. *XIX^e Journée Scientifique REALITER – Terminologie et interdisciplinarité : défis et perspectives de recherche futures*. Paris, 13 octobre 2023.
- Humbert-Droz, J., Picton, A. (2022). Du Higgs à la particule de Dieu : Une analyse de la variation terminologique dans un contexte de détermination. *Congrès Mondial de Linguistique Française (CMLF 2022)*.

- Ilschner, S., Neeman, T., Parker, M., Phillips, C. (2022). Communicating endometriosis pain in France and Australia: An interview study. *Frontiers in Global Women's Health*, 3.
- Jacques, M.-P. (2016). Une linguistique outillée, pour quels objets ? *Histoire Épistémologie Langage*, 38(2) : 87-99.
- Kennedy, G. (1998). *An Introduction to Corpus Linguistics*. Londres, New York : Longman.
- Kilgariff, A., Rychly, P., Smrz, P., Tugwell, D. (2004). The Sketch Engine. *Proceedings of the 11th EURALEX International Congress* : 105-116.
- King, M., Watson, K. (2005). Introduction. Dans M. King & K. Watsin (Éds.), *Representing Health: Discourses of Health and Illness in the Media* : 1-22. Londres : Palgrave.
- Lamy, A. (2017). Mise en cause de l'autorité médicale et légitimation du discours d'expérience sur les forums de discussion en ligne. *Quaderni. Communication, technologies, pouvoir*, 93 : 43-52.
- Le Draoulec, A., Péry-Woodley, M.-P., & Rebeyrolle, J. (2014). Glissements progressifs de « sémantique ». *Le Discours et la Langue*, 1(6) : 109-126.
- Leech, G. (2007). New Resources, or Just Better Old Ones? The Holy Grail of Representativeness. Dans M. Hundt, N. Nesselhauf, & C. Biewer (Éds.), *Corpus Linguistics and the Web*, 133-149. Amsterdam : Rodopi.
- Lefevre, L. (2017). *Analyse des marqueurs de relations conceptuelles en corpus spécialisé : Recensement, évaluation et caractérisation en fonction du domaine et du genre textuel*. Thèse de doctorat. Université Toulouse - Jean Jaurès.
- L'Homme, M.-C. (2020). *La terminologie : Principes et techniques*. 2e édition, revue et mise à jour. Montréal : Les Presses de l'Université de Montréal.
- Meyer, I. (2000). Computer words in our everyday live: How are they interesting for terminography and lexicography? *Proceedings of EURALEX 2000* : 39-58.
- Meyer, I. (2001). Extracting knowledge-rich contexts for terminography: A conceptual and methodological framework. Dans D. Bourigault, C. Jacquemin, & M.-C. L'Homme (Éds.), *Recent advances in computational terminology* : 279-302. Amsterdam : John Benjamins.
- Moirand, S. (2007). *Les discours de la presse quotidienne. Observer, analyser, comprendre*. Paris : Presses universitaires de France, Linguistique nouvelle.
- Moirand, S., Reboul-Touré, S. (2023). Itinéraires discursifs d'une nomination : La santé publique en temps de pandémie. *Langages*, 231(3), « Textes et discours face à des questions de santé publique » : 25-41.
- Pearson, J. (1998). *Terms in context*. Amsterdam/Philadelphie : John Benjamins.
- Rebeyrolle, J. (2000). *Forme et fonction de la définition en discours*. Thèse de doctorat, Université Toulouse 2.
- Rouillard, C.-A. (2016). Usage spécialisé ou usage courant des désignations de la maladie mentale dans le discours des non-spécialistes : Le cas des termes autisme et hystérie. *ScriptUM : la revue du colloque VocUM*, 2 : 21-40.
- Sanselme-Cardenas, M.-É. (2022). Postface – Et si des linguistes pouvaient participer à la réhumanisation du soin ? Point de vue d'un médecin. Dans M. Blasco (Éd.), *Parler à l'hôpital. Écouter ce qui est dit, décrypter ce qui se dit* : 273-287. Münster : Nodus Publikationen.
- Scott, M. (1997). PC Analysis of Key Words—And Key Key Words. *System*, 25(2) : 233-245.
- Sierra, G., Torres-Moreno, J.-M., Lorente, M. (2022). Exploration de contextes définitoires en français : Cas d'étude sur la COVID-19. *Actes des 16èmes Journées Internationales d'Analyse Statistique de Données Textuelles* : 817-826.
- Sinclair, J. (1991). *Corpus, Concordance, Collocation*. Oxford : Oxford University Press.
- Sinclair, J. (1994). *Corpus typology: A framework for classification*. EAGLES document : 1-18.
- Ungureanu, L. (2006). *L'interpénétration langue générale-langue spécialisée dans le discours d'internet*. Paris : Connaissances et Savoirs.

Young, K., Fisher, J., Kirkman, M. (2015). Women's experiences of endometriosis: A systematic review and synthesis of qualitative research. *Journal of Family Planning and Reproductive Health Care*, 41(3): 225-234.

¹ Ce projet, intitulé « La circulation des termes de l'endométriose : de la compréhension à l'appropriation des termes », est financé par le Fonds national suisse (FNS).

² Nous entendons ici des principes de sélection des textes élaborés pour que les corpus représentent au mieux les discours visés par l'analyse (les discours de la presse française et les discours des professionnels de santé sur l'endométriose), autrement dit pour que les corpus servent d'*échantillons représentatifs* (p. ex. Sinclair 1991, Biber 1993) de ces deux types de discours, ou du moins, tenant compte des critiques de la notion de représentativité en linguistique de corpus (p. ex. Kennedy 1998, Leech 2007), pour que les corpus reflètent au mieux ces discours.

³ EndoFrance, Nos actions, <https://www.endofrance.org/nos-actions/endofrance-redactrice-groupe-de-travail-cngof-has-2017/> (dernière consultation : 22 novembre 2023).

⁴ Bien que le corpus *EndoPresse* ne soit pas constitué de textes spécialisés, contrairement aux corpus habituellement exploités avec des extracteurs de termes, TermoStat reste pertinent pour identifier les unités lexicales qui se distinguent par leur fréquence dans ce corpus, en comparaison avec le corpus de référence de l'outil (Drouin et al. 2020 : 2).

⁵ Par exemple le verbe *être* suivi d'un déterminant, dans des séquences comportant un hyperonyme et les caractéristiques spécifiques de son hyponyme, comme dans « L'endométriose est une maladie inflammatoire chronique œstrogéno-dépendante » (corpus *EndoExp*), où *maladie* est présenté comme un hyperonyme d'*endométriose* et où les adjectifs suivant *maladie* expriment certaines caractéristiques spécifiques de l'hyponyme. L'hyperonyme n'est pas systématiquement présent avec d'autres verbes fonctionnant comme marqueurs, par exemple le verbe *se définir*, lorsqu'il est suivi d'une préposition, comme dans « L'endométriose se définit par la présence de tissu endométrial en dehors de la cavité utérine » (corpus *EndoExp*). Pour plus de détails, voir notamment Rebeyrolle (2000), Meyer (2001), Lefevre (2017).

⁶ Ressources développées à CLLE, MAR-REL – Marqueurs de relations, http://redac.univ-tlse2.fr/misc/mar-rel_fr.html (dernière consultation : 26 octobre 2023).