

Présence des noms d'agents au féminin dans un corpus de presse. Étude comparée Québec-Belgique.

Anne Dister¹, Hubert Naets², Patrick Watrin²

¹Université Saint-Louis – Bruxelles – anne.dister@usaintlouis.be

²Université Catholique de Louvain – hubert.naets@uclouvain.be
patrick.watrin@uclouvain.be

Abstract

In this article, we are interested in the implementation of feminine names of professions, titles, ranks and functions in the press. We analyse two French-language daily newspapers, one Belgian (*Le Soir*) and the other Quebecois (*Le Devoir*) at a 20-year interval (2001 and 2021). The results we obtain show that female names are much more numerous in 2021, but also indicate a difference between Quebec and Belgium.

Keywords: feminization, corpus, diachrony, Belgium, Quebec, Unitex

Résumé

Dans cet article, nous nous intéressons à l'implantation des noms de professions, titres, grades et fonctions au féminin dans la presse. Pour ce faire, nous analysons deux quotidiens francophones, l'un belge (*Le Soir*) et l'autre québécois (*Le Devoir*) à 20 années d'intervalle (2001 et 2021). Les résultats que nous obtenons montrent que les noms au féminin sont beaucoup plus nombreux en 2021, mais indiquent aussi une différence entre le Québec et la Belgique.

Mots clés: féminisation, corpus, diachronie, Belgique, Québec, Unitex.

1. Introduction

La féminisation des noms de professions, titres, grades et fonctions est loin d'être un sujet de d'études récent, que ce soit du point de vue social ou linguistique. C'est dans les années 1970 que le Québec légifère sur la question, suivi en cela par les autres pays de la Francophonie Nord (pour un aperçu historique, voir Dister et Moreau 2009). Le principe est simple, et guidé par le souci de visibiliser les femmes dans la sphère professionnelle, conformément à la morphologie du français et au fonctionnement de la langue : on préconise d'utiliser un nom au féminin pour désigner une ou des femmes. On dira ainsi *Madame l'ambassadrice* et non **Madame l'ambassadeur*, *une avocate* et non **un avocat*, *une juge assermentée* et non **un juge assermenté*. Les décisions politiques ont la plupart

du temps été suivies de la publication de guides et encouragées par des personnalités publiques qui ont revendiqué un titre au féminin. La presse est, pour la question qui nous occupe, un bon observatoire des pratiques, comme l'a montré Fujimura (2005) qui a étudié l'évolution des noms d'agent au féminin dans deux journaux français, avant et après la prise de décision politique de 1998 par le gouvernement de Lionel Jospin. Dans cet article, c'est sur la presse belge francophone et québécoise que notre attention s'est portée, dans une perspective également diachronique.

2. Méthodologie

2.1. Collecte des données

Nous avons constitué notre corpus à l'aide du service *Europresse* (<http://www.europresse.com/fr/>), une base de données payante d'informations en ligne disponible sur internet. Nous avons récupéré manuellement tous les articles proposés par *Europresse*¹ pour le journal montréalais *Le Devoir* et pour le journal belge francophone *Le Soir* pour les années 2001 et 2021. Les deux quotidiens ont un lectorat et un positionnement assez comparables. Nous avons choisi, d'une part, l'année 2021, car il s'agit de l'année complète la plus récente et, d'autre part, l'année 2001, car il nous a semblé qu'un écart de vingt ans devrait permettre d'observer une évolution significative des usages.

Tableau 1. Détail des corpus collectés

	2001		2021	
	#Articles	#Tokens	#Articles	#Tokens
<i>Le Devoir</i>	24 939	13 617 976	14 549	10 152 902
<i>Le Soir</i>	48 800	21 394 200	26 949	11 328 943

2.2. Constitution des ressources

Afin d'identifier et d'extraire les phrases contenant un nom de métier repris dans Moreau et Dister (2014), nous avons utilisé Unitex (Paumier, 2021), une suite logicielle libre d'analyse de corpus. Cet outil repose sur la création de ressources linguistiques :

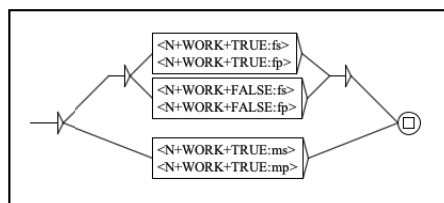
- des dictionnaires (Courtois et Silberztein, 1990) qui associent des informations morphosyntaxiques à chaque mot du texte ;
- des grammaires locales (Gross, 1997) qui formalisent des structures lexico-syntaxiques intégrant les mots et leurs informations morphosyntaxiques.

¹ Nous sommes tributaires des données rendues disponibles par Europresse, qui ne propose pas tous les articles publiés.

Outre les dictionnaires du français intégrés à l'outil, nous avons développé un dictionnaire de noms de métiers qui, à chaque entrée non épicène², associe des informations flexionnelles (de manière à permettre la distinction masculin/féminin) et un ensemble d'informations sémantiques. Ces dernières ont pour objectif d'identifier les noms de métier au sein de nos grammaires locales (étiquette +*WORK*) mais également de distinguer les formes recommandées (étiquette +*TRUE*) des formes construites artificiellement (étiquette +*FALSE*). En effet, pour permettre l'étude de la créativité lexicale, il nous fallait être à même d'observer l'ensemble des possibles, mais également de les distinguer lors de nos analyses. Notons, finalement, que nous avons séparé (dans deux dictionnaires distincts, filtré et non filtré) les formes féminines sémantiquement non ambiguës des formes potentiellement ambiguës (c'est-à-dire ne désignant pas spécifiquement un nom de métier, comme *couvreuse* ou *moissonneuse* qui désignent soit le métier soit l'outil).

Nos grammaires locales sont quant à elles rudimentaires dans la mesure où leur seule fonction est l'identification des formes reprises dans le dictionnaire des noms de métiers (cf. Figure 1).

Figure 1. Grammaire locale principale



2.3. Extraction des candidats

L'application des dictionnaires et de notre grammaire locale sur nos corpus³ nous a permis d'extraire les phrases contenant un nom de métier. Certaines formes pouvant également être utilisées en tant qu'adjectif (culture *religieuse*, par exemple), nous avons utilisé un étiqueteur morphosyntaxique⁴ afin d'identifier les formes nominales et d'ainsi faciliter l'analyse des résultats.

L'ensemble des phrases extraites a été stocké pour chaque corpus au sein d'un fichier CSV (cf. Table 2). Chaque ligne est composée de quatre colonnes spécifiant le lemme, la forme, la nature nominale de la forme ("1" si la forme

² Notre objectif étant l'observation de la féminisation des noms de métier, il nous a semblé difficile d'intégrer les formes qui ne marquent pas graphiquement le féminin. Ces formes ont donc été manuellement supprimées de la liste.

³ Pour permettre l'intégration d'Unitex dans notre outil d'extraction, nous avons utilisé la librairie python-unitex (<https://github.com/patwat/python-unitex>).

⁴ <https://spacy.io/>

est un nom et “0” dans le cas contraire) et la phrase d’origine. Si une phrase contient plus d’une forme de nom de métier, plusieurs lignes sont créées.

Tableau 2. Exemple d’extraction pour *Le Devoir* 2021

Lemme	Forme	Nom	Phrase
auteur	autrice	1	Après sa drôle et intime Année exemplaire, l' autrice marseillaise lance une structure d'édition plus favorable aux dessinateurs.
auteur	auteures	1	Diane Gistal mentionne des auteures afro-québécoises qui gagneraient à être connues : Stéphane Martelly ou Chloé Savoie-Bernard, par exemple.

3. Analyse des résultats

3.1. Analyse quantitative

Pour éviter tout biais a priori, nous nous sommes concentrés ici sur les formes féminines filtrées (sémantiquement non ambiguës, *cf.* Section 2.2) et étiquetées comme nom par le tagueur (*cf.* Section 2.3). Les comptes sont rapportés dans la table 3⁵. Sur base de ces comptes, nous avons calculé différentes valeurs de chi-carré afin d’évaluer la tendance à la féminisation en fonction des journaux et des années.

Tableau 3. Comptes pour les corpus *Le Devoir* et *Le Soir*

	2001			2021		
	TRUE	FALSE	OTHERS	TRUE	FALSE	OTHERS
Le Devoir	3061	20	13614895	5661	357	10146884
Le Soir	3713	4	21390483	3325	179	11325439

Pour 2001, le chi-carré ($\chi^2(2)=132,8$) montre un lien fort entre la féminisation et le journal. En effet, si l’on regarde les résidus standardisés, on constate un emploi nettement plus important de noms de métiers féminisés au sein du journal *Le Devoir* ($Z=10,6$ pour TRUE ; $Z=4,5$ pour FALSE). Cette différence s’accroît plus encore en 2021 ($\chi^2(2)=973,5$; $Z=29,9$ pour TRUE ; $Z=8,9$ pour FALSE). Par ailleurs, si l’on oublie les différences entre journaux (en fusionnant les comptes pour *Le Soir* et *Le Devoir*), on constate un lien très marqué entre l’année et la tendance à la féminisation ($\chi^2(2)=3227,1$). À nouveau, les résidus

⁵ Rappelons que TRUE et FALSE désignent respectivement les formes recommandées par Moreau et Dister (2014) et les formes construites artificiellement. OTHERS désigne le reste des mots.

standardisés permettent d'apprécier un emploi de la féminisation nettement plus marqué en 2021 ($Z=49,4$ pour TRUE ; $Z=28,1$ pour FALSE).

3.2. La catégorie des FALSE

Nous avons appelé "FALSE" les formes féminines non préconisées par le guide de 2014 qui nous sert de base. Nous manquons de place pour une analyse détaillée, mais l'essentiel de ces formes sont des mots en -eure (*acupuncteure, sculpteure, contrôleure, vérificateure...*) et en 2021 la majorité des items est représentée par le féminin *autrice*, forme qui était tombée en désuétude et n'était pas recommandée par les guides de féminisation. Ce terme, parce qu'il est un féminin ostentatoire, marqué aussi à l'oral, revient en force dans l'usage.

3.3. Le cas du lemme chercheur

Nous nous penchons dans cette section sur les formes du lemme chercheur au masculin et au féminin, uniquement au singulier (Table 3). En effet, les emplois du masculin pluriel sont très majoritairement utilisés pour neutraliser l'opposition de sexe. Dans un corpus de près de 7 heures de conversation, Branca et Dister (à paraître, 2022) relèvent que 98, 4% des emplois du masculin pluriel fonctionnent de cette manière. Les emplois au féminin pluriel sont quant à eux toujours exclusifs : ils ne permettent de désigner que des référents de sexe féminin, ce qui est le cas dans nos données.

Tableau 4. Occurrences des formes du lemme chercheur pour les corpus *Le Devoir* et *Le Soir* de 2001 et 2021

	<i>Le Devoir</i>				<i>Le Soir</i>			
	2001		2021		2001		2021	
	#	%	#	%	#	%	#	%
<i>chercheur</i>								
homme	340	69,7	376	51,6	291	78,9	224	55,6
femme (déterminant masculin)	3	0,6						
femme (déterminant féminin)							1	0,2
femme (apposition)	2	0,4			7	1,9		
générique	69	14,1	28	3,8	34	9,2	19	4,7
<i>chercheuse</i>	61	12,5	322	44,2	37	10	159	39,5
<i>chercheure</i>	13	2,7						
<i>chercheur et chercheuse</i>			3	0,4				

Qu'il s'agisse du *Devoir* ou du *Soir*, on constate que, entre 2001 et 2021, la proportion de *chercheur* pour désigner un homme diminue, tandis que celle de chercheuse augmente, ce qui semble correspondre au fait que le nombre de femmes chercheuses augmente et/ou que leur visibilité est plus importante. Cela va dans le sens de la tendance générale observée sur toutes les formes du corpus.

4. Conclusions

Notre étude n'avait pas pour ambition de comparer l'emploi du masculin et du féminin pour désigner des femmes, mais de voir l'implantation des termes au féminin dans un corpus de presse. De 2001 à 2021, on constate une évolution significative de l'utilisation des formes au féminin, tant en Belgique qu'au Québec. Les mots au féminin sont beaucoup plus nombreux dans les textes. Sans doute est-ce pour partie dû à l'évolution de la société en général, et des positions plus nombreuses qu'y occupent les femmes. Précurseur en la matière, le Québec faisait déjà plus de place aux femmes, à travers des mots au féminin, dans les données de 2001. Cette tendance générale s'observe dans l'analyse plus fine des emplois de *chercheur* et *chercheuse*.

Bibliographie

- Branca S. et Dister A. (à paraître, 2022). Le genre grammatical des noms désignant des référents humains. Ce que nous montre l'analyse de données orales non planifiées, *Description de l'oral et méthodes d'analyse en linguistique*.
- Courtois A. et Silberztein M. (1990). Dictionnaires électroniques du français. *Langue française*, 87 : 3-4.
- Dister A. et Moreau M.-L. (2021). Madame l'Administrateur, c'est presque fini... La dénomination des candidates lors des élections : étude diachronique. In Fagard B. et Le Tallec G. Éditeurs, *Entre masculin et féminin... Approche contrastive : français et langues romanes*. Paris, Presses de la Sorbonne Nouvelle : 31-54.
- Dister Anne and Moreau Marie-Louise (2009). *Féminiser ? Vraiment pas sorcier ! La féminisation des noms de métiers, fonctions, grades et titres*, De Boeck Dicolot, Bruxelles (coll. « Entre guillemets »).
- Fujimura I. (2005). Politique de la langue : la féminisation des noms de métiers et des titres dans la presse française (1988-2001), *Mots. Les langages du politique*, 78 : 37-52.
- Gross M. (1997). The Construction of Local Grammars. In Roche E. et Schabès Y. Éditeurs, *Finite-State Language Processing*. Cambridge, MIT Press : 329-354.
- Honnibal M. and Johnson M. (2015). An Improved Non-monotonic Transition System for Dependency Parsing, *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics*. Lisbon : 1373-1378.
- Moreau M. and Dister A. (2014). *Mettre au féminin : guide de féminisation des noms de métier, fonction, grade ou titre* (3^e édition). Fédération Wallonie-Bruxelles.
- Paumier S. (2021), Unitex 3.3. Manuel d'utilisation. URL : <https://unitexgramlab.org/releases/3.3/man/Unitex-GramLab-3.3-usermanual-fr.pdf>