

Les ‘femmes’ dans *Le Monde diplomatique* (1990-2008)

Les thèmes des femmes ou les ‘femmes’ dans les thèmes ?

Heba Metwally

Alexandria University – heba.metwally77@gmail.com

Abstract

This article reviews one of the most popular approaches in textual data analysis. The study of collocation and lexical ties supports the description of different text levels. While the description of the generalized collocation (Viprey 1997) and the thematic clusters (Reinert 1993) offer an insight into the macrostructures of the meaning, the co-occurrences of a certain word shows the lexical network related to this word. We sport here a gap between these two approaches operating on two different text levels. Therefore, we conduct a semantic investigation about the word ‘femme’ (in French) in a chronological corpus of *Le Monde diplomatique* (1990-2008). Our goal is to link thematic meso-text analysis to semantic macro-text analysis by confrontation of different lexical networks and multi-level contextualization.

Keywords: corpus semantics – thematic analysis – clustering – Reinert method – Tgen – chronological corpus – scale model corpus

Résumé

Cet article revient sur une des démarches les plus populaires en ADT. Les recherches thématiques mobilisent le calcul de la cooccurrence pour étudier deux différents niveaux de description textuelle. Tandis que la description des paradigmes thématiques s'appuie sur la cooccurrence généralisée (Viprey 1997) et les mondes lexicaux stabilisés (Reinert 1993) qui entérinent une sémantique interne visant par-là l'architectonique du sens, les recherches thématiques partant d'un mot-pôle explorent les réseaux lexicaux que le calcul de la cooccurrence juge significatifs et qui infèrent sur les univers sémantiques qui abritent le mot-pôle. Et entre les deux démarches un hiatus sépare les caractéristiques thématiques locales des macrostructures sémantiques. Pour cela, on mène une investigation sémantique à partir le lemme ‘femme’ dans une Série Textuelle Chronologique du *Monde diplomatique* (1990-2008). Par la contextualisation des faits textuels significatifs et la confrontation de différents niveaux d’analyse, nous appelons de nos vœux une sémantique de thèmes dûment balisée.

Mots clés : sémantique textuelle – sémantique des thèmes – mondes lexicaux stabilisés – corpus maquette – Tgen – Série Textuelle Chronologique

1. Introduction

« Si l’on nomme thème une structure stable de traits sémantiques (sèmes) récurrente dans un corpus et susceptible de lexicalisations diverses, cette définition doit être spécifiée, en caractérisant ces traits, ce corpus et ces lexicalisations, puis en posant les questions de l’interaction des thèmes entre eux et la constitution des paradigmes thématiques. » (Rastier, 2001, p. 197). Dans une sémantique de corpus (Rastier, 2011), l’étude d’un thème passant matériellement par ses lexicalisations s’inscrit inéluctablement dans l’étude du corpus et de ses paradigmes thématiques. Si la cooccurrence est l’objectivation minimale du thème, elle l’est de par sa nature contextuelle qui jalonne la démarche interprétative. Celle-ci dicte la reconstitution des liens de *contextualité* pour définir le sens (Rastier, 2001). Ainsi, le rapprochement entre les passages assemblés à partir du calcul des cooccurrences d’une lexie (*contextualité homoplane*) gagne à être complété par leur contextualisation dans les paradigmes thématiques sur lesquels se détachent les thèmes (*contextualité hétéroplane*).

Cette dernière procédure relativise les observations locales. Elle vise, dans un premier temps, à mettre en évidence le rapport entre les logiques discursives globales et les logiques thématiques ponctuelles, et dans un second temps, à étudier les conditions de leur convergence et les circonstances de leur autonomisation.

Dans cette optique, on mène une investigation textuelle sur le lemme 'femme' dans une Série Textuelle Chronologique du mensuel français *Le Monde diplomatique* (1990-2008). Le corpus compte plus de 5000 articles, plus de 11 millions d'occurrences. Il est reparti selon une périodisation endogène en quatre sous-corpus de stabilité sémantique¹. Dans ces quatre temps sémantiques, les analyses statistiques multi-niveaux se focalisent moins sur l'évolution du discours – tout en y prenant appui – que sur la place du lemme 'femme' dans les paradigmes thématiques et sur l'hétérogénéité des régimes sémantiques qui actualisent ce lemme au cours du temps. L'objectivation et la description des macrostructures sémantiques se fait à partir les mondes lexicaux stabilisés de la méthode Alceste (Reinert, 1983, 1993) implémentée dans le logiciel libre Iramuteq (Ratinaud et Marchand, 2012) et qui s'appuient sur les substantifs les plus fréquents du corpus. Les attractions lexicales du lemme 'femme' seront envisagées tantôt à partir de la population substantive importante du corpus, tantôt – sans discrimination – à partir de l'ensemble de la population statistique. Et on aura recours aux fonctionnalités d'Hyperbase mobilisant le calcul de la cooccurrence (Brunet, 1995, 2008, 2012).

2. Les corpus de travail et l'importance de 'femme'

La Série Textuelle Chronologique donne lieu à quatre sous-corpus de travail qui correspondent aux quatre périodes de stabilité sémantique. Ils sont partitionnés par année ; et on distingue pour chaque corpus deux versions : une version intégrale qui comprend les textes dans leur intégralité lexicale ; une deuxième version abrégée de la première qui tout en conservant la structure sérielle du corpus et séquentielle des textes ne laisse voir que ses substantifs les plus fréquents, le reste de la matière lexicale étant provisoirement mis à l'ombre. Il s'agit simplement d'une version comprimée du corpus, d'une représentation à échelle réduite du corpus, i.e. d'une *maquette*. On l'appelle *maquette* pour les raisons formelles qu'on vient d'évoquer et pour des raisons praxéologiques qu'on précisera tout de suite. Pour les architectes, les *maquettes* constituent un champ d'étude pour la conception et la validation d'un édifice en projet dans ses détails. Au-delà de la phase de la conception, et une fois l'édifice exécuté, une *maquette* donne lieu à d'autres applications. Dans cette dernière perspective, face au corpus textuel, qui est un construit, un regard rétrospectif permet à travers une *maquette* d'abord d'identifier le fond sémantique qui sous-tend la construction et ensuite de faire l'interprétation formelle des principes de son fonctionnement. Cette démarche vise à la mise en exergue de la substance du discours à partir des briques nominales² les plus importantes qui, dans un premier temps, permettent un meilleur accès aux corpus et à leurs macrostructures sémantiques et, dans un second temps, autorisent les recherches thématiques

¹ La stabilité sémantique se laisse mesurer à partir de la cooccurrence devenue unité de décompte généralisée (Brunet, 2012). La distance intertextuelle calculée sur la cooccurrence donne à voir quatre périodes : 1990-1993 ; 1994-1997 ; 1998-2001 ; 2002-2008. Ces périodes correspondent à quatre sous-corpus qui constituent nos corpus de travail.

² Le choix d'une seule catégorie grammaticale nous permet de cibler des espaces contraints où les relations syntaxiques restent minimales.

à partir de la même matière lexicale pour étudier les thèmes dans le contexte de leurs classes sémantiques.

L'accroissement du lemme 'femme' au fil du temps est évident. Les fréquences relatives à la taille des corpus du lemme confirment une progression définitive du lemme et un pic d'afflux dans la troisième période (Figure 1)³. Convenons que l'indice fréquentiel de l'occurrence est un indice unique, un chiffre synthétique qui renvoie aux effectifs d'une lexie qui s'investit dans différents régimes textuels qui s'accumulent au fil des textes. Dans les lignes suivantes, on va analyser les contextes importants de l'emploi du lemme 'femme' pour évaluer le rapport entre l'importance quantitative d'une lexie dans un corpus et son importance qualitative dans la structure sémantique du corpus et les champs sémantiques qui l'activent ou qu'il active.

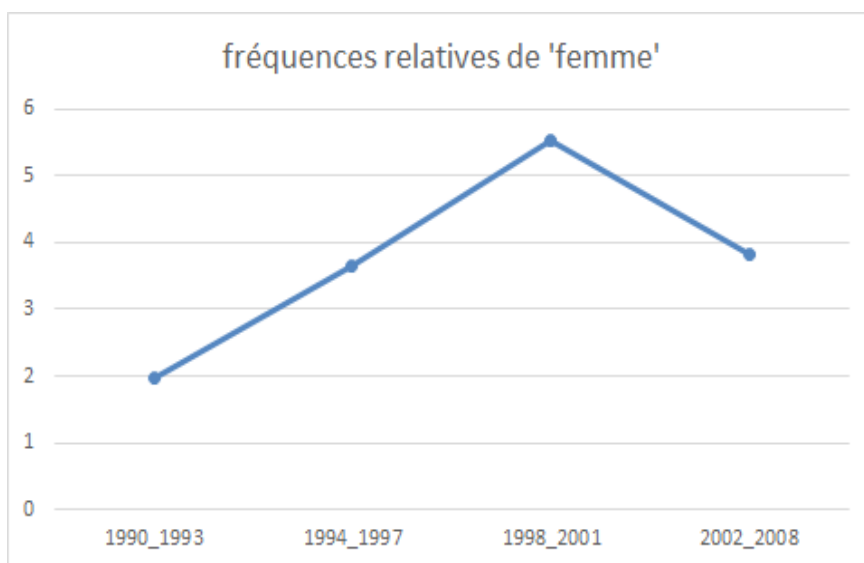


Figure 1 : Les fréquences relatives à la taille des sous-corpus de 'femme'⁴

3. 'Femme' et les macrostructures sémantiques

La mise en évidence des macrostructures sémantiques des quatre sous-corpus se réalise à partir de leurs *maquettes* respectives qui sont traitées par la classification hiérarchique descendante (CHD) de la méthode Reinert. En peu de mots, la méthode consiste en une analyse lexicale par contexte. Plus précisément, les textes sont découpés en segments pour délimiter les contextes d'étude. La CHD compare les profils cooccurentiels de ces contextes pour repérer les traces de contenu et aboutir à un classement qui distingue et hiérarchise les mondes lexicaux. Ceux-ci correspondent à l'issue du calcul à des classes lexicales caractérisées chacune par une *logique locale* propre (Reinert 1993). Ainsi, les *maquettes* sont coupées en des segments de textes (ST) de 10 substantifs⁵. L'algorithme suit une démarche itérative de bipartitions qui part d'un tableau rectangulaire croisant les ST et les formes

³ Voir aussi Annexe 1 pour une représentation linéaire et topologique de cette version abrégée. Voir également Annexe 2 pour une synthèse de la taille des corpus, de leurs *maquettes*, des effectifs des substantifs des *maquettes* et de l'importance du lemme 'femme' dans les corpus et dans les listes hiérarchiques des substantifs au fil du temps.

⁴ Les fréquences relatives ont été multipliées par 10 000 pour la commodité de lecture.

lexicales choisies. L'algorithme poursuit la bipartition à partir de la plus grande classe disponible et s'arrête quand le nombre de classes demandées est obtenu. Après traitement, on obtient dans les quatre *maquettes* 7 classes terminales. Le tableau ci-dessous rend compte de l'importance des contenus classés⁶.

	1990-1993	1994-1997	1998-2001	2002-2008
Nombre des ST	26540	15023	22657	37767
Nombre des ST classés	17942	9875	15143	22720
Pourcentage des ST classés	67.60%	65.73%	66.84%	60.16%
Pourcentage des thèmes politiques dans les ST classés ; nombre de classes	56,57% ; 4	45,46% ; 3	52,02% ; 4	62,54% ; 4
Pourcentage des thèmes économiques dans les ST classés ; nombre de classes	31,01% ; 2	41,67% ; 3	30,82% ; 2	12,55% ; 1
Pourcentage des thèmes socio-culturels dans les ST classés ; nombre de classes	12,42% ; 1	12,86% ; 1	17,15% ; 1	23,90% ; 2

Tableau 1 : Description des 4 CHD

Du point de vue qualitatif, trois domaines sémantiques se différencient dans chaque période. Les contenus politiques sont systématiquement les plus importants, viennent après les thèmes économiques et socio-culturels. Tandis que les contenus économiques diminuent avec le temps, les contenus socio-historiques augmentent sans atteindre une importance symétrique. Cette modification est aussi quantitative que qualitative. Les contenus économiques ne connaissent dans la dernière période qu'une seule classe par opposition aux thèmes socio-culturels qui se regroupent pour la première fois en deux classes.

D'une classification à l'autre, le pouvoir caractérisant du lemme 'femme' dans les mondes lexicaux stabilisés change quantitativement et qualitativement. On retient les indices de spécificités positives du lemme dans les mondes lexicaux stabilisés des quatre périodes pour avoir un avant-goût de la couleur sémantique des passages de 'femme' (ci-dessous Figure 2). D'une manière générale, 'femme' lexicalise de préférence les thématiques économiques et socio-culturelles. L'indice de spécificité (en χ^2) dit l'importance du mot dans les ST classés de tel ou tel groupement thématique. Il est l'expression statistique de la centralité de 'femme' dans les réseaux lexicaux de la classe sémantique et mesure sa capacité à configurer ses réalités si diverses qu'elles soient.

⁵ On a estimé la succession de 10 substantifs dans le corpus-maquette comme l'équivalent dans le corpus intégral de la fenêtre contextuelle de 33 mots. On vise par là un espace intermédiaire entre la phrase et le paragraphe.

⁶ Pour une description des classes thématiques voir (Metwally, 2017) et (Metwally, 2018) pour une synthèse rapide.

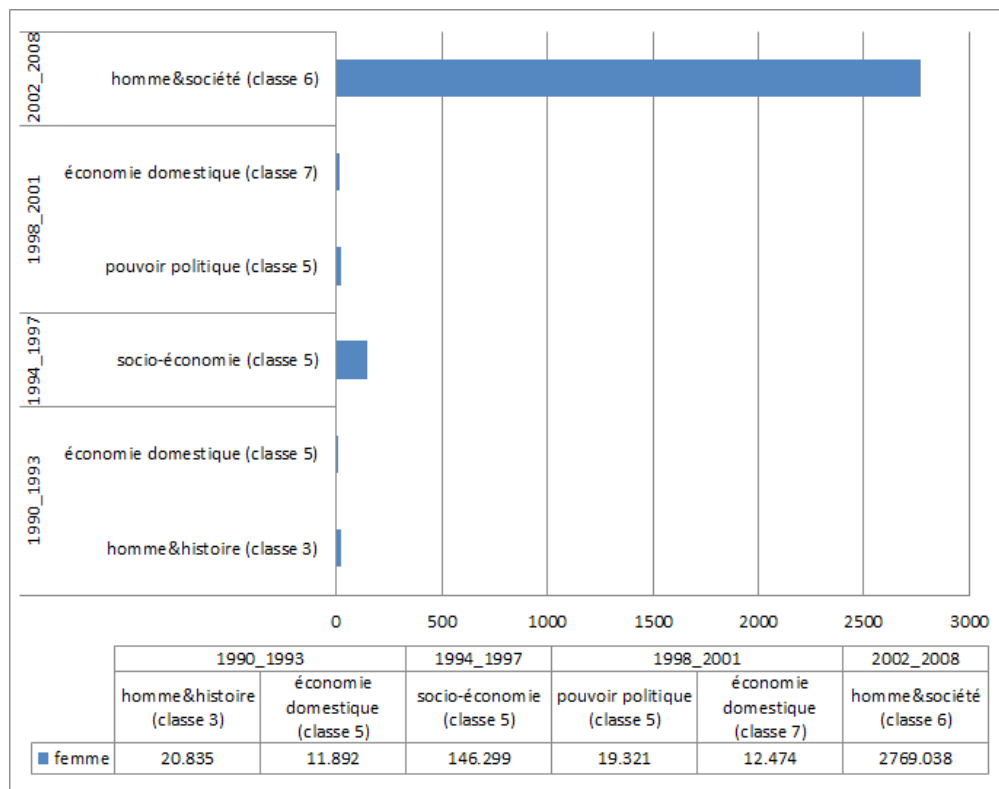


Figure 2 : 'Femme' dans les mondes lexicaux stabilisés (indices de spécificités en χ^2)

Pour identifier la place de 'femme' dans la logique sémantique des textes, dans les *logiques locales* des mondes lexicaux, on calcule les associations privilégiées de 'femme' au sein des mots de la *maquette*, traces lexicales spécifiques des mondes lexicaux. On obtient ainsi les passages que caractérisent à la fois 'femme' et les traces lexicales des mondes lexicaux. Au cœur des textes, au fil des périodes et des années, on observe les associations privilégiées de 'femme' et leur chronologie. Pour la confrontation entre les passages de 'femme' (niveau méso) et les classes sémantiques (niveau macro), on compare les contenus nominaux des passages avec les contenus spécifiques des mondes lexicaux stabilisés (i) par une simple lecture qualitative ; (ii) ou en utilisant les TGenS (Lamelle et Salem, 2002) pour la mesure de la proximité sémantique entre les passages d'un côté et les classes sémantiques de l'autre côté.

3.1 Les thèmes économiques

Les traces lexicales des thèmes économiques font systématiquement partie de l'environnement lexical de 'femme' qu'on calcule les cooccurents les plus proches à partir l'ensemble de la population statistique ou qu'en envisageant exclusivement les substantifs de la *maquette*. La consultation des listes des cooccurents les plus proches aux deux niveaux nous donne à voir : 'travail', 'travailler', 'emploi', 'salaire', 'chômage', 'temps' (voir Annexes 3, 4). Ce sont des membres permanents du réseau lexical de 'femme'. Au fil du temps, 'femme' continue à lexicaliser les thèmes économiques. En Afrique, les femmes sont les premières victimes des discriminations et des pratiques abusives en espace de travail :⁷

⁷ Sont mises en rouge les traces lexicales spécifiques des classes économiques de la classification correspondante.

« Des *femmes*, bêtes de somme sous-payées, journalières mal logées, mal nourries, travaillant sur des plantations de thé ou de café possédées par des *étrangers*. Le décalage est énorme, dans les cultures de rente, entre le *travail* fourni par des *femmes* et les *salaires* qu'elles perçoivent. Pourtant, c'est avec leur sueur que se gagnent les devises confortant les *économies* nationales et les superprofits des multinationales. Les *femmes* travaillent, les hommes possèdent la *terre* et contrôlent l'*argent* » (Mai 1993)

La discrimination contre les femmes s'inscrit dans une tendance opportuniste générale de l'exploitation des faibles par les plus forts. Et la déferlante de l'économisme continue à balayer les opportunités d'accès à une vie digne :

« Le *rapport* provisoire des Nations unies laisse aussi entrevoir que les *pratiques* liées à la mondialisation se sont appuyées sur une idéologie patriarcale qui existait précédemment mais que la mondialisation a su intégrer, quand elle ne l'a pas renforcée. Ainsi, trois des phénomènes liés à la mondialisation – la multiplication des *sweat shops* (ateliers de sous-traitance), des *emplois* à temps partiel et des *formes* de *travail* les plus précaires – frappent avant tout les *femmes*, notamment celles du *Sud* et les immigrés. » (Juin 2000)

A partir des années 1980, la marchandisation du travail, la précarisation de l'emploi et la détérioration des conditions de travail continuent à fragiliser la situation des femmes sur le marché de travail en Europe. Ceci suscite de vifs débats dans l'espace européen et dans le discours public en France. L'année 1997 connaît un afflux d'articles sur l'emploi féminin. Ceci se traduit topologiquement sur la chaîne textuelle par une *rafale*⁸ (Brunet 2006 ; voir Annexe 1). Dans le numéro de mars 1997, un dossier intitulé 'Les bastilles toujours dressées de l'inégalité' décline les questions féminines d'actualité en quatre articles. L'article consacré à l'emploi féminin, reprend le développement en Europe et en France du temps partiel au début des années 80 face à la montée du chômage. L'emploi féminin en est fortement affecté :

« Quand les *femmes* sont entrées sur le *marché* du *travail* au *début* des *années* 60, c'était à *temps* plein pour la grande majorité d'entre elles. Le *temps* partiel s'est développé chez nous non parce que les *femmes* l'ont choisi, mais parce que de nombreux *emplois* de ce *type* se sont créés dans les *secteurs* fortement féministes comme le *commerce*, les grandes surfaces, l'hôtellerie-restauration et les *services* aux *particuliers*. *Temps* partiel en *France* rime donc bien souvent avec professions sous-qualifiées et *secteurs* majoritairement féminin. » (Mars 1997)

Les trois passages précédents relèvent de deux logiques sémantiques. Dans les deux premiers passages, le travail de la femme est un argument dans la critique de la mondialisation et des injustices qu'elle engendre ou favorise. Le travail de la femme est un opérateur descriptif qui s'inscrit dans la critique de l'économisme, c'est un sous-thème du paradigme thématique de l'économie du marché condamnée systématiquement dans les pages du mensuel. A cette logique de sous-détermination sémantique s'ajoutent les thèmes de l'emploi féminin qui montent à la surface grâce à un *moment discursif* (Moirand, 2007) pour s'aligner parfaitement sur la *logique locale* de la classe consacrée aux thèmes de l'emploi et du travail de la deuxième période. D'où la place importante qu'occupe 'femme' dans les ST du monde lexical socio-économique de la deuxième période (Figure 2). Notons que ce *moment discursif* sera suivi par d'autres dans la suite de la Série Textuelle Chronologique, tous traitant de thèmes

⁸ Il s'agit de la représentation graphique de la fonction 'Topologie' dans Hyperbase. C'est « une concentration des occurrences due à quelque cause locale, thématique ou stylistique » (Brunet, 2006).

féministes que ce soit du point de vue économique ou du point de vue politique (Voir tout de suite infra 3.2)

3.2 Les thèmes politiques

Les thèmes politiques sont les plus importants du discours au cours du temps et les plus productifs au niveau des classifications (voir Tableau 1). Leurs classes mettent en scène les questions principales du pouvoir national et international : la circulation du pouvoir, l'instabilité politique, les guerres civiles et les conflits internationaux. Dans la troisième période, 'femme' est une trace lexicale des thèmes politiques (classe 5) avec un indice de spécificité faible (chi 2 : 19,3 ; rang : 50). La faible intégration des réseaux de 'femme' dans les thèmes politiques reste un reflet signifiant de la couleur sémantique des thèmes des femmes dans cette période. On a déjà identifié en 1997 un dossier de 4 articles consacrés aux thèmes de la parité (voir supra 3.1) ; ceux-ci se spécialisent dans le domaine politique en 1998 avec la révision par le Parlement des articles 3 et 4 de la Constitution relatifs à la parité hommes-femmes dans le domaine politique et qui était prévue pour la fin de l'année 1998 :⁹

*« Le projet de révision constitutionnelle sur la parité entre **hommes** et **femmes** sera discuté par le **Parlement** français au début du **mois** de **décembre**. Si cette revendication soulève des questions de **principe**, elle est de plus en plus partagée par les féministes du monde entier. Malgré quelques évolutions, le **pouvoir** politique demeure en effet dominé de manière écrasante par les **hommes**, qui constituent 90% de la représentation nationale. »*
(Novembre 1998)

A l'occasion de cet événement constitutionnel, les débats se multiplient dans le discours public autour d'une diversité de questions sur la parité. Dans les deux dernières périodes, certains cooccurents qui configurent ces questions s'approchent davantage de 'femme'. Citons, à titre d'exemple, 'homme', 'droit', 'travail', 'journaliste', 'syndicat' et 'mouvement'. (Annexes 3, 4). Ceci coïncide avec une seconde *rafale* de 'femme' (Annexe 1), celles-ci emportant une grappe de cooccurents qui déclinent les différents domaines de la parité. Et l'expression ponctuelle et sporadique de leur distribution est évidente sur le graphique ci-dessous (Fig. 3).

⁹ Sont mises en rouge les traces lexicales spécifiques de la classe 5 (pouvoir politique) de la troisième période.

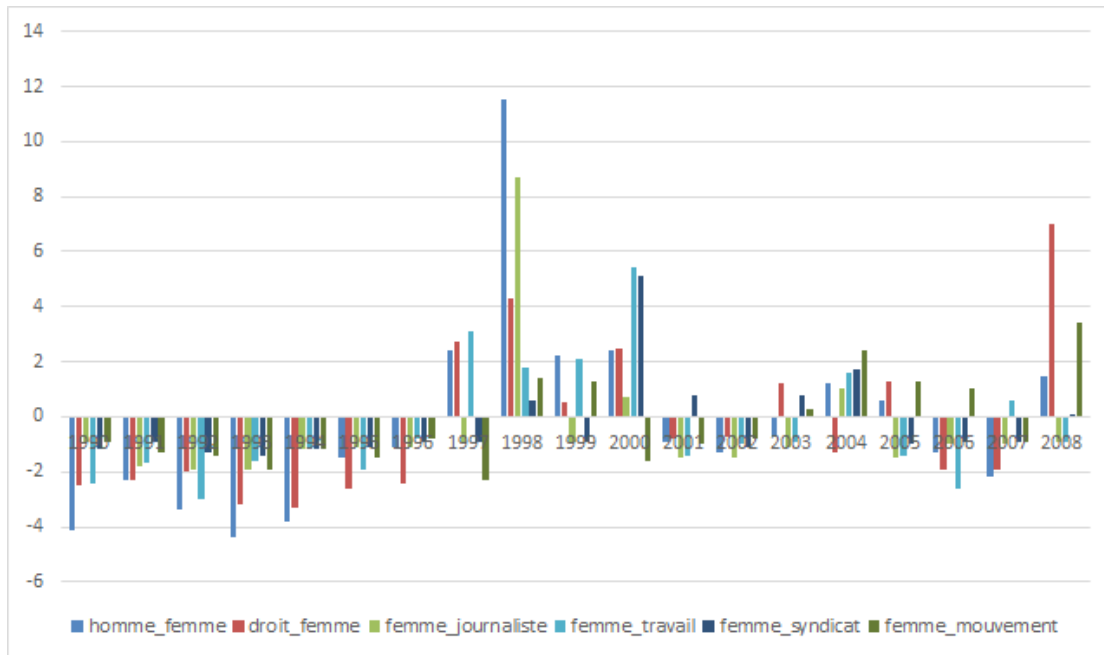


Figure 3 : Distribution de cooccurents de « femme » relatifs aux questions féministes

Il s'agit de 6 couples qui se combinent différemment à partir de l'année 1997. En 1997, c'est l'emploi féminin (femme-homme-droit-travail). En 1998, avec 'journaliste' qui s'ajoute à la cette communauté, il est question de « *l'accès égal des femmes et des hommes aux mandats et aux fonctions, que devrait préconiser la Constitution française [et qui] reste à gagner dans les médias comme en politique.* » ('Le deuxième sexe du journalisme'¹⁰, Juillet 1998). En 2000, dans le prolongement des thèmes de la parité, la représentation des femmes dans les syndicats laisse à désirer : « *Chacun sait qu'il ne suffit pas d'avoir une femme à la tête d'un syndicat pour que la place de la gent féminine dans l'entreprise, le service ou le syndicat en soit bouleversée.* » ('D'autres bastilles à abattre. Les syndicats n'ont pas la fibre féministe', Juin 2000). En 2008, le combat féministe culmine avec les deux binômes 'femme_droit' et 'femme_mouvement'. Dans le numéro de novembre 2008, une référence intense à Olympe de Gouges est évidente ; un dossier de quatre articles est intitulé 'Olympe de Gouges, une femme du XXI^e siècle'. On y lit : « *Auteure de la Déclaration des droits de la femme et de citoyenne, elle est devenue une figure emblématique des mouvements pour la libération des femmes.* ». Il y est question de pouvoir ce qui justifie l'alignement des thèmes de la parité et du combat féministe sur les thèmes politiques qu'ils soient nationaux ou internationaux.

3.3 Les thèmes socio-culturels

Au fil du temps, les lemmes 'enfant' et 'homme' restent les cooccurents les plus proches de 'femme' quelle que soit la population statistique analysée. Les trois lexies configurent l'humanité et sont activées par des contenus qui font le bilan de la condition humaine. La guerre est une cause principale de la souffrance humaine. La guerre du Golfe en 1991 en est un exemple qui multiplie pertes et dégâts humains :

¹⁰ Cet article contient à lui seul 11 cooccurrences du binôme 'femme_journaliste' ; ce score est redoublé lorsque le même article est republié le mois suivant.

« J'ai rencontré plusieurs des victimes civiles dans les hôpitaux de Bagdad : *hommes, femmes et enfants* souffrant de terribles blessures infligées par des tirs et des éclats de fusées et d'obus. Version officielle : c'était l'œuvre de saboteurs. Les victimes ne purent que me donner leurs noms et ceux de leur ville d'origine. » (Juin 1991)

Face à la lutte armée en Amérique latine qui persiste depuis les années 60 entre les guérillas et les régimes militaires, les innocents paient le prix :

« Au Pérou, un nombre incalculable d'*hommes, de femmes et d'enfants* vivent depuis une dizaine d'années dans la terreur, après qu'eut été instauré l'état d'urgence dans le cadre de la lutte anti-insurrectionnelle. Dans ce rapport daté de novembre 1991, Amnesty international montre comment les forces gouvernementales bafouent le droit international, et dénonce les exactions commises par le Sentier Lumineux. » (Février 1992)

Dans les années 90, l'Algérie connaît des cycles de violence qui enfoncent le pays dans une guerre civile. La population algérienne est minée par le terrorisme et la pauvreté :

« En Algérie, 28 millions de *femmes, d'hommes et d'enfants* voient avec terreur surgir la mort au quotidien, mais c'est avec une terreur égale que l'écrasante majorité de la population voit refuser le droit à la décence la plus élémentaire. Au détour des grands artères de la capitale, des Algériens vivent par milliers – *hommes, femmes et jeunes* – sous des tentes sur des terrains insalubres : d'autres milliers d'Algériens ont, depuis des années, squatté des conteneurs – ayant servi à transporter les marchandises dont le commerce enrichit l'oligarchie au pouvoir – et s'y entassent par familles entières. » (Février 1998)

Les trois passages caractérisés par les trois lemmes en question comportent des sèmes de la souffrance matérielle et morale. Pour une évaluation statistique de la couleur sémantique des passages qui abritent ce trio lexical dans les trois premières périodes, on projette une liste de ces trois lemmes regroupés dans un TGen successivement sur les classes de leurs classifications respectives¹¹. On ne retient ci-dessous que les indices positifs (Figure 4).

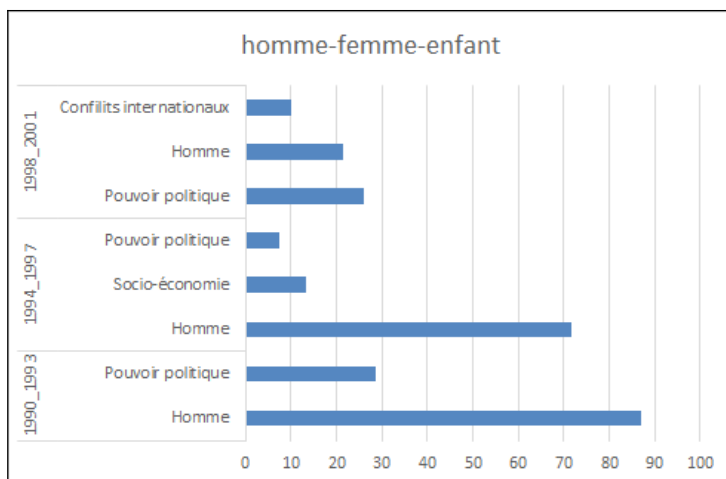


Figure 4 : Projection du TGen (homme-femme-enfant) sur les classes des trois premières périodes

¹¹ On exclut de la comparaison la quatrième période qui sera abordée à part vu l'indice de spécificité trop important du lemme 'femme' dans la classe 6 de la classification (voir Figure 2).

Dans cette analyse, on identifie l'origine sémantique des contenus des passages qui traitent de la misère matérielle de l'Homme, de la précarité de ses moyens de survie et de sa situation de première victime des guerres civiles, de l'instabilité politique nationale et internationale ce qui renoue avec nos observations globales des thèmes socio-historiques dans le mensuel¹².

Dans la dernière période, 'femme' et 'enfant' figurent pour la première fois en tête des traces lexicales les plus spécifiques d'une classe centrée sur les thèmes sociaux (Figure 2, classe 6). Cette classe n'a pas d'équivalent dans les périodes précédentes. De la condition de l'Homme dans le monde et dans l'Histoire on passe ici aux sorts des individus dans telle ou telle société. Le cadre se déplace dans la ville, dans le quartier, voire dans la maison. Le passage suivant correspond à l'un des ST les plus spécifiques de la classe en question. A l'occasion des débats parlementaires autour de la loi sur le port des signes religieux dans l'établissement scolaire public en France, le regard se dirige vers les cités et les facteurs favorisant l'isolement de la population maghrébine. La situation dans des quartiers de plus en plus stigmatisés par cette frange de la société française se laisse décrire comme suit :¹³

« Le *quartier* étudié est devenu, comme tant d'autres, l'habitat des laissés-pour-compte de la mobilité résidentielle : *familles* immigrées comptant de nombreux *enfants*, retraités et *femmes* seules avec *enfants*, *salariés* pauvres, etc. Le *nombre* des *familles* immigrées y est plus élevé que dans les autres *quartiers* de la *ville*. En outre, les *jeunes* issus de ces *familles* y sont particulièrement visibles en raison de leur poids démographique dans la *population* juvénile totale. Et, de fait, ils polarisent une attention inquiète. Les fils d'immigrés privés d'*emploi* ont pris peu à peu possession des *lieux*, imposant leurs usages aux autres *habitants*, qui n'ont plus la force sociale suffisante pour les contrôler. » (Janvier 2004)

La singularité de la classe 6 dans cette dernière période réside dans le fait de mêler des traces lexicales sociales, économiques, urbaines et de resserrer la focale pour accorder une place importante à la vie privée ce qui favorise la montée à la surface du thème de la violence domestique objectivé par le cooccurrent 'femme_violence'. Ce cooccurrent gagne en importance notamment dans la deuxième moitié ce que montre sa distribution dans le corpus (ci-dessous Figure 5). L'indice de corrélation, en l'occurrence celui de Bravais-Pearson, confirme une progression significative (+0.6). D'une période à l'autre, on peut observer également au niveau de la *proxémie lexicale*, dans le tableau synthétique de l'environnement lexical de 'femme', l'ascension du lemme 'violence' dans la liste des cooccurrents les plus proches du mot-pôle (Annexes 2, 3). Dans la première période, celle du sous-emploi, le lemme 'violence' est absent de l'environnement lexical de 'femme'.

¹² Voir (Metwally, 2017)

¹³ Sont mises en rouge les traces lexicales spécifiques de la classe 6 de la quatrième période. Précisons qu'un ST ne dépasse pas 10 substantifs, ceux qui sont issus de la liste des substantifs les plus fréquents. Pourtant, pour une contextualisation raisonnable, on se permet de lire au-delà des limites du ST.

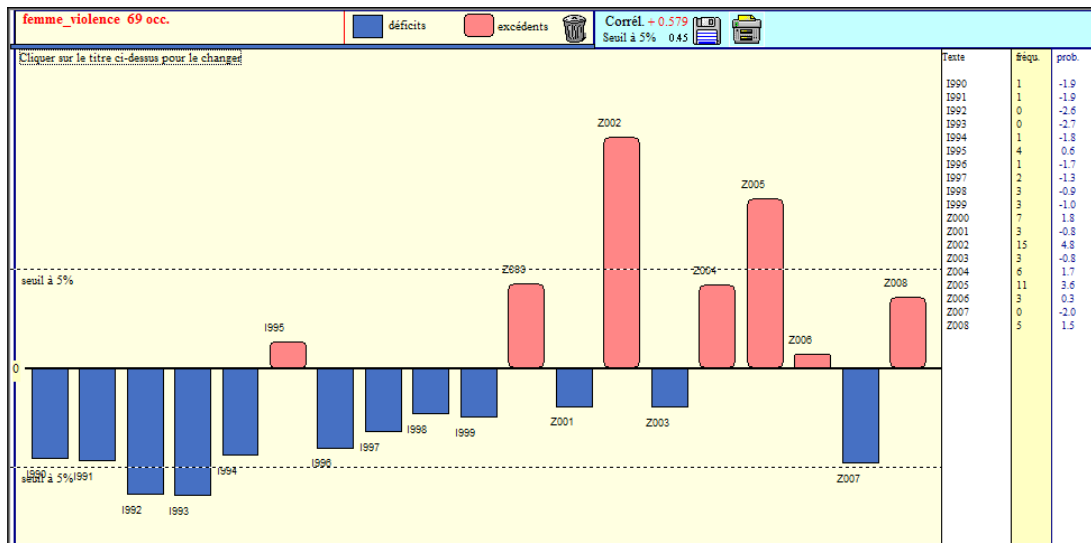


Figure 5 : Distribution du cooccurrent 'femme_violence' dans le corpus

Dans la première période dominée par les bouleversements mondiaux et les conflits internationaux, le thème de la violence contre la femme n'est pas encore à la mode. Avec la croissance progressive de la place textuelle réservée aux femmes dans le mensuel, le cooccurrent 'femme_violence' passe dans la zone du suremploi. Dans la dernière période, le mot 'violence' est un cooccurrent très proche du lemme 'femme'. En 2002 et en 2005, un débordement du cooccurrent 'femme_violence' traduit l'importance de la violence domestique dans le discours social :

« Si beaucoup de *femmes* sont démunies et souvent sans *travail*, le drame touche absolument toutes les catégories sociales. En témoigne l'*Enquête* nationale sur les *violences* envers les *femmes* en *France* (Enveff), qui a fait l'effet d'une petite bombe lors de sa publication en juin 2001 : une *femme* sur dix est victime de *violences* conjugales, et six *femmes* meurent chaque mois. Les *violences* conjugales sont une des causes principales de mort par homicide des *femmes*, souligne Mme Maryse Jaspard, qui a coordonné cette *enquête*. » (Juillet 2002)

Dans la dernière période, la structure sémantique du sous-corpus offre les conditions textuelles idéales pour la stabilisation des thèmes des femmes dans la classe des thèmes sociaux. La *logique locale* de la classe 6 assimile facilement les thèmes du privé. D'où l'harmonie entre les thèmes des femmes et les macrostructures sémantiques.

4. Conclusions

Le lemme 'femme' dans les passages du *Monde diplomatique* renvoie aux femmes dans le monde et lexicalise les contenus sémantiques importants du discours en marquant les *moments discursifs* alimentés par les thèmes des femmes. Les passages étudiés constituent un réquisitoire contre des régimes politiques, économiques et sociales qui réduisent les femmes au statut de victime. Dans les pages du *Monde diplomatique*, les femmes sont les premières victimes de la tyrannie de l'économisme et l'enfer des guerres. Leur vulnérabilité les expose dans le privé à la violence domestique. Ainsi à travers le lemme 'femme' est mis en texte un

rapport constant de domination (économique, politique, culturelle, éthique) par le mensuel. Dans cette optique, les *moments discursifs* cristallisent l'autre versant de la domination, à savoir la lutte. La lutte pour la parité dans les domaines publics et privé. La visibilité dans les pages du mensuel des mouvements féministes s'inscrit dans le soutien qu'accorde le mensuel à tous ceux qui sont à contre-courant. Olympe de Gouges, à la fois, figure féministe française et symbole national de la Révolution, incarne la fusion entre le combat féministe et les mouvements contestataires. La critique de tous les systèmes de domination est ce leitmotiv qui sous-tend les contenus des passages de 'femme' quelle que soit leur place dans les macrostructures sémantiques et qui fonde l'harmonie entre les logiques discursives globales et les logiques thématiques ponctuelles.

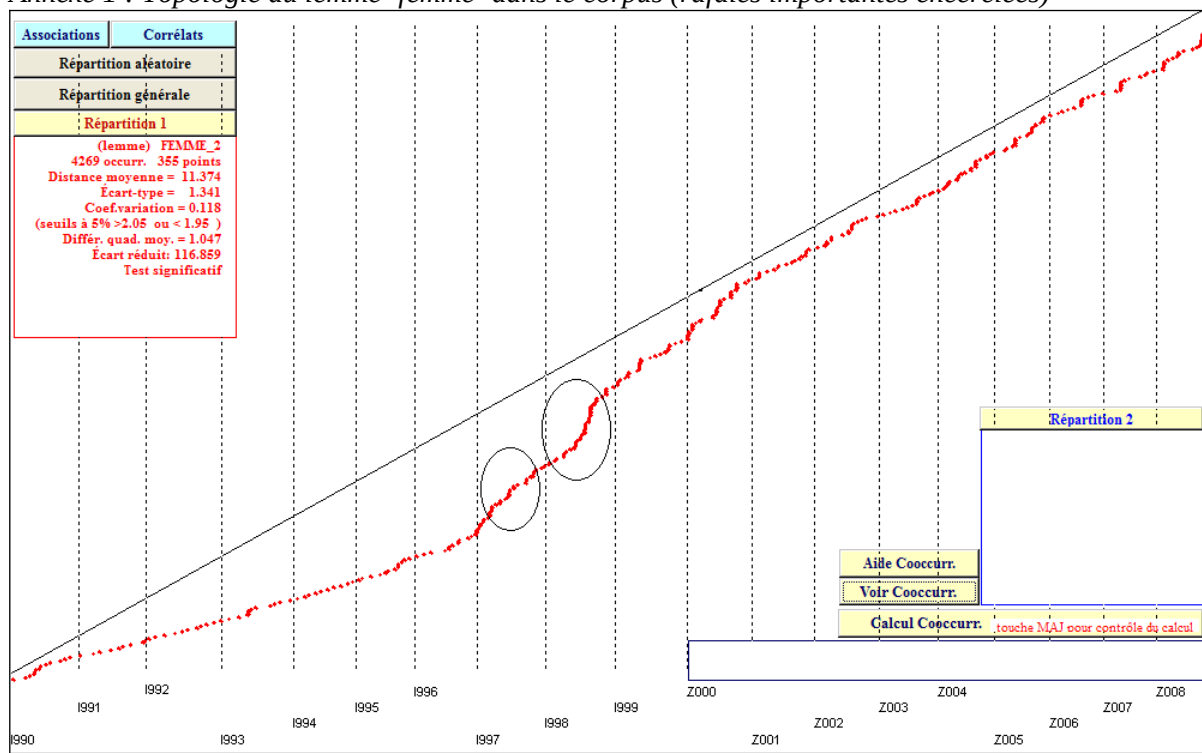
Le rapport entre les mots et les thèmes n'est nullement transparent, mais le mot reste notre point d'entrée dans les réseaux lexicaux, matière première des thèmes. L'investigation textuelle qu'on a menée dans les lignes précédentes s'appuie principalement sur la *maquette*, un modèle réduit du corpus intégral qui calque la structure sérielle du corpus, séquentielle de ses textes, qui remanie la partition sérielle pour l'adapter au temps sémantique et met en évidence leur structure non-séquentielle, i.e. la texture, en poursuivant sa voierie lexicale. Dans ce sens, la *maquette* se propose comme une infrastructure qui facilite la circulation entre les niveaux textuels méso et macro du corpus. On juge que la *maquette* est un dispositif nécessaire à une étude thématique qui contextualise les traits locaux dans les faits globaux. La stratégie principale qu'on a employée pour lier les caractéristiques locales (niveau méso) aux classes sémantiques (niveau macro) réside dans le recours à la *maquette* et pour les mésoanalyses (les thèmes objectivés par les réseaux lexicaux significatifs de 'femme') et pour les macroanalyses (les classes sémantiques). La comparaison en est le résultat immédiat. De ce fait, la *maquette* est dans la praxis un trait d'union entre les macrostructures et les mésostructures sémantiques.

Les Séries Textuelles Chronologiques appellent naturellement les réflexions sur le temps lexical (Salem, 1988) et l'évolution du discours. L'étude des régimes textuels de 'femme', de sa topologie simple dans le continuum textuel, de sa topographie complexe au sein de ses réseaux, par rapport aux classes sémantiques et dans la rhétorique du discours, confirme et renseigne sur l'historicité des mots dans le discours. Chaque mot a une histoire dans le corpus, une histoire quantitative résultant de sa différenciation qualitative et qui se caractérise et se réaffirme au fil des textes. Un mot est doté d'une identité discursive qui se définit par ses connivences lexicales au niveau du passage, sa place par rapport à la structure sémantique du corpus et son emplacement dans la chronologie du corpus. Et, du coup, au fil du temps, au fil du texte, le mot ne circule pas sans bagage. Il a une identité sémantique faite de ses tendances isotopiques et isotropiques (Viprey 2005) et une identité numérique que dit son importance au sein de la communauté lexicale du corpus et se reflète topologiquement par son accroissement au fil du texte. C'est cette identité complexe qui, au fil du temps, est en jeu pour s'accommoder aux aléas du temps. Les modifications quantitatives et qualitatives que parcourt un mot dans sa circulation sur la chaîne textuelle ressemble aux impressions que produit un kaléidoscope. Le stock lexical change peu d'une période à l'autre car il reste le propre du genre ; ce qui change ce sont les tris lexicaux et les relations hiérarchiques de détermination sémantique au sein des réseaux lexicaux qui reproduisent périodiquement de nouvelles combinaisons. Ceci dit, le temps lexical n'est que l'expression quantitative finale d'un mouvement plus profond, celui des classes sémantiques, des leurs relations intra-classes et des relations inter-classes, et qui matérialisent l'évolution du discours au niveau des corpus textuels.

Références

- Brunet E. (1995). *Cardiogrammes*. In Rastier, F., *L'analyse thématique des données textuelles*. Paris : Didier.
- Brunet E. (2006). *Navigation dans les rafales*. *Actes des 8^e Journées internationales d'Analyse des Données Textuelles*.
- Brunet E. (2008). *Les séquences (suite)*. *Actes des 9^e Journées internationales d'Analyse des Données Textuelles*.
- Brunet E. (2012). *Nouveau traitement des cooccurrences dans Hyperbase*. *Corpus*, (11).
- Lamalle C. et Salem A. (2002). *Types généralisés et topographie textuelle dans l'analyse quantitative des corpus textuels*. *Actes des 6^e Journées internationales d'Analyse des Données Textuelles*, Lausanne : EPLF, 85-94.
- Mayaffre D. (2014). *Plaidoyer en faveur de l'Analyse de Données co(n)Textuelles. Parcours cooccurentiels dans le discours présidentiel français (1958-2014)*. *Actes des 12^{es} Journées internationales d'Analyse des Données Textuelles*.
- Metwally H. (2017). *Les thèmes et le temps dans Le Monde diplomatique (1990-2008)*. Thèse de doctorat, Université Côte d'Azur.
- Metwally H. (2018). *L'apport du corpus-maquette à la mise en évidence des niveaux descriptifs de la chronologie du sens. Essai sur une Série Textuelle Chronologique du Monde diplomatique (1990-2008)*. *Actes des 14^{es} Journées internationales d'Analyse des Données Textuelles*.
- Rastier F. (2011). *La mesure et le grain. Sémantique de corpus*. Paris : Champion.
- Ratinaud P. et Marchand P. (2012). *Application de la méthode ALCESTE aux « gros » corpus et stabilité des « mondes lexicaux » : analyse du « CableGate » avec IRAMUTEQ*. *Actes des 11^{es} Journées internationales d'Analyse des Données Textuelles*.
- Ratinaud P. et Marchand P. (2016). *Quelques méthodes pour l'étude des relations entre classifications lexicales de corpus hétérogènes : application aux débats à l'Assemblée Nationale et aux sites web de partis politiques*. In Mayaffre D. et al. (éds.), *Actes des 13^{es} Journées internationales d'Analyse des Données Textuelles*.
- Reinert M. (1983) *Une méthode de classification descendante hiérarchique : application à l'analyse lexicale par contexte*. *Les cahiers de l'analyse des données*, 8(2), pp. 187-198.
- Reinert M. (1993). *Les « mondes lexicaux » et leur « logique » à travers l'analyse statistique d'un corpus de récits de cauchemars*. *Langage et société*, (66) : 5-93.
- Salem A. (1988). *Approches du temps lexical. Statistique textuelle et séries chronologiques*. *Mots* (17). pp. 105-143.
- Viprey J. -M. (2005). *Corpus et sémantique discursive : éléments de méthode pour la lecture des corpus*. In A. Condamines, *Sémantique et corpus*. Paris : Lavoisier.

Annexe 1 : Topologie du lemme 'femme' dans le corpus (rafales importantes encerclées)



Annexe 2 : Taille des corpus et importance de 'femme'

	1990-1993	1994-1997	1998-2001	2002-2008
Taille des corpus (occurrences, N)	2697013	2402434	2552998	3765908
Maquettes (effectifs des substantifs, V)	307	282	290	375
Maquettes (occurrences, N)	266439	218643	229119	382298
'femme' dans les corpus (occurrences, N)	534	876	1415	1444
Rang de 'femme' dans la liste des substantifs des maquettes	228	73	28	59

Annexe 3 : Les cooccurrents les plus proches de 'femme' classés selon l'ordre hiérarchique de l'attirance lexicale (corpus maquettes)

	1994_1997	1998_2001	2002_2008
enfant homme famille millier droit nombre travail emploi an vie	homme enfant Mme droit conférence travail emploi famille vie action	homme enfant droit journaliste jeune travail rôle presse taux association	homme enfant droit violence vie an corps association statut mouvement
campagne rôle association salaire mort exemple million personne minorité main	millier nombre association victime an million société film chômage monde	étude famille sein accès fonction syndicat statut violence égard éducation	loi travail emploi famille victime classe temps combat militant libération

Annexe 4 : Les cooccurrents les plus proches de 'femme' classés selon l'ordre hiérarchique de l'attirance lexicale (corpus intégraux)

1990_1993	1994_1997	1998_2001	2002_2008
homme_2* enfant_2 enceint_3 travailler_1 %_2 jeune_3 âge_2 stérilisation_2 naissance_2 famille_2	homme_2 enfant_2 jeune_3 sexe_2 %_2 fille_2 parité_2 mère_2 mme_2 sexuel_2	homme_2 enfant_2 jeune_3 féminin_3 égalité_2 sexe_2 masculin_3 droit_2 enceint_2 alphabétisation_2	homme_2 enfant_2 droit_2 féministe_2 jeune_3 égalité_2 mari_2 Gouges_2 viol_2 vilolence_2
fécondité_2 âgé_3 jeune_2 alphabétisation_2 voile_2 mari_2 souris_2 prostitution_2 espérance_2	masculin_3 droit_2 foyer_2 scientifique_3 patricarcal_3 travailler_1 ingénieur_2 grossesse_2 féminin_3	jeune_2 égal_3 journaliste_2 rédaction_2 grossesse_2 confrère_2 mari_2 travail_2 parité_2	oppression_2 enceint_3 fille_2 féministe_3 sexe_2 discrimination_2 avortement_2 mère_2 et_8

* Les indices numériques renvoient aux codes grammaticaux selon la norme de Cordial Analyseur (1 verbe, 2 substantif, 3 adjectif, 4 numéral, 5 pronom, 6 adverbe, 7 déterminant, 8 conjonction, 9 préposition).