

Le Monde et le « Grenelle de l'Environnement » : pistes pour l'analyse sémantique assistée par ordinateur d'un corpus de presse

Marie-Pierre Escoubas-Benveniste

« Sapienza » Università di Roma, Dipartimento di Geo.Sta.Sto, Facoltà di Economia

Résumé

« Grenelle Environnement » est le nom donné à une vaste consultation multipartite lancée par N. Sarkozy après son élection sur le thème de l'environnement dans le but de définir un plan d'action national contre la double crise économique et écologique. De 2007 à 2009 le journal *Le Monde* s'est abondamment fait l'écho de cet événement et de ses suites politiques comme en témoignent les 270 articles produits sur le sujet. L'objectif de cette étude est d'explorer les possibilités d'une analyse sémantique du corpus assistée par ordinateur. Etant donné la complexité énonciative de l'écrit journalistique et l'absence d'une typologie du genre auquel le rattacher, la première condition méthodologique est d'élaborer le corpus de manière à ce que les dimensions pragmatiques des textes soient accessibles au cours de l'analyse. La deuxième condition est d'identifier au niveau local des formes graphiques les moyens d'expression du sens, afin d'explorer la possibilité d'une catégorisation sémantique fiable. À ce stade de notre étude nous avons identifié trois phénomènes néologiques qui fournissent des indices sémantiques pour le repérage des motifs sémantiques.

Abstract

“Grenelle Environnement” is the name given to a large multipartite consultation organized by N. Sarkozy after his election in order to define some political and social answers to the economic and ecological crisis. From 2007 to 2009 the French national daily paper *Le Monde* has largely echoed the event and its political follow up publishing 270 articles on the subject matter. The purpose of this paper is to explore the possibilities of a computer aided semantic analysis of this corpus. Given the complex plurivocity of news paper article and the lack of a clear categorization, the first methodological condition is to built up the corpus in such a way that certain pragmatic dimensions of the fragments might be reachable during the analysis. The second condition will be to track at local graphic form level the formal expression of meaning in order to explore the possibility of establishing a reliable semantic categorization. At this point of our study we have identified three types of neologisms that provide strong semantic hints for our purpose.

Keywords: corpus linguistics, textual genre, assisted semantic analysis, journal texts, political ecology language.

1. Introduction

La présidence de Nicolas Sarkozy ¹ est marquée, dès le début, par un discours écologiste officiel. Au lendemain de son élection, le 18 mai 2007, le président crée un grand ministère de l'Écologie, du développement et de l'aménagement durables confié à Alain Juppé, seul Ministre d'Etat du premier gouvernement Fillon et à ce titre deuxième personnage du gouvernement.

¹ Élu le 6 mai 2007.

Au vu de l'importance de la charge ministérielle, il semble donc que le président Sarkozy tienne les promesses du candidat Sarkozy, signataire pendant la campagne électorale du *Pacte écologique* rédigé par la Fondation Nicolas-Hulot et le Comité de veille écologique². Le nouveau ministère organise alors une série de réunions pour la mise en place d'un processus de consultation baptisé « Grenelle Environnement ». Il s'agit d'une concertation multipartite entre des représentants de l'État, des collectivités locales, de certaines ONG engagées sur le front de la défense de l'environnement, des employeurs et des salariés. L'objectif déclaré est de définir un plan d'action écologique national qui puisse aller de pair avec « une production et une consommation durables » et favoriser « l'emploi et à la compétitivité »³. Le 6 juillet 2007, Jean-Louis Borloo, qui depuis juin 2007 est le successeur d'Alain Juppé dans le deuxième gouvernement Fillon, procède au « Lancement du Grenelle Environnement » et présente à la presse les grandes lignes du projet : « Le “Grenelle Environnement” réunira pour la première fois l'Etat et les représentants de la société civile afin de définir une feuille de route en faveur de l'écologie, du développement et de l'aménagement durables ».

L'initiative présidentielle est fortement médiatisée par les soins des services d'e-communication du gouvernement⁴. La presse et en particulier le quotidien national *Le Monde* suit de près l'événement et ses nombreux développements, se faisant ainsi le vecteur d'un discours politique à plusieurs voix sur les enjeux de l'écologie.

L'objectif de cette étude est de saisir les caractéristiques sémantiques de la production du *Monde* pendant deux ans autour de l'événement et de définir les conditions d'une analyse sémantique assistée par un logiciel d'analyse statistique lexicale et textuelle automatique d'un corpus de presse. Le texte journalistique est par nature polyphonique, les auteurs de ces textes sont différents et nombreux : le corpus doit donc être préparé de manière à permettre l'accès à toutes les informations disponibles concernant la production du texte. La première étape consiste à rechercher à la source documentaire les indices utilisables pour cette première catégorisation. Les marqueurs de l'énonciation complexe des textes ne sont pas immédiatement accessibles au traitement automatique. Il faut donc, dans un deuxième temps, détecter au niveau des formes graphiques, les seules accessibles par le logiciel, les manifestations de ces énonciations.

2. Caractéristiques du corpus LMGrenelle0709

Le point de vue qui a été adopté pour l'élaboration du corpus résulte d'une tentative de réponse à la question suivante : quels indices formels caractérisent les codes sémiotiques du *Monde* dans les discours écologistes qu'il a rendus publics avant, pendant, et après la consultation

² Publié le 7 juin 2006, le Pacte écologique, initiative du militant écologiste et animateur de télévision Nicolas Hulot, est défini par ses fondateurs comme une « démarche d'interpellation des décideurs politiques et des citoyens, visant à placer l'écologie au cœur de l'action publique ». Le Comité de veille écologique, co-auteur du Pacte, se définit comme « un comité d'experts scientifiques et techniques » et regroupe actuellement des économistes, sociologues, historiens, écologues, philosophes, architectes. Cf. <http://www.pacte-ecologique.org/>. En signant le Pacte écologique, les candidats aux élections présidentielles de 2007 prennent acte de la demande toujours croissante de la part des citoyens français de réponses politiques aux questions environnementales, et en saisissent du même coup le considérable enjeu électoral.

³ Parmi les six thèmes de réflexion sur lesquels ont travaillé les six groupes constitués pour l'élaboration de propositions (changement climatique et maîtrise énergétique ; biodiversité et ressources naturelles ; environnement et santé ; démocratie écologique) deux en particulier concernent l'adoption « de modes de production et de consommation durables » et la promotion « de modes de développement écologiques favorables à l'emploi et à la compétitivité ».

⁴ Le même jour est créé le site officiel www.legrenelle-environnement.fr sous l'égide du Ministère.

Grenelle Environnement? Une recherche automatique à partir du mot clé «Grenelle» dans l'archive électronique ⁵ du journal a produit 270 textes, écrits entre septembre 2007 et juin 2009, que nous avons retenus comme constitutifs d'un tout cohérent et pertinent à l'objet de l'étude (Bommier-Pincemin, 1999 : 416) ⁶. Afin de garantir que ces signes soient repérables et interprétables par une analyse automatisée il est nécessaire que les textes qui le constituent soient correctement et suffisamment documentés.

2.1. Métadonnées du corpus

En effet, selon la perspective théorique de la sémantique interprétative : tout texte est un objet social issu d'une pratique normée ⁷. Il produit du sens selon deux axes logiques. D'une part il est la manifestation particulière d'un genre et d'un discours issus d'une pratique sociale établie et reconnue. Il est donc à ce titre soumis à ces deux « instances de normativité textuelle » (Rastier, 2008) et porteur de caractéristiques sémiotiques définitoires de son genre et de son discours ; d'autre part le texte se situe dans un rapport d'intertextualité au sein de sa catégorie générique et, tout comme le signe saussurien, il tire sa valeur de cette coexistence et de cette résonance avec ses congénères. Saisir les aspects du sens textuel à travers la linguistique de corpus, généralement assistée par un traitement automatique, suppose une élaboration rigoureuse de l'objet d'étude : le corpus doit être un regroupement structuré de textes intégraux, documentés, éventuellement enrichis par des étiquetages, et rassemblés selon : a) une démarche théorique qui tienne compte des discours et des genres, et b) selon une démarche pratique en vue d'une gamme d'applications (Rastier, 2004, cité par Duteil-Mougel, 2005).

Constituer un corpus pour l'analyse sémantique de textes signifie donc définir et élaborer des « données » homogènes du point de vue du genre et du discours, afin que l'interprétation sémantique des résultats puisse tenir compte de la composante pragmatique du sens déterminée par les « conditions de production du texte ». Reconduire le texte à son genre prédéfini permet de faire la part entre les formes imposées par la norme et les formes qui s'en écartent et éviter ainsi le danger d'une surinterprétation sémantique. Or les genres, à quelques exceptions près, n'étant pas exhaustivement décrits, tout particulièrement les genres supposés constitutifs d'un discours de presse, s'il existe, la documentation du corpus nécessite la recherche de métadonnées exploitables à la source des textes.

2.2. Variables a priori et structuration interne des fragments

L'écrit journalistique résiste à une catégorisation scientifique par genre. Les classifications établies par les journalistes et les rédacteurs ne convergent pas vers une typologie unanimement reconnue par les linguistes (Bonnafous, 1991). Même si elles sont explicitement signalées dans les pages du journal (éditorial, horizons-débats analyse...) ces classifications restent à vocation didactique et ne sont pas opératoires pour une typologie unificatrice satisfaisante, puisqu'elles nécessitent un deuxième type de classement, par sujet, à travers les rubriques traditionnelles de la presse écrite : économie, international, culture, société. (Adam, 1997 ; 2001).

⁵ L'archive du journal à laquelle nous avons eu accès contient les numéros de Monde publiés à partir de l'année 2006.

⁶ Selon l'auteur, la pertinence et la cohérence (constitutives de la signifiante du corpus par rapport au projet d'analyse) représentent la première des trois conditions auxquelles celui-ci doit répondre.

⁷ La sémantique interprétative de François Rastier propose un appareil théorique pour la description des différents paliers sémiotiques par lesquels les textes signifient. (cf. Hébert, 2001) pour une présentation détaillée.

La possibilité d'une première caractérisation a priori des fragments se pose donc avec acuité. Si l'on peut supposer qu'un certain degré d'homogénéité thématique (écologie, politique, économie) est garanti aux textes en vertu du mode de sélection et de la « fenêtre temporelle » restreinte dans laquelle ils se situent, il n'en va pas de même en ce qui concerne leur forme : le volume des fragments en est un indice, le plus court étant la légende d'un dessin humoristique. Or le format adopté par l'archive électronique ne rend absolument pas compte de cette hétérogénéité : tous les textes obéissent à une structuration minimale⁸ qui ne prévoit pas de champ *author* et tous sont uniformément étiquetés comme *Doc Type : article*.

Pourtant l'éditeur opère des distinctions non explicitées mais stables dans le laps de temps considéré pour la mention de l'auteur à la fin du corps du texte : *By X ; Propos recueillis par X ; Y, [expression d'une fonction]*. Pour une première catégorisation des fragments constitutifs du corpus nous faisons l'hypothèse que *By X* signale la présence d'un auteur journaliste, donc d'un article (A) ; *Propos recueillis par X*, celle d'une interview (I) menée par un auteur journaliste ; *Y, [expression d'une fonction ou toponyme]* manifeste la prise de parole d'un énonciateur en dehors d'une fonction journalistique, que nous catégorisons comme tribune (T). Enfin pour 58 textes, l'auteur n'est pas précisé (np). Nous attribuons donc 3 variables en amont de l'analyse : a) Type de document (A, I, T, np, DH) ; b) auteur : 89 auteurs nommés, caractérisables 55 auteurs de A, 33 auteurs de T, 1 auteur de DH ; c) date (codée mm/jour/an pour permettre des regroupements de fragments à l'échelle du mois). Le corpus compte 212 articles, 17 interviews, 40 tribunes et une légende.

Ainsi documenté, le corpus LMGrenelle0709 doit permettre a) la mise en relation de phénomènes sémantiques observés localement avec des indications de rang supérieur, aussi bien à l'échelle de chaque fragment, à travers les sections (*Titre, Résumé, Texte*) qu'à l'échelle du corpus, à travers les variables affectées aux fragments (*Type, Auteur, Date*) ; 2) d'affiner ultérieurement l'analyse sur des sous-parties du corpus, obtenus par regroupements de fragments, en croisant les variables prédéfinies avec les premiers résultats de l'analyse.

3. Problématique et méthode d'exploration et d'analyse

3.1. Comment détecter les manifestations locales du sens

Au niveau local, les formes graphiques du vocabulaire manifestent différents types de sens : le plus immédiat lorsqu'on parcourt la longue liste de « mots » identifiés dans le corpus est le sens référentiel. C'est celui qui permet de dresser la liste des noms propres de lieux, de personnes, d'aliments (cf. Bolasco, 2002), celui par lequel le mot assure une partie de sa fonction, celle de pointeur vers le monde. L'autre aspect du sens, qui nous intéresse ici, est le sens non référentiel des formes linguistiques, dit dialogique, cette épaisseur sémantique constituée par la trace d'une énonciation, d'une intention, d'un jugement, d'une action par les mots et qui constitue un code discursif.

Les mots du lexique habituellement cités comme les plus immédiatement signifiants à cet égard sont les déictiques (nommés « embrayeurs » par E. Benveniste ou « indexicaux » par F. Rastier), par lesquels de manière directe, mais plus ou moins explicite, l'énonciateur s'identifie comme tel (*je, nous, mon, notre, ...*). Nous faisons l'hypothèse qu'ils sont très fréquents dans le

⁸ Métadonnées documentaires de l'archive électronique du Monde : *Title, Language, Source, Document Type, Abstract (French), Full Text Word Count, Full Text Word Count, ISSN, Accession Number, Persistent link to this record (Permalink), Bookmark and Share, Database*.

corpus, étant donné le nombre de tribunes et la place du discours cité dans les écrits de presse. En raison même de cette polyphonie, leur fréquence ne peut être significative qu'une fois leurs occurrences rapportées aux différents énonciateurs. Si la structuration du corpus permet en partie de corrélérer ces déictiques aux instances de l'énonciation dans le cas des tribunes, cela est impossible dans le cas des articles, du moins dans le cadre d'une analyse assistée. Il faut donc rechercher d'autres types de mots sur lesquels centrer la première phase de l'exploration.

Les néologismes sont de bons candidats car ils manifestent la présence d'un énonciateur innovant ou promoteur d'un lexique inédit. Cette créativité peut être issue a) de la forme seule de l'expression, b) de sa forme et de son sens, c) d'un changement de sens d'une forme existante :

- a) l'emprunt et la siglaison sont deux procédés de néologie purement morphologique. Le premier importe une forme issue d'un autre système, le second crée un mot nouveau par réduction de la dénomination complète ;
- b) l'affixation, la composition, la dérivation de mots préexistants sont des procédés néologiques morphosémantiques productifs ;
- c) la néologie par changement de sens est un procédé sémantique plus délicat à saisir. Il peut être détecté à travers l'emploi du mot, en analysant ses cooccurrents, selon les préceptes de la linguistique de corpus (Sinclair). La coprésence inattendue de formes graphiques oblige à une interprétation susceptible de modifier le sens des formes si le phénomène se répète dans le temps. Les figures tropiques, en particulier l'antonomase, sont des procédés néologiques par changement de sens.

La presse en tant que lieu de débat où se croisent les prises de parole individuelles et officielles, est un lieu privilégié de production et de divulgation des néologismes, ce qui nous conforte dans notre choix méthodologique de repérage de ces phénomènes lexicaux. Mais le choix même des mots pour référer aux choses est une manifestation de l'énonciateur. Ainsi l'analyse s'intéressera aux expressions dont le sémantisme, à travers leur connotation, manifeste une instance langagière agissante (évaluatrice, accusatrice, ...) et fait émerger une dimension dialogique et des motifs sémantiques qui constituent autant de pistes à explorer.

3.2. Les fonctions d'exploration pour l'analyse sémantique

L'analyse du corpus est assistée par le logiciel TALTAC 2.9 (Bolasco et al., 2004) qui tient compte de la structuration en amont du corpus, de la structuration interne des fragments et permet l'extraction d'informations selon des catégorisations définies en aval des premières phases d'exploration. L'analyse automatique opère dans notre cas sur des formes graphiques (formes fléchies, variantes graphiques). Le corpus n'est pas lemmatisé ni annoté pour deux raisons théoriques : 1) afin de ne pas limiter son emploi en intégrant au corpus des annotations qui ne sont pas indispensables pour notre analyse ; 2) parce que la définition des notions de «mot» et de « parties du discours » ne font pas l'unanimité parmi les linguistes et ne permettent pas de fournir une approche unitaire directement exploitable dans le cadre de l'analyse de corpus (Sinclair, 1991) ⁹.

Le logiciel calcule les segments répétés (suites de formes graphiques cooccurrentes deux fois au moins). Il permet d'extraire des informations, qui sont transformables en données (après analyse et catégorisation) et, selon le principe de récursivité, réinsérées successivement dans le vocabulaire du corpus afin d'extraire de nouvelles formes. Les fonctions d'exploration opèrent à deux niveaux : sur l'axe paradigmatique (la liste des formes graphiques du corpus), et sur l'axe syntagmatique (segments répétés, concordances). Il permet en outre l'extraction d'information complexe par combinaison de plusieurs variables catégorielles affectées à des classes de formes graphiques.

⁹ À propos du nettoyage du corpus Sinclair écrit «the safest policy is to keep the text as it is, unprocessed and clean of any other codes. These can be added for further investigations» (*Ibid.* : 21).

Les résultats que nous présentons ici sont issus de l'analyse du vocabulaire (repérage des thèmes sémantiques par la morphologie lexicale), des segments répétés et des concordances. Ils suggèrent des pistes concernant la présence de motifs sémantiques, les possibilités de lexicalisation des segments répétés et le choix de traits sémantiques pertinents pour la catégorisation des formes.

4. Premiers résultats et pistes à explorer

Le vocabulaire du corpus est composé de 16 030 formes graphiques (*types*) et de 180 980 occurrences totales de ces formes (*token*), le pourcentage de hapax est de 43%. Une première analyse fait apparaître trois processus néologiques caractéristiques du corpus : l'antonomase, la « labellisation » écologique des mots par préfixation (*éco-*) ou adjectivation (*vert/green*), la siglaison, qui affecte toutes sortes de syntagmes et toutes sortes de référents. Ils constituent les premiers signes dialogiques.

4.1. Antonomase et hyperproductivité néologique ou le Grenelle des ondes n'est pas le Grenelle de la mer

Le mot plein le plus fréquent (785 occ. 26^e rang) est naturellement le mot *Grenelle*. L'analyse des segments répétés incluant cette forme (Tab. 1) révèle que a) la dénomination de l'événement dans le corpus ne correspond pas à son appellation officielle ; b) le syntagme nominal construit *Grenelle de N* est productif et donne lieu à 6 autres types de segments isoformes

<i>Segment</i>	<i>Nbre d'occ.</i>	<i>Segment</i>	<i>Nbre d'occ.</i>
Grenelle de l'environnement	371	Grenelle (les accords de /rue de)	7
Grenelle de la mer (Mer)	47	Grenelle de l'insertion	4
Grenelle des ondes	33	Grenelle de la fiscalité	4
Grenelle du pouvoir d'achat (un)	6	Grenelle des territoires	4

Tableau 1 : Segments répétés contenant « Grenelle »

L'analyse des concordances du nom propre confirme la productivité du syntagme *Grenelle de N* et révèle 15 autres types de syntagmes forgés sur ce modèle : *G. de la chasse*, *G. de la culture*, *G. de la diversité*, *G. de la fiscalité*, *G. de la Guadeloupe*, *G. de la Presse*, *G. de la recherche et de l'innovation*, *G. de la relance*, *G. des antennes*, *G. des ports de plaisance*, *G. des stades*, *G. du centre*, *G. du pouvoir d'achat*, *G. du socialisme*, *G. du sport*.

c) d'autres types syntaxiques de segments répétés sont présents : *les accords de Grenelle*, *rue de Grenelle* ; et *Grenelle 1*, *Grenelle 2*, *loi Grenelle 1*, *loi Grenelle 2* (Tab. 1 et Tab. 2) :

<i>Segment</i>	<i>Nbre d'occ.</i>	<i>Segment</i>	<i>Nbre d'occ.</i>
Grenelle 1 (Grenelle I de l'environnement)	38	Grenelle 2 (Grenelle II)	16
Grenelle 1 (loi) / Grenelle I (loi)	12	Grenelle II (loi)	2
Grenelle 1 (projet de loi)	4	Grenelle II (le)	4

Tableau 2 : Segments répétés contenant « Grenelle » et un numéral

L'analyse des formes morphologiquement apparentées (Tab. 3) montre que d) le nom propre est fléchi (pluriel), déterminé (démonstratif) et entre dans des mécanismes de dérivation et de composition ; e) certains de ces mots construits sont sémantiquement connotés.

<i>Segment</i>	<i>Nbre d'occ.</i>	<i>Rang</i>	<i>Zone</i>
Grenelle	785	26	Haute
contre-Grenelle	4	4008	Basse
pseudo-Grenelle	4	4008	Basse
après-Grenelle	2	6242	Basse
Grenello-incompatibles	2	6242	Basse
Grenelle-là	1	9132	Basse
Grenelles	1	9132	Basse
grenellienne	1	9132	Basse

Tableau 3 : *Formes morphologiquement apparentées à «Grenelle»*

Tout d'abord ces résultats nous prouvent que nous sommes en présence d'une antonomase, figure rhétorique qui consiste à utiliser un nom propre comme un nom commun. Ressort puissant de la créativité discursive et lexicale ¹⁰, l'antonomase est une des manifestations de la dimension dialogique de la dénomination. À ce titre elle représente un fait observable précieux pour l'analyse sémantique de corpus.

Sur le plan morphosyntaxique, le Np *Grenelle* présente les caractéristiques d'emploi du nom commun : c'est une forme fléchie, dérivable, composable, passible de détermination comme le montrent les concordances (*un tel Grenelle, un deuxième Grenelle*). Sur le plan morphosémantique, le Np entre dans des syntagmes nominaux qui permettent au niveau du corpus d'en ébaucher l'évolution chronologique : *rue de Grenelle* ¹¹ : *on parle à l'époque de "l'accord de la rue de Grenelle"* (concordance). L'antonomase *Grenelle de l'environnement* établit volontairement une analogie entre les deux référents et invite le lecteur à transférer certaines caractéristiques sémantiques du référent initial au référent secondaire. Le problème de son interprétabilité sémantique dépend en partie de l'accessibilité du référent initial (Siblot et al., 2000). Ici, la chronologie facilite les choses : dans un contexte de célébrations du 40^e anniversaire des événements de Mai 68, le référent initial « *accords de Grenelle* » ¹² est réévoqué, ses traits sémantiques réactivés et le transfert sémantique au nouvel événement par l'entremise d'un nom propre néologique « *Grenelle de l'Environnement* » peut opérer. *Grenelle* a pris à son compte exclusif les traits sémantiques « accords multipartites qui aboutissent à la dissolution de la crise de mai 68 ». Le Np néologique *Grenelle Environnement* institué officiellement en 2007 opère un double transfert sémantique : 1) la crise déclarée est environnementale ; 2) la consultation multipartite sera aussi garantie de succès.

Comment interpréter le fait que le corpus n'atteste pas l'appellation officielle « Grenelle Environnement » adoptée par le Ministère? Le syntagme NN fait pourtant bien partie des schémas lexicaux productifs du français ¹³. On peut probablement y lire une adaptation syntaxique qui répondrait à des motivations divulgatives justifiées au sein d'un média. Le néologisme officiel *Grenelle Environnement* étant potentiellement doublement opaque, selon le bagage culturel du lecteur : il cumule le flou sémantique attaché à un nom propre décontextualisé

¹⁰ L'histoire du mot *grève* en est un exemple éclairant.

¹¹ Désignation du Ministère du Travail, situé dans un hôtel particulier de la rue de Grenelle, où furent signés les accords en 1968.

¹² Ce syntagme lexicalisé désigne les accords conclus, pendant la crise de mai 68, entre les représentants du gouvernement Pompidou, des syndicats et des organisations patronales. Ils aboutissent à des avancées salariales (augmentation de 25% du SMIG et de 10% des salaires réels, création de la section syndicale d'entreprise) et contribuent à la dissolution de la crise socio-politique.

¹³ Cf. *cadeau surprise, produit grand public, régime minceur...*

Grenelle, et celui de la relation syntaxique non exprimée qui le relie à un autre nom. Le transformer en un composé N de N permet de supprimer un palier d'opacité sémantique, de le rendre plus lisible et en quelque sorte de lui enlever son étrangeté.

Ainsi le néologisme officiel est modifié et médiatisé à travers une matrice syntaxique hautement productive. La prolifération des types de N classifieurs révèle une hétérogénéité et une ambiguïté sémantiques intéressantes : lexique de la nature (*environnement, mer*) ; lexique économique et social ¹⁴ (*fiscalité, pouvoir d'achat, relance, insertion*) ; lexique politique (*socialisme*), lexique ambigu (*territoires, ondes, centre*) ; lexique des activités (*chasse, sport*), lexique abstrait (*diversité*) et concret (*antennes, stades, ports de plaisances*). Sur le plan paradigmatique s'instaure ainsi un rapport implicite entre ces signes et l'association quantitativement fréquente de *Grenelle* avec tout type d'objet sémantique, même les plus éloignés sémantiquement du classifieur de référence : l'environnement.

L'antonomase suit son cours morphologique (*grenellien*) non sans quelque résonance tragique... ¹⁵ Du point de vue sémantique, se manifestent au niveau des formes graphiques issues du mot pivot *Grenelle* la présence des motifs de l'opposition active (*contre-Grenelle*), de la fausseté et de l'imposture (*pseudo-Grenelle*) ainsi que le motif de l'incompatibilité à travers un adjectif néologique par imitation (*Grenello-incompatibles*) (Tab. 4). Le contexte d'emploi permet d'en retracer l'origine énonciative et suggère le motif associé de méfiance.

LM_010809	Il' Alliance pour la planète met en ligne sur son site la carte de France des projets, selon elle , “	Grenello-incompatibles	“ . La méfiance s' installe entre le gouvernement et les écologistes
LM_010809	le sujet aux chefs d' état et aux ministres de l' environnement . La croissance est-elle	Kyoto-compatible	Dix ans après la signature du protocole de Kyoto , voici venu le temps des travaux pratiques

Tableau 4 : « *Grenello-incompatibles* » et « *Kyoto-compatible* »

On ne peut pas ne pas voir ici une création mimétique de la préfixation par formants, autre processus néologique très répandu, qui a lui aussi valeur d'indice dialogique.

3.2. Les formes graphiques « labellisées » du vocabulaire

La recherche automatisée des formes graphiques contenant le formant « *éco-* » dans le vocabulaire du corpus produit 44 types de formes graphiques, variantes graphiques confondues et 575 occurrences totales. Bon nombre de ces formes sont impliquées dans des syntagmes répétés (noms propres et communs en particuliers d'acteurs, que nous ne traiterons pas ici). Leur confrontation à un dictionnaire de langue courante (PR 2008) permet d'identifier les néologismes (ou du moins les formes non encore stabilisées lexicalement dans le dictionnaire de langue générale). Ces unités sont au nombre de 19 (variantes et formes fléchies confondues) dont 13 hapax répartis en 2 n. propres et 11 n. communs (Tab. 5).

Les « mots » ainsi sémantiquement motivés sont hétérogènes • morphologiquement (N communs, déverbaux et déadjectivaux ; Adj.) ; • lexicalement : urbanisme (*écoquartier, écocités, écoconstruction, ...*) ; économie (*écocroissance, écoprêt, écoredevance*) et plus ou moins

¹⁴ Il existe un précédent *le Grenelle des retraites*, cité par Jean Pruvost, dont la première mention n'est pas datée (Pruvost and Sablayrolles, 2003 : 106).

¹⁵ Économie ou écologie? Le choix est cornélien.

hermétiques selon l'opacité sémantique de l'élément préfixé (*écopastille*, *éco-PTZ*, *Ecophyto 2018*). Certains travaux ont démontré comment la préfixation à partir de formants (*éco-*) assure une sorte de pré-marquage sémantique des mots en amont des discours. Ce procédé de «préqualification» qui a pour effet de conférer une apparence de technicité et de scientificité aux mots peut servir à dissimuler l'usage idéologique et politique qui en est fait. (Baggioni, 1977 : 37). Ainsi la forme même du mot, sémantiquement motivée par la préfixation qualifiante, propose de manière synthétique ce que nous appelons une labellisation sémantique par la forme graphique. Tout se passe comme si le nom devait servir de garant pour l'homologation de la chose écologique, et comme si la dénomination avait pour seule fonction de faire exister le concept.

<i>Forme graphique</i>	<i>Nbre d'occ.</i>	<i>Forme graphique</i>	<i>Nbre d'occ.</i>
écoquartier(s), éco-quartier	8	éco-innovations	1
écoprêt(s), éco prêt, éco-prêt	6	éco-label	1
écoblanchiment, éco-blanchiment	5	Ecophyto_2018	1
écocités, éco-cités	4	éco-PTZ	1
écocroissance	4	écoredevance	1
écoresponsables	2	écoresponsabilité	1
Eco_Artisan	1	éco-tartufes	1
écoconception	1	écotoxicologie	1
écoconstruction	1	écopastille	1
écodéconstruction	1	Total	42

Tableau 5 : Les 19 types de néologismes en « éco »

Comme dans le cas précédent le mécanisme néologique s'applique à des unités fortement connotées, qui manifestent clairement la présence d'une énonciation : lexique évaluatif (*responsable*, *responsabilité*) qui ne sont pas clairement interprétables ; deux noms aux connotations très négatives (*blanchiment*, *tartufe*)¹⁶. À travers l'instance énonciative critique émerge le motif sémantique de l'accusation d'imposture, de fausseté et d'hypocrisie : *éco-tartufe* (cf. *pseudo-Grenelle*) et le double motif de la fraude et de la purification (*écoblanchiment*). Ces deux formes indiquent la présence d'un détournement sémantique du procédé au service d'un contre-discours : elles convoquent des associations de traits sémantiques dysphoriques. La labellisation écologique des mots est imitée et renversée sémantiquement donc démasquée comme une manœuvre dissimulatrice des actions frauduleuses de faux écologistes animés par d'autres intérêts.

La qualification par l'épithète *vert* concourt aussi à cet effet sémiotique de labellisation du nom cooccurrent. Deux des 4 segments répétés les plus fréquents (Tab. 6) produisent des néologismes du lexique économique : *croissance verte* et *fiscalité verte*.

<i>Forme graphique</i>	<i>Nbre d'occ.</i>
croissance verte	23
fiscalité verte	10
trame verte	2
énergies vertes	2

Tableau 6 : Segments répétés de type Nom vert

¹⁶ PR (sv *tartufe*) hypocrite, faux dévôt. PR (sv *blanchiment*) 2) Action de décolorer pour rendre blanc. 3) Fig. Action de blanchir (de l'argent). Blanchiment de fonds d'origine frauduleuse, des narcodollars.

Les concordances de l'adjectif (Tab. 7) confirment au niveau des discours l'usage de vert au sens de « relatif à l'environnement naturel » : Grenelle vert, mais aussi sa cooccurrence avec le lexique économique (emplois, New Deal, prêt, critères, carburant), politique (à droite ... à gauche) ainsi que le motif sémantique de la mise en doute de la vérité (se veut vert, mais) et du travestissement (habiller, repeindre en vert) et l'explicitation du contre-discours ironique.

LM_052808	remis par Hélène Jourda lors du <i>Grenelle l'</i>	vert	et les interrogations de plus en
LM_032709	Un <i>carburant</i> un peu plus “	vert	“ bientôt à la pompe : Le “ E10 “
0LM_032409	connectés à grande vitesse. Le <i>New Deal</i>	vert	passé par l' investissement
0LM_051609	ceux qui seront exemplaires . Le <i>prêt</i>	vert	à taux zéro a été créé et 5 000
LM_121607	Mais au-delà des stricts <i>critères</i> “	verts	“, la “ ville durable “ se donne pour objectif
0LM_032409	estimés à 500 000 . Un jour , les “ <i>emplois</i>	verts	“ viendront compléter le tertiaire
LM_061109	promis , juré : <i>tous</i> “	verts	“, tous préoccupés, à <i>droite comme à gauche</i>
0LM_051609	annoncé par le gouvernement <i>se veut</i> “	vert	“, mais les discours du président
LM_111208	et ce n' est pas en <i>habillant</i> Total en	vert	Qu' on va changer quoi que ce soit
LM_100909	Hulot propose de <i>repeindre le capitalisme en</i>	vert	, a <i>ironisé</i> Sophie Divry, journaliste à <i>La Décroissance</i>

Tableau 7 : Concordances de «vert»

LM_040108	aux yeux de certains Américains pour qui “ <i>green_is_gold</i> “ - ne semble pas beaucoup plus réaliste.
0LM_032409	ont récemment appelé de leurs vœux une nouvelle croissance , la “ <i>green_growth</i> “ , <i>que je préfère traduire par</i> : « <i>écocroissance</i> »
LM_110408	mettre en sourdine ses ambitions d'être leader dans le “ <i>green_business</i> “ , rendant plus difficile la mise en application en France du Grenelle de l' en
LM_102908	Vaste opération de “ <i>green-washing</i> “ , consistant à <i>barder de références environnementales</i> ces univers dédiés à la

Tableau 8 : Concordances de «green-»

Les concordances de l'adjectif *green* (Tab. 8) confirment le motif du travestissement écologique d'objets qui ne le sont pas, ainsi que l'association de l'épithète au lexique économique. On apprend ainsi que *écocroissance* est préféré au calque français *croissance verte*. Tous deux sont pourtant avantageusement ambigus, mais l'apparente technicité du premier s'oppose aux accents bucoliques du second. Il est alors éclairant de pouvoir identifier l'énonciateur de ces positions linguistiques affirmées : il s'agit de Michel Barnier, Ministre de l'Agriculture du second gouvernement Fillon. Le corpus de presse ne reflète pas cette préférence : 4 occ. pour *écocroissance* (Tab. 5) et 23 pour *croissance verte* (Tab. 6).

3.3. La siglaison : *éco-PTZ, Medad, Meedat, HQE, chimères néologiques*

Le corpus contient 271 types de formes graphiques (sigles, acronymes ou composés) pour 1359 occurrences totales. Les dix formes les plus fréquentes (Tab. 9) sont des objets concrets (OGM, CO2, Mon810), des organisations (UMP, CGT, ONG), des entreprises (EDF, SNCF), des termes économiques (PIB, RSA) qui sont au centre des débats. Elles ne sont généralement plus accompagnées de l'expression étendue du syntagme originel.

<i>Forme graphique</i>	<i>Nbre d'occ.</i>	<i>Rang</i>
OGM	202	81
UMP	115	148
CO2	67	284
ONG	63	302
EDF	41	500
Mon810	39	518
CGT	31	656
PIB	30	684
SNCF	25	847
RSA	22	977

Tableau 9 : les 10 sigles les plus fréquents

Les sigles qui attirent l'attention sont les plus hermétiques et les moins fréquents, qui ne font pas nécessairement partie du lexique disponible du locuteur moyen. Ils recouvrent des réalités complexes, des acteurs (*Agence française de sécurité sanitaire de l'environnement et du travail* (Afsset)) ou des appellations dont le référent exact reste à identifier (*Transport développement intermodalité environnement* (TDIE) ; THPE). L'analyse des concordances des sigles les plus opaques, comme le sigle labellisé éco-PTZ (Tab. 10), permet de les associer à la dénomination complète de leur référent.

LM_091308	du prêt à taux zéro écologique , l'	éco-PTZ	, à hauteur de 300 euros par mètre
LM_090508	le président ne croyait qu' à moitié au	Medad	- le ministère géant qui regroupe
LM_102908	labellisée haute qualité environnementale (HQE) de l' enseigne Système_U devait ouvrir ses
LM_102908	galerie marchande . En France, un label	HQE	commerce a été élaboré . Pour l' obtenir, un
LM_121107	“ très haute performance énergétique “ (THPE), qui impose une consommation inférieure de

Tableau 10 : Concordances des sigles opaques peu fréquents

Néologisme de forme qui relève de l'abréviation : la siglaison consiste généralement à regrouper les initiales des constituants d'un syntagme long et l'économie linguistique est souvent invoquée pour la justifier. La presse fait un usage massif des sigles et acronymes (dénomination des sociétés, des partis, des organismes, des associations, des hommes et femmes politiques) et se doit aussi de les définir s'ils ne sont pas entrés dans l'usage commun. Elle se fait ainsi le relais des discours dominants mais ne se prive pas de commenter les appellations (Tab. 11, *ministère géant*). Le sigle est révélateur d'un statut face à la langue : « pouvant éventuellement emprunter des sigles à la langue écrite mais ne pouvant pas les créer le locuteur moyen [...] est ici en situation de faiblesse, est dominé. » (Calvet, 1980 : 116-117, cité par Pires, 2007 : 295). Non seulement sa création reste le privilège d'une instance qui détient un pouvoir linguistique, mais son contenu est de plus très hermétique et inaccessible au non initié (Pirès, 2007 ; Percebois, 2009). Il peut même être ambigu et source de malentendus : ainsi par exemple «UMP» ne désigne pas le même mouvement politique en 2002 et en 2007¹⁷.

Désigner par un sigle un Ministère, un homme politique, une « label » (Tab.10) c'est à la fois les faire entrer dans la durée, l'officialité et la technicité, tout en les situant, en raison de leur opacité, en dehors de la sphère des « non initiés ». L'enjeu est important si l'on en juge par la

¹⁷ En 2002, il s'agit de l'« Union pour la majorité présidentielle » rassemblement autour de Jacques Chirac ; en 2007, de l' « Union pour un Mouvement Populaire » de Nicolas Sarkozy.

productivité du sigle politique : entre janvier et novembre 2008, deux sigles différents désignent le Ministère de l'écologie (Tab. 11).

LM_013008	A signé, avec le <i>ministère de l'écologie</i> , du <i>développement et de l'aménagement durables</i> (Medad), une convention, élaborée dans le cadre du Grenelle de l'environnement, sur leurs engagements mutuels
LM_112708	Écologie, <i>de l'Énergie</i> , du développement durable et de l'aménagement du territoire (Meedat) met actuellement en place une " démarche écocités " pour accompagner l'émergence

Tableau 11 : Medad et Meedat

5. Conclusions

Au terme de l'analyse des formes graphiques en suivant la piste de 3 procédés particulièrement saillants dans le corpus LMGrenelle20072009 : l'antonomase, la labellisation des mots par préfixation et épithète, la siglaison, nous sommes en possession de données qui vont nous permettre de procéder à une deuxième phase de l'analyse. • Des thèmes (*grenelle*, *éco-*, *vert*, *green*) mais aussi (*bio-* *agro-*) vont nous permettre de catégoriser les formes graphiques et de comparer leur distribution (cf. *écocroissance* vs *croissance verte*). • Des motifs sémantiques tels que le mensonge (vrai/faux + le dire), le doute (donc le croire), la pureté, la morale, la loi, mais aussi l'antagonisme suggéré par *contre-Grenelle* (*contre-*, *anti-*, *dé-*) vont nous permettre de définir des catégories sémantiques pertinentes au sein du corpus. De nombreuses questions restent ouvertes en ce qui concerne les choix théoriques et pratiques que nous devrions privilégier en particulier pour l'annotation sémantique connotative. À ce stade de l'analyse, les caractéristiques de la forme de l'expression révèlent deux tendances sémiotiques contraires dans le corpus : la technicité, l'opacité et la codification par les formes d'un discours écologique dominant et l'émergence d'un retournement caricatural des mêmes procédés au service d'un contre-discours qui introduit les motifs du mensonge, du doute, de la méfiance en démasquant les stratégies discursives manipulatrices.

Références

- Adam J.-M. (1997). Unités rédactionnelles et genres discursifs : cadre général pour une approche de la presse écrite. *Pratiques*, vol. 94, Metz.
- Adam J.-M. (2001). Genres de la presse écrite et analyse de discours. *Semen*, 13, [En ligne], mis en ligne le 30 avril 2007. URL : <http://semen.revues.org/document2597.html>. Consulté le 30 juillet 2009.
- Baggioni D. (1977). Aspects sociolinguistiques de la néologie lexicale : le vocabulaire 'à connotations scientifiques' et le 'fétichisme des mots'. *LAB* 17 : 33-40.
- Bolasco S. (2002). Integrazione statistico-linguistica nell'analisi del contenuto. In Mazzara, B., editor, *Metodi qualitativi in Psicologia Sociale. Prospettive teoriche e strumenti operativi*, Roma : Carocci, pp. 329-342.
- Bolasco S., Biscaglia B. and Baiocchi F. (2004). Estrazione di informazione dai testi. *Mondo Digitale*, vol. III, 1 : 27-43.
- Bommier-Pincemin B. (1999). Diffusion ciblée automatique d'informations : conception et mise en œuvre d'une linguistique textuelle pour la caractérisation des destinataires et des documents. Thèse, Paris IV.
- Bonnafous S. (1991). *L'immigration prise aux mots*. Paris : Kimé.

- Duteil-Mougel C. (2005). Les mécanismes persuasifs des textes politiques. *Corpus*, vol. 4, Consulté le 29 juillet 2009. URL : <http://corpus.revues.org/index357.html>.
- Hébert L. (2001). *Introduction à la sémantique des textes*. Paris : Honoré Champion.
- Le Monde (2006-2009). *Le Monde Interactif*, 21 Bis Rue Claude Bernard, Paris 75005.
- Percebois J. (2009). Les procédés d'économie linguistique : quels partenaires pour la communication? In Greenstein, R. editor, *Langue et culture, mariage de raison?*. Paris : Presses de la Sorbonne, pp. 17-44.
- Pires M. (2007). Le détournement de sigle : le cas de CPE. *Langage et société*, vol. 121-122 : 290-301.
- Pruvost J. and Sablayrolles J.-F. (2003). *Les néologismes*. coll. « Que sais-je ? ». Paris: PUF.
- Rastier F. (2008). Que cachent les « données textuelles »?. In *JADT 2008*, Lyon.
- Siblot P. and Leroy S. (2000). L'antonomase entre nom propre et catégorisation. *Mots*, vol. 63 : 89-104.
- Sinclair J. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.

