



HAL
open science

Exploration textuelle du discours d'un quotidien régional au carrefour du XIXe et du XXe siècles : Le Petit Comtois (1883-1903)

Virginie Lethier

► **To cite this version:**

Virginie Lethier. Exploration textuelle du discours d'un quotidien régional au carrefour du XIXe et du XXe siècles : Le Petit Comtois (1883-1903). Sciences de l'Homme et Société. Université de Franche-Comté, 2009. Français. NNT : . tel-01339066

HAL Id: tel-01339066

<https://shs.hal.science/tel-01339066>

Submitted on 29 Jun 2016

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

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



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

UNIVERSITÉ DE FRANCHE-COMTÉ

ÉCOLE DOCTORALE « LANGAGES, ESPACES, TEMPS, SOCIÉTÉ »

Thèse en vue de l'obtention du titre de docteur en

SCIENCES DU LANGAGE

EXPLORATION TEXTUELLE DU DISCOURS D'UN
QUOTIDIEN RÉGIONAL AU CARREFOUR DU XIX^e ET DU
XX^e SIÈCLES : *LE PETIT COMTOIS* (1883-1903)

Présentée et soutenue publiquement par

Virginie LETHIER

Le 3 décembre 2009

Sous la direction de MM. Les Professeurs Jean-Marie VIPREY et Philippe
SCHEPENS

Membres du Jury :

Sonia BRANCA-ROSOFF, Professeure à l'Université Paris III, Sorbonne nouvelle.
Rapporteur.

Jacques GUILHAUMOU, Directeur de recherche au CNRS, UMR « Triangles »,
ENS-LSH, Lyon. Rapporteur.

Pierre FIALA, Maître de conférences à l'Université Paris 12.

Michael PALMER, Professeur à l'Université Paris 3.

Alain RABATEL, Professeur à l'Université Claude-Bernard, Lyon 1.

Philippe SCHEPENS, Professeur à l'Université de Franche-Comté.

Jean-Marie VIPREY, Professeur à l'Université de Franche-Comté.

Tome 1

Remerciements

Je remercie mes directeurs de recherche, Jean-Marie Viprey et Philippe Schepens, qui m'ont permis de mener cette thèse à son terme. Qu'ils reçoivent le témoignage de ma sincère reconnaissance pour m'avoir guidée et conseillée, en me faisant l'honneur de me confier plusieurs responsabilités dont j'espère avoir été à la hauteur.

Mes remerciements s'adressent également à Sonia Branca-Rosoff, Jacques Guilhaumou, Pierre Fiala, Michael Palmer, ainsi qu'à Alain Rabatel, qui m'ont fait l'honneur de participer au jury de cette thèse.

Pour leurs conseils et l'intérêt qu'ils ont porté à mes recherches, j'adresse mes remerciements à François Daoust, Lou Burnard, Margareta Kastberg-Sjöblom, Alpha Ousmane Barry et Andrée Chauvin-Vileno, ainsi qu'à l'ensemble des membres du Laseldi et du Centre Jacques Petit, que j'ai eu le plaisir de côtoyer pendant ces années.

Enfin, je tiens à témoigner ma profonde gratitude à tous ceux qui, à des degrés divers, mais avec une égale bienveillance, ont contribué à ce travail : ma famille, mes proches et mes collègues. Une pensée toute particulière à mes complices Céline Lambert et Adèle Petitclerc pour leur soutien et leur bonne humeur quotidienne.

Introduction

Exhumer le discours d'un quotidien régional *Le Petit Comtois* (1883-1944) paru sous la Troisième République : telle est la visée de notre recherche. « Exhumer » au sens plein du terme d'abord, puisqu'il s'agit bel et bien d'extraire un *corpus* de son support matériel initial, tombé dans un oubli presque complet, pour l'intégrer à un nouvel environnement numérique, que nous entendons comme la condition de la valorisation et du renouvellement des lectures interdisciplinaires de cette ressource. « Exhumer » en un sens métaphorique ensuite : notre projet de recherche vise à proposer un éclairage sur le discours d'un quotidien qui est normalement – c'est le destin des quotidiens – voué à une vie éphémère, compte tenu des conditions de production qui sont propres à cette presse.

Le type de discours que nous visons à étudier, celui de la presse régionale, semble avoir toujours pâti d'une faible légitimité sociale. L'attestent les récriminations de ses contemporains du début du XIX^e siècle à la fin du XX^e siècle à son sujet. En effet, au début du XIX^e siècle, le quotidien régional est accusé d'être une presse de qualité inférieure en raison de sa grande dépendance aux sources parisiennes qui lui fournissent la majeure partie de son contenu ; on stigmatise le journaliste de province, cet « âne du métier » qui ne fait preuve d'aucune distance vis-à-vis des informations qu'il se contente de rapporter ; on condamne enfin le manque de perspective et le contenu peu digne d'intérêt de la feuille : « Le journaliste départemental, à d'honorables exceptions près, est de sa nature un être excessivement paresseux ; il fume, se promène, joue considérablement au billard, mais rédige rarement. Il reçoit tous les jours de Paris une correspondance qui lui apporte articles politiques et littéraires, entrefilets, faits Paris [sic] et causeries de Salon, il n'a qu'à choisir ; la paire de ciseaux fait le reste... Ainsi, en réalité, ce sont MM. Delaire, de Vigny

et Degouve-Denuncques, qui sont les rédacteurs en chefs des journaux de province¹ ».

Au XX^e siècle, les critiques sont sensiblement identiques : l'horizon restreint du contenu du quotidien régional, focalisé sur la vie sociale et culturelle locale, est régulièrement dénoncé ; la qualité de la plume qui la produit est encore mise en doute : « Un agrégat de bienséance, de mondanités, d'urbanité, de déférence remplit les colonnes grises et froides de ces canards sans vie (...) Ces photos rangs d'oignons, ce vocabulaire aride et ampoulé, ces lieux communs, ces « beaux succès » et ces « grandes manifestations » ne racontent pas la vie. C'est du toc ! » (Grandmaison, (de), 1973 : 134).

De telles représentations de la presse de province, qui manifestent son déficit de légitimité sociale, ne sauraient compromettre l'intérêt socio-historique de l'étude de la presse régionale de la Troisième République. Bien au contraire, nous voulons montrer qu'elle constitue une remarquable ressource pour l'analyse du discours, que cette dernière soit centrée sur l'étude des genres journalistiques appréhendés dans leur *historicité*, ou qu'elle s'attache à observer la lutte des forces en présence pour « prendre la langue » (Bonnafous & Tournier, 1995). Plongeant ses racines dans *l'analyse du discours du côté de l'histoire* (Guilhaumou, *et al.*, 1974), l'analyse de données textuelles (Lebart & Salem, 1994) et *l'analyse textuelle du discours* (Adam, Heidmann, 2005), notre travail entend contribuer à sa mesure au mouvement de constitution raisonnée de très grandes bases textuelles et au développement de méthodologies adaptées aux défis que ces très grandes bases constituent. Ce faisant, notre recherche, qui vise à fournir à la communauté de nouvelles archives à dimension historique, s'inscrit dans un programme de renouvellement des analyses du discours et de développement de l'histoire des concepts de la Troisième République.

¹ *Physiologie de la presse. Biographie des journalistes et des journaux de Paris et de la province*, (1841), Paris : Jules Laisné Editeur, pp. 103-104.
Disponible sur <<http://gallica.bnf.fr/ark:/12148/bpt6k1113858>>

Un premier rappel nous semble essentiel : le contingent le plus important des productions discursives de l'« âge d'or du journalisme » n'est pas celui des grands titres nationaux, des illustres signatures, des genres nobles, mais celui des journaux de province et de leurs formes d'écriture hétérogènes. Une seconde remarque justifie pour sa part l'intérêt que peut représenter la presse en général et plus spécifiquement la presse régionale de la Troisième République pour l'analyse du discours du côté de l'histoire. Les opinions légitimes, s'il en est, ne sont pas assignées dans les grands textes à prétention intellectuelle : la dimension réflexive des acteurs d'une société s'exerce dans les pratiques les plus ordinaires et les plus immédiates, ainsi que dans les types de discours les plus communs comme nous le rappellent les historiens (notamment Chartier, 1990 : 105-110). Centrée sur la vie locale et associative, la presse régionale de la Troisième République constitue une source qui offre un accès original à l'appréhension du fonctionnement de l'espace public et de la formation des opinions. Rappelons à cet égard le rôle spécifique assuré par la presse régionale à la fin du XIX^e siècle. Organe porte-parole d'une sensibilité politique dont les historiens s'accordent à reconnaître l'influence décisive en période d'élections (Lagoueyte, 1997 : 73), le quotidien régional du XIX^e siècle œuvre quotidiennement pour gagner l'adhésion de son (é)lectorat à ses combats politiques. Pour ce faire, il s'adapte aux caractéristiques socio-culturelles et politiques de celui-ci : « [...] Au contact direct des populations, ces feuilles savaient adapter les mots d'ordre des partis aux tendances et aux besoins de ceux auxquelles elles s'adressaient. Elles savaient utiliser au bénéfice de la politique qu'elles défendaient tous les incidents de la vie locale, tous les faux pas de leurs adversaires » (Goguel, 1958 : VII).

Parce qu'il est une co-construction structurée par la connaissance que la rédaction a de son lectorat, le discours de la presse régionale se distingue très nettement du discours de presse d'un grand titre parisien à une heure où l'opposition rural vs urbain a pour pendant de nets clivages sociologiques et politiques. De par les polémiques qu'elles alimentent, de par la reproduction

des comptes-rendus des élections et des réunions, des professions de foi et les programmes politiques dans les élections, la presse régionale constitue par ailleurs un remarquable dispositif d'observation des opinions et des avis individuels des différentes figures dont la voix se donne à entendre dans ses colonnes (lecteur/correspondant-citoyen, élu, expert).

Bien conscients de la valeur de cette source, le LASELDI (EA 2181) et le Centre Jacques Petit (EA 3187) ont initié en 2003 au sein du pôle *Archives, Base, Corpus* de la MSHE Claude-Nicolas Ledoux un ambitieux projet de constitution d'une archive de presse numérisée à partir de l'ensemble des titres de la région Franche-Comté du XIX^e et du XX^e siècle conservés par les bibliothèques et archives du ressort régional, en partenariat avec ces dernières (notamment la Bibliothèque d'Etude et de Conservation de Besançon). À terme, c'est l'archivage numérique en mode plein texte de dizaines de titres de presse de sensibilité politique diverses tels que *La Gazette de Franche-Comté (1831-1834)*, *L'Abeille franc-comtoise (1876-1888)*, *Le Flambeau (1903-1911)*, *L'Eclair Comtois (1903-1939)* qui est visé par cette expérience pionnière. Au-delà d'une solution de conservation pérenne de ce patrimoine éminemment précieux, le projet vise à renforcer le statut scientifique de ces ressources en explorant et en développant les possibilités offertes par l'archivage numérique.

C'est dans le cadre de ce projet de laboratoire que s'inscrit notre recherche, débutée en 2005, figurant les développements inauguraux de l'acquisition des premiers fonds de la base et de leur exploration, portant sur le quotidien *Le Petit Comtois (1883-1944)*. Loin de correspondre à deux volets distincts, - d'un côté l'acquisition de données textuelles numérisées et de l'autre, l'analyse proprement dite de ces données réunies en corpus -, ces deux axes de recherche sont, dans notre perspective, complémentaires et s'unifient autour d'un programme qui est celui d'une *analyse textuelle du discours* (Adam, Heidmann, 2005). Visant à dépasser le déficit textuel de l'analyse du discours, l'analyse textuelle du discours intègre à son programme celui d'une *nouvelle*

philologie numérique, telle que J.-M. Viprey la nomme (2005b) à la suite de F. Rastier (2000, 2001). Dans ce cadre, le moment d'établissement philologique s'impose comme le garde-fou crucial contre l'oubli de l'historicité du texte. Numériser une collection de presse ancienne nous confronte à l'inscription du texte sur un document, nous met face au rôle structurant de son co(n)texte et réfléchit les opérations qu'effectue le chercheur sur le texte. Ce changement de support nous invite ainsi à une réflexion critique sur l'établissement numérique du texte, qui non seulement n'entraînerait aucune perte d'informations significatives, mais qui enrichirait le texte de ses lectures. Si ce volet de nos recherches constitue donc un laboratoire d'expérimentations en vue de l'acquisition de données textuelles finement établies, il est aussi un temps de réflexion sur les conditions de normalisation des données, permettant le partage et la confrontation des lectures du texte ainsi que des outils méthodologiques des sciences des textes, dont dépendent leurs avancées.

Notre recherche a également pour objectif d'ouvrir et de parcourir des pistes méthodologiques appropriées à l'étude de grands volumes de données textuelles de presse ancienne selon les modalités de l'analyse textuelle du discours, tel que le permet le passage du support-papier au numérique. Il s'agit en effet de pouvoir proposer un éclairage sur le discours du *Petit Comtois*, observé à partir d'un corpus de quelque 26000 articles parus de 1883 à 1903, de natures et de tailles extrêmement hétérogènes, pour un volume total excédant les cinq millions de mots. Par le syntagme « exploration textuelle du discours »², nous ne voulons pas tant insister sur la prise pour objet d'un discours peu connu, que sur un certain parcours de lecture à visée exploratoire au sein de grands ensembles textuels, ayant pour objectif de recueillir des informations objectivées par le texte lui-même. Notre démarche convoque pour ce faire les outils de la statistique textuelle (Lebart & Salem, 1994), qui offrent le socle méthodologique d'un traitement exhaustif et systématique auquel ne peut prétendre une lecture traditionnelle. Plus : seuls ces outils permettent de

² Syntagme emprunté à J.-M. Viprey (2006a : 181) et que nous faisons nôtre pleinement dans ces travaux.

construire des vues synthétiques de complexes ensembles de données, non pas entendues comme des arguments probatoires, mais comme des pistes en vue de la poursuite de l'analyse. C'est donc le potentiel heuristique de ces outils, mais aussi leur capacité à construire différentes prises de vue sur la *textualité* définie comme « la combinaison de parcours linéaire et réticulaire » (Adam, 2006b : 5) que nous voulons mettre à l'épreuve. Par exploration textuelle du discours enfin, nous souhaitons insister sur un certain parcours d'interprétation, ayant pour pierre de touche le co(n)texte, dans un mouvement allant du corpus au hors-corpus, où le primat est donné à l'environnement linguistique immédiat, au cotexte comme réseau de textes « autres », mis en série dans le corpus (Mayaffre, 2002a).

En raison du caractère expérimental du projet dans lequel elle s'inscrit et de la complexité du discours de presse qu'elle vise à appréhender, notre recherche s'est naturellement orientée vers la constitution d'une série de rapports d'activité scientifique. De ce fait, cette thèse s'articule en cinq chapitres, dont les lignes directrices seront les suivantes :

Chapitre I. Explorer la presse régionale de la Troisième République : État de l'art

Notre thèse débutera par une mise en relief des enjeux dont sont porteuses la constitution et l'exploration d'une base de presse régionale de la Troisième République en mode plein-texte. Il s'agira pour nous de montrer l'intérêt de l'exploration d'une telle source pour les sciences humaines en général et particulièrement pour les disciplines et interdisciplines soucieuses de la dimension historique du discours. Notre propos sera de montrer en quoi et sous quelles conditions la constitution d'une base de presse textuelle peut dynamiser ces champs de recherche. Ce faisant, nous identifierons des obstacles méthodologico-techniques, qui sont autant de défis à lever pour l'avancement des sciences textuelles.

Chapitre II. Constitution du dispositif d'exploration du *Petit Comtois*

Proposer un compte-rendu de nos expérimentations visant à dépasser les difficultés soulevées par l'acquisition de grands ensembles de données établies philologiquement selon les normes contemporaines : tel sera l'objectif assigné au premier temps de ce chapitre. Nous nous attacherons ensuite à présenter notre corpus extrait de la base ainsi constituée. Après avoir présenté brièvement l'histoire du quotidien *Le Petit Comtois* (1883-1944) et son contexte général de parution, nous présenterons les caractéristiques du corpus sur lequel se construit cette étude.

Chapitre III. Examen de la mise en forme de l'information dans *Le Petit Comtois*

Notre démarche d'exploration textuelle du discours du *Petit Comtois* se propose comme une série de prises de vue sur sa matérialité. Postulant que celle-ci nous renvoie tout d'abord à l'inscription d'un texte sur un document dont l'organisation participe à en construire le sens, nous consacrerons un troisième chapitre à l'examen de l'« aire scripturale » (Peytard, 1975) du quotidien. Nous nous intéresserons ensuite aux différents niveaux péritextuels importants que sont le rubriquage et l'appareil des titres. Il s'agira pour nous d'observer les ressorts de la mise en forme de l'information dans *Le Petit Comtois*, envisagés comme des points d'accès privilégiés à l'*ethos* du quotidien.

Chapitre IV. Explorations du vocabulaire du *Petit Comtois*.

S'appuyant sur les outils de l'analyse statistique textuelle, le chapitre IV sera composé d'une série d'expérimentations sur l'organisation du vocabulaire dans notre corpus. Après avoir brièvement rappelé les évolutions conceptuelles et méthodologiques de la lexicométrie, nous évoquerons les caractéristiques globales du vocabulaire du corpus pour ensuite examiner l'évolution diachronique (« macro »-distribution). Nous nous tournerons ensuite vers une analyse (« micro »-) distributionnelle des occurrences et co-occurrences, qui nous permettra d'ouvrir des perspectives sur les réseaux lexico-thématiques

endogènes, qui constituent un niveau de structuration trop souvent négligé du vocabulaire d'un vaste corpus. Puis, nous analyserons l'influence de la diachronie sur cette configuration fine du vocabulaire, en la confrontant aux indications livrées par la première classe d'analyse, fondée sur l'observation du niveau macro-distributionnel. Enfin, nous tenterons d'examiner les facteurs responsables de la remarquable stabilité de l'organisation micro-distributionnelle du vocabulaire, en interrogeant plus particulièrement l'influence exercée par le rubriquage.

Chapitre V. Approches du niveau morphosyntaxique sur le corpus *Petit Comtois*. Essai de typologisation textuelle.

Dans le dernier chapitre, nous examinerons les apports d'une analyse fondée sur le niveau morphosyntaxique à l'exploration textuelle du discours du *Petit Comtois*. Visant, dans la continuité des indications pointées par l'entrée du vocabulaire, à constituer une typologie des rubriques du journal, ce volet de nos recherches se présente comme une investigation méthodologique préalable à la construction d'un protocole d'observation des formes génériques du *Petit Comtois*.

Chapitre I.

Explorer la presse régionale de la Troisième République : État de l'art

Le présent chapitre vise à mettre en relief les enjeux interdisciplinaires d'une exploration de la presse de la Troisième République telle que permise par une *nouvelle philologie numérique* (Viprey, 2005b) et les obstacles rencontrés dans ce cadre.

À cet effet, nous ferons tout d'abord état des travaux menés en sciences humaines sur la presse régionale de la Troisième République, en orientant notre réflexion sur une première difficulté : cette source donne lieu à un champ de recherche qui souffre d'un manque de dynamique.

Nous évoquerons ensuite les apports féconds d'une histoire linguistique et ses liens avec une analyse du discours soucieuse de la matérialité textuelle, en vue de pointer dans quel contexte la constitution d'une base de presse ancienne se révèle un chantier de premier ordre pour l'analyse du discours et les sciences des textes.

Enfin, nous examinerons la question de la numérisation de la presse de la Troisième République sous ses aspects méthodologiques et techniques. Il s'agira de faire état des obstacles pesant sur cette étape décisive, obstacles dont la levée est la condition d'un renouveau interdisciplinaire de la lecture de la presse ancienne.

1.1. Approches de la presse régionale de la Troisième République par les sciences humaines

La presse régionale de la fin du XIX^e siècle et du début du XX^e siècle, comme tout objet médiatique, constitue un vaste domaine de recherches plus qu'elle n'est l'objet d'une discipline particulière. Pour mieux définir notre démarche, il convient donc de rappeler l'horizon que constitue l'ensemble des études menées en sciences humaines sur cet objet.

1.1.1. Un champ de recherche peu investi par les sciences de l'information et de la communication ainsi que par l'analyse du discours

Force est de constater que la période et le type de presse que nous souhaitons étudier n'ont fait l'objet que de très rares études en analyse du discours et en sciences de l'information et de la communication, disciplines pourtant axées vers l'analyse des médias.

1.1.1.1. En analyse du discours

Ayant pour pierre angulaire les problématiques de l'assujettissement du « sujet » à l'idéologie, de l'interdiscours, des rapports des formations discursives aux formations idéologiques, l'analyse du discours s'est d'abord essentiellement développée autour de recherches prenant pour objet le discours politique au sens institutionnel (discours de syndicats, de partis, etc.).

L'analyse du discours, définissant le sens comme historique et social par essence, et portant de fait une attention toute particulière aux processus de circulation discursive, n'a que progressivement constitué la presse comme terrain d'observation. Des études comme celle de D. Maldidier (1971), qui compare les lectures du discours du général De Gaulle dans six grands quotidiens, illustrent la perspective d'une analyse linguistique du discours

médiatique intégrant une forte dimension historique. On se référera encore à l'analyse de la construction de l'événement du meeting de Charléty (27 mai 1968) dans les principales formes rhétoriques de l'écriture journalistique réalisée par R. Robin et D. Maldidier (1977). Mis à part l'exemplaire étude pionnière, sur laquelle nous aurons l'occasion de revenir plus en détails, menée par le laboratoire de lexicologie politique de l'ENS de Saint-Cloud sur le *Père Duchesne* d'Hébert, principal organe de la presse pamphlétaire de la Révolution française, l'analyse du discours « à la française » ne compte guère de travaux portant sur la presse ancienne. De façon très nette, les recherches menées en analyse du discours médiatique prennent pour objet la presse contemporaine de leur réalisation. Pour illustrer notre propos, on citera les travaux de M. Ebel et de P. Fiala (1983) qui explorent la xénophobie dans la formation sociale suisse à partir d'un corpus de deux quotidiens pendant la votation sur les initiatives référendaires de 1970, 1974 et 1977, et d'un corpus du courrier des lecteurs de ces journaux aux mêmes périodes. À quelques notables exceptions près, telles que les recherches récentes d'Alice Krieg-Planque (2003), retraçant le trajet discursif de la formule « purification ethnique » de 1980 à 1994 dans un corpus médiatique, il convient de remarquer que l'analyse du discours médiatique, à partir des années 1980-1990, tend de surcroît à se détourner de corpus journalistique à dimension historique.

De fait, si on observe à la fin du XX^e siècle une massification des travaux en analyse du discours prenant pour objet des corpus centrés sur la notion de genre médiatique, ces derniers tendent à relever d'un contexte historique proche de celui du temps de l'analyse. Il est à cet égard symptomatique que les recherches d'un numéro phare de la revue *Semen*, consacré aux *Genres de la presse écrite et analyse de discours*, portent pour la grande majorité sur des corpus de presse du XX^e siècle. Dans ce numéro, Nicole Juper et Thierry Herman (2000) mènent ainsi une étude rigoureuse sur l'éditorial, à partir d'une analyse contrastive de journaux suisses et français de 1997 à 1999 ; Sylvie

Durrer (2000), pour sa part, propose un éclairage du billet à partir de l'examen d'un corpus s'échelonnant de 1996 à 1998.

La presse de la Troisième République, et plus encore la presse régionale de la Troisième République, demeure ainsi en analyse du discours un terrain de recherche inexploré, tant comme dispositif d'observation de la circulation des stratégies discursives que des formes génériques de cette diachronie.

1.1.1.2. En sciences de l'information et de la communication

De façon tout aussi paradoxale, la presse régionale du XIX^e siècle est un média très peu investi par les Sciences de l'Information et de la Communication (SIC). Instituées en catégorie universitaire en 1975¹, les Sic traitent pourtant initialement de l'accès à l'information, de son traitement et de sa gestion. Le fort caractère interdisciplinaire des Sic, qu'on les envisage comme discipline ou comme domaine de recherche en structuration², nous semble un premier facteur qui justifie la diversité et l'hétérogénéité de ces objets de recherche, et qui explique, par suite, le décalage quantitatif entre le nombre d'études menées dans ce champ et le nombre d'analyses consacrées à la presse régionale du XIX^e siècle. Outre ce premier trait structurel, il nous semble légitime de supposer que la production de la recherche des Sic, qui est très fortement liée à la demande sociale, s'est dès lors détournée d'un média datant d'une époque révolue. Songeons notamment aux nombreux travaux financés par diverses institutions s'inspirant des *mass communication research*, qui procédaient jusqu'à la fin des années 1980 à la mesure des effets des médias sur l'opinion publique à partir de méthodes quantitatives.

Par ailleurs, les Sic ont émergé à la fin d'un XX^e siècle qui a pris acte des profondes mutations relatives aux technologies de l'information et de la

¹ Depuis 1975, les SIC sont la 71^e section du Conseil National des Universités qui réunit les formations au journalisme et à la communication, les spécialistes des bibliothèques et de la documentation)

² Voir à ce sujet Davallon (2004 : 30).

communication, mutations qu'elles se sont attachées à décrire. Dans ce cadre, c'est la télévision qui a occupé une place prépondérante dans le champ disciplinaire (Jost & Leblanc, 1998 ; Lochard & Boyer, 1998b). Le second objet privilégié par les Sic a été les pratiques d'information dans les organisations (documentation scolaire, information scientifique, etc.). On comprend dès lors pourquoi les travaux des sémiologues comme ceux de J. Tétu et de M. Mouillaud (1989) sur la morphologie du journal, ainsi que ceux de l'historien de la presse M. Palmer (1983) sur la *naissance du journalisme moderne* (1863-1914) ne peuvent être considérés comme pleinement représentatifs. Ils ont cependant considérablement enrichi l'état des savoirs sur la presse du XIX^e siècle.

1.1.2. Une presse essentiellement étudiée par les sciences historiques

C'est dans le champ des sciences historiques que se concentre la majeure partie des études de la presse régionale du XIX^e siècle. Néanmoins, il ressort d'une analyse bibliométrique de ce champ que la presse quotidienne régionale de la Troisième République (désormais PQR3R) constitue un champ de recherches peu développé au profit d'une presse nationale de même période. Empressons-nous de préciser qu'au clivage presse de province vs presse nationale, s'ajoute un décalage quantitatif entre le nombre de recherches menées sur la PQR3R et celles relevant d'autres périodes comme le XVII^e et le début du XIX^e (Beloeil, 1994 ; Albertan & Albertan, 1989 ; Trénard, 1989). Preuve en est la structuration des publications d'histoire générale de la presse : révélatrice du retard historiographique accumulé vis-à-vis de la presse régionale en général, et plus spécifiquement de la PQR3R, la publication de référence de C. Bellanger et J. Godechot (1972) édifie une histoire générale de la presse très largement orientée vers la presse *nationale*, tandis que le récent ouvrage de Marc Martin (2002) est le premier ouvrage entièrement consacré à

l'histoire générale de la presse régionale en France dont une part substantielle soit consacrée aux journaux de la Troisième République.

Pour organiser notre présentation des travaux menés en histoire sur la presse de la Troisième République, nous distinguerons à la suite de P. Albert (1965) deux axes de recherches, suivant que l'objectif des études vise à une « *histoire à travers la presse* » ou à « *faire l'histoire de la presse* » (Albert, 1965 : 24).

1.1.2.1. Bref panorama des recherches en sciences historiques sur et/ou par la presse régionale de la Troisième République

Le premier axe de recherche distingué regroupe les études relevant d'une *histoire par la presse*, qui participent d'un intérêt général pour l'état d'une société et de ses représentations. Ces études utilisent le témoignage des journaux comme éclairage d'un fait de l'histoire générale et envisagent ainsi la presse comme *medium*, tandis que la priorité est donnée au « contenu » plus qu'au « contenant », pour reprendre le *distinguo*³ établi par P. Albert (1965 : 23).

L'histoire par la PQR3R souffre d'un manque de systématisme

Pour l'historien, la presse régionale constitue une source de connaissance irremplaçable d'une époque donnée. L'historien trouvera dans les rubriques du quotidien régional, y compris celles présentant un intérêt textuel moindre, d'innombrables renseignements qui lui permettent de reconstituer la vie politique, économique et sociale locale, comme le signale J. Kayser :

[...] que de précisions utiles ressortent de l'étude des journaux au jour le jour et notamment des actes de l'état civil, de détails de vie sociale locale, des comptes rendus de manifestations publiques ou de réunions privées (parfois même de leurs seules convocations), du cours des marchés, des plans de travaux publics et des articles consacrés à leur état d'avancement, de l'analyse de la publicité locale et des petites annonces (Kayser, 1957 : 301).

³ Très largement admis dans la communauté historique, ce *distinguo* donne bien entendu lieu à discussion.

Ce type de presse peut également servir de source documentaire pour l'étude des représentations et de l'opinion⁴ de la province au sujet d'un événement, d'un thème social ou culturel, principaux axes de recherches pour lesquels l'historien convoque la presse régionale. À titre d'exemple, on citera l'étude de J. Kayser (1957), qui procède à une mesure de l'impopularité en province de J. Ferry à l'heure de sa chute en mars 1885 à la suite de sa prise de position en faveur de l'expansion coloniale, prise de position qui entraîne une violente opposition parlementaire. Pour ce faire, l'historien opère un sondage dans plusieurs quotidiens régionaux, dont les lieux de publication et la sensibilité politique varient. Ainsi, au terme du dépouillement de titres de presse de la région du Calvados, tels que *Le Journal de Caen*, *L'Eclaireur du Calvados*, *Le Moniteur du Calvados*, de la presse de l'Isère, du Lot-et-Garonne et du Var, J. Kayser observe que d'une part, l'impopularité de J. Ferry est aussi intense en province qu'à Paris, et d'autre part, que celle-ci est partagée au-delà des sensibilités politiques, puisque tant les quotidiens de gauche que de droite attaquent avec virulence l'homme politique.

Une autre recherche emblématique de cette perspective historiographique nous est livrée par J. Ponty (1974). L'historienne propose un essai de typologie des positions de la presse française parisienne et provinciale à l'égard de l'affaire Dreyfus à trois moments clés : l'affaire Esterhazy-Zola (janvier-février 1898), le suicide du Colonel Henry (septembre 1898) et le Procès de Rennes (août-septembre 1899). Cette étude se fonde sur un panel de 55 titres de presse sélectionnés en fonction de leur notoriété et de leur tirage, dont 29 quotidiens régionaux. J. Ponty distingue six positions : les antidreyfusards non-violents, les antidreyfusards haineux, les antidreyfusards haineux et antisémites, les révisionnistes et les dreyfusards. L'étude montre qu'en 1898, l'antidreyfusisme domine très largement, tant à Paris qu'en

⁴ Pour une discussion des problèmes méthodologiques soulevés par l'analyse de l'opinion publique à travers la presse problèmes méthodologiques, nous renvoyons à J.-J. Becker (1988 : 171-182).

province. Près d'un tiers des titres font preuve d'un antidreyfusisme non-violent motivé par le « goût de l'ordre, [l'] autorité de la chose jugée, [le] sentiment national et [l'] amour de l'armée » (Ponty, 1974 : 202). La grande majorité des titres adopte une position antidreyfusarde violente, à relier à un sentiment national exacerbé ayant pour corollaire des manifestations de xénophobie et de haine. Sur l'échantillon de 55 titres observé par J. Ponty, seuls sept sont dreyfusards : six de ces quotidiens sont parisiens. Parmi les 29 titres régionaux, seul *Le Réveil du Nord*, journal guesdiste et antimilitariste, refuse alors de s'engager dans l'affaire. L'arrestation du lieutenant-colonel Henry (30 août 1898) et son suicide, qui entraînent la démission du général de Boisdeffre et celle du ministre de la guerre Cavaignac, le 3 septembre, sont source d'un renversement des positions. Les journaux antidreyfusards non-violents rejoignent le camp révisionniste. Ce dernier représente dès lors presque 40% du nombre d'exemplaires du panel de presse quotidienne observé, contre à peine 2% le mois précédent. Les événements de septembre 1898 renforcent la haine des antidreyfusards violents. Le procès de Rennes (août-septembre 1899) marque pour sa part le revirement des quotidiens antidreyfusards non-violents au camp révisionniste. J. Ponty remarque sur ce point qu'il s'agit essentiellement de quotidiens régionaux, plus longtemps réfractaires au dreyfusisme que les quotidiens parisiens. En conclusion, J. Ponty remarque que l'adoucissement de la grâce de Dreyfus s'est fait « en dépit de l'opinion de la majorité de la presse » (Ponty, 1974 : 220).

On aura remarqué que les exemples d'études que nous venons d'évoquer sont déjà anciens (1957, 1974), ce qui nous semble indiquer que la recherche contemporaine en histoire « par la presse » souffre d'un manque de systématisme. En effet, si actuellement de très nombreux mémoires de Maîtrise et de Master référencés prennent pour source documentaire la presse régionale du XIX^e siècle, il n'existe, à notre connaissance, que peu de recherches d'envergure relevant de cette perspective.

L'histoire de la PQR3R accuse un sérieux retard historiographique

La deuxième orientation des recherches historiques regroupe quant à elle les travaux prenant la presse pour objet : elle se focalise sur la description et l'analyse du monde des journaux en lui-même. La priorité est ainsi donnée au « contenant » au détriment du « contenu », qui ne reste abordé que dans la mesure où il apporte un éclairage explicite sur les tendances de la publication considérée. Il s'agit ainsi d'une *histoire de la presse* au sens strict, dont le statut est double. Selon une conception faible, elle pourra être envisagée comme une science auxiliaire de l'histoire contemporaine, aidant les historiens à utiliser les témoignages des journaux. Selon une conception forte mettant l'accent sur le fait que ce secteur disciplinaire replace les journaux dans leur contexte historique, l'histoire de la presse peut être perçue comme un des nombreux secteurs de l'histoire générale.

Il convient ici de noter que le développement de l'histoire de la presse et le renouvellement de son champ d'investigation après la seconde guerre mondiale doit beaucoup à la création de la section d'histoire de l'Institut français de presse (IFP), sous l'impulsion de Jacques Kayser et de Jacques Godechot. Ces derniers insistent, avec P. Albert, sur la signification culturelle de la PQR3R. Les travaux de ces pionniers restent la référence dans le domaine, à l'instar de l'ouvrage de J. Kayser (1958) *La presse de province sous la IIIe République*, et de la parution sous la direction de P. Albert, *Documents pour l'histoire de la presse de province dans la seconde moitié du XIXe siècle* (1973).

Si l'on cherche à dresser les orientations des chantiers de l'histoire de la presse régionale couvrant *a minima* la période de la Troisième République, un premier constat s'impose : très nettement, les monographies historiques de journaux régionaux publiées depuis 30 ans forment le contingent le plus nombreux des études sur la PQR3R. Se focalisant sur la description des aspects économiques et techniques de la vie d'un quotidien et ses modes de diffusion, ce premier chantier participe d'une histoire des journaux. *L'Histoire de l'Est Républicain (1889-1914)* de Colette Hirtz (1973) fournit un exemple de

premier ordre de cette approche. On trouve dans cet ouvrage une documentation précise sur la première période d'existence d'un modeste quotidien régional devenu une institution, dont l'implantation sur la région Est s'opère sur fond de concurrence féroce et de rachat de nombreuses feuilles locales comme *Le Comtois* publié de 1945 au début des années 1970, journal qui a succédé au *Petit Comtois* dans des circonstances que nous préciserons ultérieurement (cf. chapitre 2). L'étude de C. Hirtz se fonde sur les divers aspects de la vie matérielle de l'*Est Républicain*, de son capital social à sa zone de rayonnement, en passant par l'évolution du tarif de ses publicités : ces informations s'avèrent d'autant plus précieuses qu'aucune archive des débuts du quotidien n'a été conservée. Parallèlement à l'histoire matérielle du quotidien, l'auteure procède, dans une perspective d'analyse du contenu, à l'examen des thèmes révélés structurants de l'identité du quotidien, parmi lesquels l'Alsace-Lorraine et de la politique coloniale. L'auteure s'attache également à décrypter la ligne politique de l'Est-Républicain en diachronie, en portant une attention toute particulière à la position adoptée par le quotidien vis-à-vis des grands événements de la vie politique française de la Troisième République comme l'affaire Dreyfus.

Autre exemple, plus récent, de monographie historique de titres régionaux : l'étude réalisée par F. Torrès (2002) sur *La Dépêche du Midi* de 1870-2000. L'historien procède à une rétrospective des événements marquants de l'histoire de l'un des plus anciens quotidiens français et des plus importants journaux de province. Reflétant l'influence de la médiologie sur les travaux menés en histoire de la presse, une part importante de l'étude de F. Torrès est consacrée au traitement des modalités de diffusion et de distribution du quotidien, ainsi qu'à l'évolution des moyens technologiques disponibles. À la différence de l'étude de l'*Est Républicain* précédemment évoquée, l'approche de F. Torrès porte par ailleurs une attention toute particulière aux acteurs individuels qui ont marqué l'histoire du quotidien et aux mutations du milieu professionnel qui se donnent à lire à travers la vie du journal. L'évolution de la

ligne éditoriale et de la sensibilité politique de *La Dépêche du Midi* est ainsi éclairée par la succession des directeurs et rédacteurs en chef de ce quotidien, de Louis Blairet, au patron de presse Jean-Michel Baylet, en passant par Arthur Huc (1854-1932) et Maurice Sarrault (1869-1943).

En dehors des monographies, un second noyau de recherche s'oriente vers la réalisation de tableaux de la presse régionale. Ces travaux consistent à recenser et à décrire tous les titres publiés dans une région durant une période généralement longue. Ils sont très souvent menés dans le cadre d'une thèse de troisième cycle. Pour la région de Franche-Comté, on citera ainsi la thèse de Marcel Vogne (1972) sur *La presse en Franche-Comté des origines à 1970*, ainsi que les travaux de Roger Marlin (1965a), proposant une histoire de la presse du Doubs et du Territoire de Belfort de 1865 à 1944. Pour la presse du Nord-Pas-de-Calais, on se référera aux travaux de J.-P Visse (2004) ; pour la presse de Lorraine mosellane de 1860-1940, aux travaux de F. Roth (1983). Enfin pour la presse de Montpellier, nous mentionnerons les travaux de R. Andréani (1989).

Dressant les grands axes des relations entre l'histoire et la PQR3R, cet état de l'art vient donc insister sur ceci : en dépit de la ressource précieuse qu'elle constitue, la PQR3R demeure un champ de recherche aujourd'hui faiblement investi. Nous souhaitons à présent proposer un regard critique sur les facteurs qui pèsent ou qui ont pesé sur le développement des études qui prennent cette source pour *medium* ou pour objet.

1.1.2.2. Les facteurs qui accentuent le retard de développement de l'étude de la presse régionale de la Troisième République

Paradoxalement, c'est ce qui confère à la PQR3R son identité et son intérêt historique, qui explique partiellement pourquoi son étude a été, et reste, malaisée.

L'essor de la presse sous la Troisième République s'est tout d'abord traduit par l'émergence d'un nombre considérable de feuilles à la plus ou moins grande longévité dans chaque département de France. Ainsi P. Albert formule l'hypothèse que le retard historiographique accumulé en histoire de la presse de la Troisième République est à relier au volume et à l'hétérogénéité des publications de cette période :

Peut-être est-ce la difficulté de dominer un ensemble aussi étendu et hétérogène qui a retardé les études historiques sur la presse (Albert, 1965 : 37).

Comme le déplore J. Kayser (1957 : 285), le volume des publications de la Troisième République est tel qu'il a rendu difficile toute création d'un inventaire général des feuilles de cette période avant la seconde moitié du XX^e siècle. Ce n'est qu'après la Seconde guerre mondiale que sont établis les inventaires généraux de la PQR3R.

À l'ampleur du défi à relever, s'ajoutait, sur le plan méthodologique, la nécessité de consolider la scientificité de ces inventaires, comme l'a souligné J. Kayser à la fin des années 1950. Pour que de tels inventaires puissent répondre aux besoins de la communauté, l'auteur (1957) met en lumière la nécessité de renouveler leur méthode de construction. Afin d'accroître la pertinence et la précision de ces inventaires, il propose de construire les inventaires à partir d'un panel de 75 points, relevant d'un travail de dépouillement du « contenu » du journal (présence ou non d'articles de fond, énumération des rubriques, articles de politique internationale ou non, etc.) ainsi que d'un travail de recherche visant à obtenir des renseignements relatifs au tirage de la feuille, à son influence et à ses dirigeants. Comme le précise Daniel J. Grange (1999), l'historien de la presse se heurte dans ce cadre à la fréquente indisponibilité des archives des rédactions et des maisons d'édition d'une part, et d'autre part, à l'état lacunaire des collections des journaux :

[...] que dire de l'état des collections de journaux, plus que lacunaires ; des archives des maisons d'édition détruites ; des papiers des rédactions, disparus. Comment après cela mesurer avec précision l'audience d'un journal, dessiner

sa zone d'influence, typer son lectorat, reconstruire sa ligne éditoriale ou l'état de ses finances? (Grange, 1999 : 7).

Répertorier les publications régionales et locales, retracer l'histoire du quotidien sous ses aspects matériels, socio-culturels et politiques, ne se révèle cependant que le préalable à la levée des difficultés majeures qui pèsent sur l'étude de la PQR3R. Si les inventaires permettent en effet à l'historien de connaître les publications locales et régionales pouvant constituer une « source » potentielle, ils ne lui indiquent nullement quel journal traite de la thématique ou de l'événement qu'il souhaite aborder. Ils lui permettent encore moins de savoir dans quel(s) numéro(s) du quotidien et à quelle(s) page(s) se localisent les articles traitant du sujet qui l'intéresse. En d'autres termes, la question du dépouillement et de l'indexation demeure problématique. Généralement, l'historien, pour cibler plus ou moins efficacement sa recherche et limiter le nombre de numéros des quotidiens à dépouiller, établit une chronologie du sujet à traiter sur la base de ses savoirs sur l'objet à analyser. Cependant, cette « méthodologie » se révèle insuffisante selon la nature du sujet que l'historien souhaite étudier :

[...] certains sujets d'histoire des idées au sens large ou d'histoire sociale qui touchent plus des thèmes diffus dans l'actualité que des événements se prêtent mal à l'établissement d'une telle grille de dates, mais il faut, dans ce cas aussi, chercher à deviner à quel moment les journaux ont pu traiter le sujet choisi (Albert, 1965 : 36).

Cette méthodologie rudimentaire est doublement contestable : d'une part, l'analyse qui en découle ne peut nullement prétendre à la systématisme et l'exhaustivité ; d'autre part, en renvoyant l'historien à des jugements de savoir, elle limite de fait le renouveau de l'état des connaissances historiographiques du sujet :

Il ne faut pourtant pas se cacher que cette méthode « par sondage », pour pratique qu'elle soit, n'est pas entièrement satisfaisante. C'est, en effet, souvent en période de creux de l'actualité que les journalistes écrivent les articles de doctrine ou d'analyse les plus originaux et les plus réfléchis. Malheureusement, il n'existe aucun moyen de les retrouver a priori, et ils se cachent souvent dans les pages intérieures du journal (Albert, 1965 : 36).

Il convient également d'insister sur le fait qu' un dépouillement « à l'œil nu » de la presse constitue une tâche écrasante, participant à faire de l'étude de la presse du XIX^e siècle, qu'elle soit nationale ou régionale, *une école de patience et de défiance* (Melmoux-Montaubin, 2001 : 327-328). L'ampleur de la tâche est accrue par la déshérence des dépouillements effectués, liée à l'hétérogénéité des protocoles existants dans la communauté historique, comme le remarque M.-E. Thérenty :

[...] chacun est contraint de refaire le chemin parcouru par d'autres, de bricoler ses propres instruments, de reprendre à partir de zéro des collections déjà étudiées par ses prédécesseurs. Cette situation assez absurde conduit à une immense déperdition d'énergie, et fait naître le sentiment que l'étude des périodiques est infiniment moins rentable — et gratifiante, à tous égards — que, par exemple, celles des grandes œuvres canoniques (Thérenty, 2001 : 345).

Si ce constat est établi au sujet de l'étude de la presse nationale du XIX^e siècle, il nous semble qu'il reste pertinent pour les recherches sur la presse régionale de la Troisième République.

Au-delà des difficultés méthodologiques que nous venons d'énumérer, il est essentiel de mettre en perspective le manque de systématisme des études menées *par* et *sur* la presse régionale de la Troisième République avec l'évolution de l'histoire en tant que discipline. Si l'on dénombre si peu d'études prenant la PQR3R pour source, c'est tout d'abord parce que la presse, qu'elle soit nationale ou régionale, est, jusqu'à la seconde moitié du XX^e siècle, une source dépréciée par les historiens :

Tant que l'histoire se cantonnait, pour l'essentiel, à l'étude des institutions, des doctrines et à la présentation des événements de la grande actualité politique, elle n'avait que peu à demander à la presse qui n'était guère qu'une source accessoire, complémentaire (Albert, 1975 : 40).

Ce n'est qu'à partir des années 1950 que la presse devient un matériau privilégié pour les historiens du politique qui la convoquent comme *medium* pour restituer les évolutions des opinions publiques. Si l'on en croit l'historien P. Albert (1975), de la fin du Second Empire à la fin de la Troisième République, la presse est dépréciée au profit du livre jugé comme la « seule

source imprimée digne d'intérêt » (Albert, 1975 : 39). Loin de faire état d'une simple opposition entre deux types de matérialités que sont le livre et le périodique, la réflexion de P. Albert nous semble devoir être interprétée en des termes plus complexes.

L'opposition du livre et du journal nous semble tout d'abord être sous-tendue par la problématique du statut de l'auteur, de son autorité et de l'influence postulée de son discours sur la société. Le livre serait un ouvrage canonique, supposé avoir été rédigé par un auteur, fournissant *via* cette matérialité un discours autorisé, auquel l'historien attribue une influence sur les débats culturels, économique, socio-politique d'une société. Le journal, lui, ne serait abordé que comme la production d'un *publiciste*, une voix de second ordre dont l'impact sur l'espace public a très longtemps été minimisé par les historiens.

Ensuite, il nous semble que la dépréciation du journal est à relier à la temporalité du discours journalistique. Aux antipodes du temps long de l'historien, le discours du journal, fondamentalement éphémère, ne transmet que la perception immédiate d'un événement, qui contraste avec la distanciation visée par l'historien. Ces raisons nous semblent éclairer pourquoi le discours du quotidien a longtemps été perçu comme une source archivistique anecdotique pour l'historien :

Ils [les historiens] ne trouvaient, dans tous les cas, dans les journaux qu'un reflet des problèmes politiques qui retenaient leur attention et n'éprouvaient pas le besoin de confronter leurs sources traditionnelles avec ces « archives du présent », à leurs yeux trop imparfaites et secondaires (Albert, 1975 : 39).

Ces remarques nous semblent valoir d'autant plus pour la presse régionale, qui dès ses origines, a été dévalorisée par ses contemporains, qui la percevaient comme une presse d'intérêt second et de qualité très inférieure à celle des quotidiens parisiens.

À cette étape de notre recensement des travaux menés en sciences humaines sur la presse quotidienne régionale de la Troisième République, il convient de constater que cet objet reste peu exploré, malgré son intérêt

historique. Imputable d'évidence aux conditions d'accès à cet objet et plus largement aux lourdes difficultés méthodologiques que son étude soulève, le faible nombre des recherches qui l'investissent est également à éclairer à la lumière de l'évolution de l'historiographie française et de son rapport aux objets langagiers. Nous rendrons ainsi compte des changements induits par le *linguistic turn* dans les années 1970 au niveau international pour ensuite nous focaliser sur son impact sur le champ historiographique français.

1.2. Le discours et sa dimension matérielle

En histoire, à l'instar de l'évolution générale dans les sciences humaines, l'attention portée au discours et à sa dimension matérielle s'est accrue au cours de la seconde moitié du XX^e siècle.

Après avoir replacé cette mutation épistémologique dans son contexte, nous présenterons les grandes directions de travail d'une histoire linguistique, en faisant état des voies ouvertes par *l'histoire des concepts* et *l'analyse du discours du côté de l'histoire* pour ensuite envisager ses liens avec l'analyse textuelle du discours.

Enfin, nous ferons état des différentes entreprises scientifiques menées pour satisfaire aux besoins de ces champs de recherches en attente de ressources textuelles historiques.

1.2.1. Regards sur le *linguistic turn*

Il convient tout d'abord de situer l'évolution du statut du texte dans le contexte englobant de la crise de la modernité dont prend acte la fin des années 1960. Dans le domaine de l'histoire, cette crise se traduit par l'émergence du débat sur le *linguistic turn*.

C'est dans le titre de l'ouvrage du philosophe américain Richard Rorty, *The linguistic Turn. Recent Essays in Philosophical Method*, publié en 1967, que l'on rencontre pour la première fois ce terme. Dans le champ historique, l'étiquette

de *linguistic turn* ne s'impose cependant qu'au début des années 1980, rassemblant un ensemble de chercheurs pour qui l'histoire doit s'intéresser au langage, à la « textualité » ou au discours. Selon G. Noiriel, s'attachant à retracer l'histoire du *linguistic turn*, les théoriciens de ce mouvement se réunissent autour du postulat selon lequel :

[...] toute réalité est médiatisée par le langage et les textes, donc toute la recherche historique est dépendante de la réflexion sur le discours (Noiriel, 1996 : 167).

Parce que sa paternité ne peut être attribuée à aucun fondateur précis et qu'il n'a pas de programme, le *linguistic turn* ne peut, en toute rigueur, être envisagé comme une école ou un courant de pensée (Müller, 2005) mais comme une réaction historiographique peu homogène, renvoyant à la formation d'un groupe d'historiens et de spécialistes de la littérature. Si, comme le signale C. Topalov (2008 : 173), une enquête historique sur ce mouvement reste à écrire, les historiens⁵ s'accordent à considérer le colloque tenu à l'Université Cornell en mars 1980 comme l'acte fondateur du *linguistic turn* et les actes de ce colloque (La Capra & Kaplan, 1982) comme son manifeste.

Pour clarifier les enjeux de cette controverse historiographique, nous nous attacherons à resituer les arguments d'un débat véhément qui n'a été que tardivement mis sur le devant de la scène en France (Noiriel, 1996, 1998 ; Müller, 2005), en insistant sur les résultats des recherches menées dans ce cadre et son influence sur les sciences historiques actuelles.

1.2.1.1. Une remise en cause des moyens de la connaissance historique

Le *linguistic turn*, émergeant dans le courant des années 1980 aux Etats-Unis et en Angleterre, et plus particulièrement dans le cercle des

⁵ dont G. Noiriel (1996 : 160)

européanistes, figure une contestation radicale des méthodes et des présupposés de la recherche historique qui prévalent jusqu'alors.

Il s'agit tout d'abord d'une réaction aux prétentions totalisantes d'une histoire culturelle abordant le fait social au travers d'un modèle explicatif causal et structural, reposant sur l'étude des infrastructures économiques et sociales. Le *linguistic turn* se cristallise également autour du refus des théories dites déterministes. Ce sont plus particulièrement les structures sociales pré-établies conçues par le marxisme, jugées inadaptées à l'appréhension des phénomènes sociaux, dont les théoriciens du tournant linguistique se proposent de faire table rase. Ces refus doivent être rapportés au contexte d'émergence du *linguistic turn*, qui est celui d'une prise de conscience, dans les milieux intellectuels de gauche, des dérives du stalinisme. L'un des théoriciens les plus importants du *linguistic turn*, L. Stone (1979) dans son article « Revival of Narrative » considéré comme le manifeste du retour au récit, explique ainsi l'origine du « tournant narratif » par cette désillusion :

The first cause of the current revival of narrative is widespread disillusionment with the economic determinist model of historical explanation (Stone, 1979 : 8)⁶.

Cette crise du modèle marxiste altéré par Staline s'accompagne inmanquablement d'une contestation de la notion d'histoire scientifique qui avait été défendue par certains marxistes dès la fin des années 1950. C'est dans cette optique que les partisans du *linguistic turn* critiquent de façon virulente l'histoire quantitative et dénoncent l'incapacité des mathématiques et des statistiques à répondre aux objectifs de la recherche historique :

Despite its unquestionable achievements it cannot be denied that quantification has not fulfilled the high hopes of twenty years ago. Most of the great problems of history remains as insoluble as ever, if not more so. [...] Quantification has

⁶ « La cause première de la résurgence actuelle du narratif est la désillusion généralisée que suscite le modèle économique déterministe de l'explication historique » (Stone, 1979 : 8, notre traduction).

told us a lot about the *what* questions of historical demography, but relatively little so far about the *why* (Stone, 1979 : 12)⁷.

Le *linguistic turn* est également à comprendre comme une réaction envers une histoire sociale qualifiée d'« empirique », accusée de ne pas tenir compte du fait que la réalité qui est accessible à l'historien est immanquablement de nature langagière. Les travaux de Gareth Stedman Jones (1983) visent à fournir des propositions méthodologiques et épistémologiques dépassant les pièges des modèles interprétatifs antérieurs (Jones, 1983 : 24). Dans *Language of Class*, publié en 1983 mais reprenant une série d'articles publiés à partir de 1975, l'auteur se montre convaincu de l'inadéquation des approches empiriques (*empiricist approaches*) pour expliquer les faits sociaux. Il pose que l'on ne peut appréhender le social hors du langage. Au rebours d'une histoire sociale traditionnelle, G.S. Jones se fixe pour objectif d'apporter un éclairage sur l'histoire des ouvriers anglais au début du XIX^e siècle, en se focalisant non plus sur l'expérience de classe mais sur les langages de classe. En rupture avec l'approche de son contemporain E. Thompson (1963, trad. fr. : 1968) selon lequel la classe ouvrière devrait être envisagée comme un produit de l'histoire, du fait d'un processus selon lequel l'expérience d'une condition partagée se transformait en « conscience », G. S. Jones affirme qu'il n'existe aucune réalité, aucune expérience, en dehors du langage qui la structure. Plus, le langage ne réfère pas à une réalité, il l'organise. Aussi G. S. Jones, postulant que la structure discursive du langage politique définit l'intérêt de classe et non l'inverse, cherche à décrire la production de l'intérêt de classe au sein même de ses conditions langagières de formation :

Language disrupts any simple notion of the determination of consciousness by social being because it is itself part of social being. We cannot therefore decode political language to read a primal and material expression of interest since it is

⁷ « Malgré des réussites indéniables que l'on ne peut remettre en question, il faut bien dire que les méthodes quantitatives n'ont pas été à la hauteur des hauts espoirs que l'on plaçait en elles vingt ans auparavant. La plupart des grands problèmes de l'histoire restent tout aussi insolubles qu'ils pouvaient l'être, si ce n'est plus. Les méthodes quantitatives nous ont appris beaucoup quant à la question du *comment* de la démographie historique, mais relativement peu quant à la question de son *pourquoi* » (Stone, 1979 : 12, notre traduction).

the discursive structure of political language which conceives and defines interest in the first place. What we must therefore do is to study the production of interest, identification, grievance and aspiration within political languages themselves. We need to map out these successive languages of radicalism, liberalism, socialism etc., both in relation to the political languages they replace and laterally in relation to the rival political languages with which they are in conflict (Jones, 1983 : 21-22)⁸.

1.2.1.2. La controverse sur l'objectivité et la vérité

On ne peut par ailleurs comprendre les polémiques relatives au *linguistic turn* sans mentionner la thèse de H. White. Son approche des procédés d'écriture des historiens a sans nul doute contribué à diviser la communauté historique et à masquer les véritables propositions et les conséquences du *linguistic turn*. S'inspirant d'une réflexion développée par R. Barthes (1960), H. White, affirme que l'histoire, tel un genre littéraire, est fiction :

There has been a reluctance to consider historical narratives as what they most manifestly are: verbal fictions, the contents of which are as much invented as found and the forms of which have more in common with their counterparts in literature than they have with those in the sciences (White, 1978 : 82)⁹.

Captif du mode langagier dont il relève, le récit de l'historien est fiction. Partant, H. White dénie au discours de l'historien toute prétention à accéder à la réalité, remettant en cause le statut de discipline scientifique de l'histoire :

[...] in any field of study not yet reduced (or elevated) to the status of genuine science, thought remains the captive of the linguistic mode in which it seeks to

⁸ « Le langage bouleverse toute notion simple de la détermination de la conscience par l'être social parce qu'il est lui-même partie prenante de cet être social. Nous ne pouvons donc pas décoder le langage politique pour y lire une expression primaire et matérielle d'un intérêt dans la mesure où c'est la structure discursive du langage politique qui conçoit et définit l'intérêt, de l'identification, du ressentiment et des aspirations à l'intérieur des langages politiques eux-mêmes. Il nous faut identifier ces langages successifs que sont le radicalisme, le libéralisme, le socialisme, etc. à la fois en relation avec les langages politiques qu'ils remplacent et littéralement en relation avec les langages politiques rivaux avec lesquels ils entrent en conflit » (Jones, 1983 : 21-22, notre traduction).

⁹ « Il y a eu de la réticence à considérer les narrations historiques comme ce qu'elles sont le plus manifestement : des fictions verbales dont le contenu est tout autant inventé qu'avéré, et dont les formes ont plus en commun avec leurs contreparties en littérature qu'en science » (White, 1978 : 82, notre traduction).

grasp the outline of objects inhabiting its field of perception (White, 1978 : 111).¹⁰

Postulant que le discours historique est déterminé par des structures qui excèdent la volonté et la conscience des historiens, le propos de H. White est de formaliser les matrices de l'imagination historique. Dans cet objectif, il identifie tout d'abord les niveaux de mise en forme dont la combinaison détermine les styles historiques : il convoque à cet effet quatre catégories, à savoir les modes de mise en intrigue (*romanesque, tragique, comique, satirique*), les différents paradigmes de l'explication historique (*formiste* (« *formism* »), *mécaniste, organiciste, contextualiste*) et les différentes positions idéologiques (*anarchiste, radicale, conservatrice, libérale*). Cependant, l'identification de ces catégories n'est que le préalable à l'éclairage du « niveau plus profond » où se joue la validation implicite et théorique de ces opérations. H. White ambitionne en effet de révéler les préfigurations linguistiques et poétiques du champ historique par lesquelles l'historien crée son objet d'analyse et qui prédéterminent la modalité des stratégies conceptuelles mobilisées. Ces présupposés sont définis à l'aide des quatre tropes classiques du langage poétique : la métaphore, la métonymie, la synecdoque et l'ironie. En attaquant ainsi les frontières entre fiction et histoire, la position relativiste de H. White a suscité de nombreuses polémiques. Elle équivaut à poser qu'il n'y a pas de vérité définitive en histoire, ce qui ouvre la voie aux dérives négationnistes, danger que signale G. Noiriel (1998 : 124). Le débat instauré par H. White autour de la réalité et du discours a eu de lourdes conséquences sur l'héritage du *linguistic turn* en France, comme le souligne J. Guilhaumou :

Focalisant leur attention sur « le tournant linguistique » à l'américaine d'Hayden White qui semble ne voir dans le discours de l'historien qu'une expression parmi d'autres de l'invention fictionnelle, ils [les chercheurs français] craignent

¹⁰ « Dans tous les domaines d'études qui ne soient pas encore réduits (ou élevés) au rang de science véritable, la pensée reste captive du mode linguistique dans lequel elle cherche à saisir les contours des objets occupant son domaine de perception » (White, 1978 : 111, notre traduction).

son pouvoir de subversion des règles du métier d'historien (Guilhaumou, 2006a : 44).

Postulant, dans une version forte, que toute réalité ne peut être pensée en dehors de la médiation du langage, le *linguistic turn* introduit dans la pratique historique un renouvellement de l'importance accordée au texte. En ce sens, le *linguistic turn* peut être envisagé et traduit comme un « tournant textuel », tout spécialement si l'on insiste sur les nouvelles alliances de l'histoire avec les sciences du langage qu'il initie.

1.2.1.3. *L'histoire des concepts*

Affiliée au *linguistic turn*, l'*histoire des concepts*, bien distincte de l'histoire des idées, constitue un champ de recherche dont l'influence n'a cessé de croître depuis 1980. Les lignes suivantes seront consacrées à une présentation très synthétique de cette perspective telle qu'initée en Allemagne par R. Koselleck et en Angleterre par l'école de Cambridge.

L'histoire des concepts (R. Koselleck)

Connu essentiellement en France par la traduction somme toute récente de ses ouvrages *Le futur passé. Contribution à la sémantique des temps historiques* (Koselleck, 1990) et *L'Expérience de l'histoire* (Koselleck, 1997), R. Koselleck est à l'origine d'une méthodologie d'histoire des concepts socio-politiques, l'histoire sémantique, dont on trouve une réalisation magistrale dans l'entreprise dictionnaire *Geschichtliche Grundbegriffe : Historisches Lexikon zur politisch-socialer Sprache in Deutschland* (Les concepts fondamentaux en histoire : Dictionnaire historique du langage politique et social en Allemagne), ouvrage de plus de 7000 pages qu'il aura fallu vingt ans à faire aboutir.

Selon R. Koselleck, les notions-concepts de l'âge préindustriel ne peuvent être comprises par le lecteur du XXI^e siècle qu'au prix d'un important travail de recontextualisation. Les changements sémantiques profonds des notions-concepts sont imputables à une triple mutation structurelle et fonctionnelle. L'*historicisation* (*Verzeitlichung*) des notions-concepts est ce

processus sémantique, dont les origines sont à trouver dans les effets de la Révolution française et de la révolution industrielle, qui contribue à enrichir les notions-concepts d'une dimension anticipatrice. La *généralisation* désigne pour sa part le mouvement par lequel se trouve modifiée la charge sémantique d'une notion-concept : au fur et à mesure que son usage s'étend au-delà des frontières des groupes sociaux, sa signification devient de plus en plus diffuse et malléable à la perspective qui l'investit. D'où une *politisation* croissante des notions-concepts : celles-ci sont les lieux et les enjeux du combat pour l'histoire que mènent les forces politiques en présence.

Partant, R. Koselleck pose que l'histoire sociale ne peut faire l'économie d'une histoire sémantique, s'attachant à observer l'usage des concepts que crée ou s'approprié une société en vue de se comprendre elle-même et qui la conditionnent en retour. Plus que le simple reflet du fait social en train de se produire, la notion-concept, selon R. Koselleck, rétroagit sur la société :

Une notion n'est pas seulement l'indicateur de l'ensemble d'éléments qu'elle saisit, elle en est également un facteur constitutif. Avec chaque notion s'établissent des horizons déterminés, mais également des frontières de l'expérience potentielle et des limites concevables de la théorie. C'est pourquoi l'histoire des notions peut apporter des connaissances que l'analyse des faits matériels ne prend pas en considération (Koselleck, 1979 : 345 cité par Reichardt, 1982 : 195).

H.S. Lüsebrink, R. Reichardt et E. Schmitt prolongent en Allemagne les travaux de R. Koselleck et sont investis, en collaboration avec certains historiens français tels que M. Vovelle, R. Chartier et J. Guilhaumou, dans une autre entreprise « dictionnaire » gigantesque depuis 1985 : celle-ci porte sur les concepts politiques et sociaux fondamentaux en France de 1680 à 1820 appréhendés dans une archive plus diversifiée que celle convoquée par R. Koselleck. Très justement, R. Reichardt (1982 : 199) remarque qu'une histoire des mots-thèmes socio-politiques ne peut se fonder exclusivement sur les différents dictionnaires d'une époque. S'ils constituent des ressources précieuses, les dictionnaires ne prennent que tardivement acte des sens nouveaux. De plus, ils n'en proposent qu'un compte-rendu succinct.

R. Reichardt pose ainsi comme pré-requis nécessaire à toute entreprise dictionnaire de s'ouvrir « à d'autres sources, moins définissantes [...] mais où certaines notions-clefs sont souvent employées, soit pour informer, soit pour inciter l'action » (Reichardt, 1982 : 199). Ces autres sources sont notamment les « mémoires-journaux », les procès-verbaux de sociétés populaires de province, les « livres de raison » ainsi que les almanachs et les chansons, offrant une perspective originale et complémentaire pour « pénétrer les concepts stéréotypés des petites gens tant à Paris qu'en province » (Reichardt, 1982 : 200), mais aussi les images. La presse écrite constitue dans ce cadre une ressource essentielle, car selon R. Reichardt, s'inspirant de J. Sgard (1981), le contenu des journaux est représentatif du discours du groupe social que constitue son lectorat :

Dans la communication périodique, l'offre doit donc s'aligner perpétuellement sur la demande ; un contact permanent s'établit entre l'auteur et son public, et c'est dans une certaine mesure le public qui décide du contenu du journal : le journal est une création collective, l'expression d'un groupe social (Sgard, 1981 : 20).

Si la valeur générale de cette constatation est bien entendu à discuter et à nuancer, notamment en fonction de la période de parution du journal et de son tirage, elle a, selon nous, le mérite de pointer l'adaptation du discours du journal à son lectorat.

Autre caractéristique innovante de ce *Dictionnaire des mots-thèmes socio-politiques en France de 1680 à 1820* dont l'orientation pragmatique et interculturelle est prononcée : le projet ne vise pas à la constitution d'un *dictionnaire* de manière à dépasser la structure statique de ce type d'ouvrages¹¹, mais à la publication d'un manuel (*Handbuch*) mettant en réseau les concepts sans lien alphabétique.

¹¹ Voir à ce sujet Guilhaumou (2006 : 50).

L'école de Cambridge

Celle-ci partage avec les travaux de R. Koselleck le fait d'appréhender l'histoire comme construction sociale. L'école de Cambridge s'est construite autour des figures de J.G.A. Pocock (1971, 1975 (trad. fr 1997)) et Q. Skinner (1996 (trad.fr 2001)). Elle constitue le second courant dominant de l'histoire des concepts. Dans le cadre de cette brève présentation, nous nous concentrons sur les travaux de Q. Skinner, dont le rayonnement est plus intense que celui des travaux, pourtant antérieurs, de J.G.A. Pocock.

Aux traditions d'interprétations des historiens déterministes, aux lectures internes du texte pensé comme système et aux reconstructions téléologiques de l'histoire, Q. Skinner oppose la contingence des faits historiques et la nécessité d'une histoire du discours politique *historicisant* la pensée politique en s'attachant à prendre en compte le contexte. Ses recherches relèvent d'un projet de renouvellement des lectures des textes de philosophie politique du passé, dont l'historicité est, selon lui, trop fréquemment négligée. Ce programme explique que l'attention de l'auteur porte exclusivement sur les textes typés comme canoniques.

Selon Q. Skinner, toute lecture d'un texte du passé doit procéder d'un travail de mise en contexte historique, entendu comme un réseau de discours, une matrice textuelle et non comme un simple cadre événementiel. Inspiré par Wittgenstein et Austin, Q. Skinner focalise son attention sur « ce que les auteurs [...] ont fait en écrivant comme ils l'ont fait » (Guilhaumou, 2006a : 53). L'historien du discours caractérise les concepts-politiques comme des actes de langage dont la dimension agissante et l'usage argumentatif ne peuvent être appréhendés que par une remise en contexte :

[...] comprendre les questions qu'affronte un auteur et ce qu'il fait des concepts dont il dispose équivaut à comprendre ses intentions premières dans l'acte d'écrire, et consiste donc à élucider ce qu'il aurait vraiment voulu dire dans ce qu'il a dit – ou n'a pas dit. En cherchant à situer ainsi un texte dans le contexte qui lui revient, ce n'est pas seulement un décor qui se propose à l'interprétation : c'est l'acte d'interprétation lui-même qui commence (Skinner, 2001 : 13).

Aussi, son approche de l'œuvre de T. Hobbes (Skinner, 1996) peut-elle être envisagée comme une « mise en contexte » de la pensée de l'auteur, dont les concepts sont resitués dans l'environnement historique, politique ainsi que rhétorique prévalant en Angleterre au XVII^e siècle. La première partie de l'étude de l'œuvre de Hobbes est ainsi celle de son contexte : il s'agit, pour Q. Skinner, de décrire la culture rhétorique de l'humanisme de la Renaissance et de montrer le rôle majeur qu'elle joue dans l'évolution des positions de T. Hobbes quant à la construction du concept de *société civile*.

1.2.2. L'analyse du discours du côté de l'histoire

En France, l'intérêt des historiens pour le champ linguistique, perceptible dans les travaux de l'École des Annales, est antérieur au *linguistic turn*. On mentionnera tout particulièrement l'appel à la rencontre de la linguistique et de l'histoire émis par L. Febvre (1953), qui met en lumière l'urgence pour l'historien de disposer d'une histoire des « mots-thèmes » :

Langage, cette autre voie cardinale d'accès au social dans l'individu ? Il faut la coopération des philologues, dressant des inventaires de langues qui ne sont pas faits pour les historiens, mais dont ceux-ci peuvent tirer un si grand parti : [...] Non moins, il faut la collaboration de ces « sémantismes » qui, en nous restituant l'histoire de mots particulièrement lourds de sens, écrivent du même coup des chapitres précis d'histoire des idées. Il faut la collaboration de ces historiens des langues, tel Meillet écrivant l'histoire de la langue grecque, tel Ferdinand Brunot suivant pas à pas les destinées de la langue française, qui notent l'apparition à certaines dates de tout un contingent de mots neufs ou de sens nouveaux donnés à de vieux mots¹².

Le paradigme du discours comme objet de l'histoire, dont prend acte la publication de *Histoire et linguistique* (1973) et celle de l'ouvrage collectif *Langage et idéologies* (Guilhaumou et al, 1974), marque cependant l'émergence d'une « figure inédite » de l'historien du discours (Guilhaumou, 2002c : 294). Opérant un tournant épistémologique majeur à l'heure où le structuralisme est le paradigme dominant, l'analyse du discours du côté de

¹² Febvre, L., 1965 (1953), « Histoire et psychologie » dans *Combats pour l'histoire*, Paris, A. Colin, p. 219, cité par Robin (1973 : 46).

l'histoire pose la nécessité de considérer le discours dans ses liens avec l'ensemble des rapports extralinguistiques qui le structure :

Nous cherchons à constituer dans le champ de l'histoire, le niveau discursif comme nouvel objet d'étude, et en même temps, contre un certain positivisme linguistique, à postuler les impasses et les insuffisances d'une analyse interne de laquelle on voudrait déduire l'interprétation du discours y compris sa fonction, son efficace et ses processus d'insertion dans sa formation sociale (Robin, 1973 : 22).

L'exigence des rapprochements entre histoire et linguistique se justifie conjointement, dans le cadre de l'analyse du discours, par l'*historicité* du discours et par la nécessité de dépasser une lecture référentielle prévalant dans la pratique historique, se limitant à observer le contenu des pratiques langagières sans en examiner les formes :

En somme, à la question habituellement posée à propos du texte : que dit-il ? nous substituons trois autres questions ambitieuses sans doute : comment est dit ce qui est dit ? Comment le discours fonctionne-t-il ? Pourquoi ce fonctionnement est-il ce qu'il est ? (Maldidier, Robin, 1974 : 13)

Dénonçant le postulat de l'*immédiateté du sens* et de la *transparence des mots* (Robin, 1973 : 38) qui conditionne d'ordinaire le geste de lecture des historiens, l'analyse du discours du côté de l'histoire affirme que toute analyse confrontée à un matériel verbal ne peut se dispenser d'une analyse linguistique.

1.2.2.1. Retour sur la démarche inaugurale de l'analyse du discours du côté de l'histoire

En réaction à la « misère méthodologique » des historiens (Guilhaumou, Maldidier, Robin, 1994 : 176), l'analyse du discours « du côté de l'histoire » s'est construite dans les années 1970 autour d'un modèle de scientificité inspirée de la linguistique distributionnelle américaine. De la tradition distributionnelle, elle adoptait en premier lieu les principes méthodologiques fondamentaux d'exhaustivité, de systématisme et de formalisation de l'analyse.

Le corpus comme moment inaugural

À l'instar des linguistiques descriptivistes, l'analyse du discours du côté de l'histoire postule un moment inaugural : la constitution du corpus. Après avoir découpé des énoncés dans la totalité des discours d'une conjoncture historique et/ ou d'un groupe social donné, l'analyste sélectionne dans un deuxième mouvement un ensemble plus restreint d'énoncés sur la base de la présence en leur sein d'un ou plusieurs mot(s)-pivot, posé(s) *a priori* comme étant stratégique(s). R. Robin et D. Maldidier (1974), dans le cadre de leur étude des *Remontrances parlementaires face aux Edits de Turgot* de 1776, organisent ainsi leur corpus autour des mots-pivots *liberté* et *règlements*, très précisément en référence à un savoir historiographique selon lequel ces notions cristallisent l'opposition entre noblesse et bourgeoisie dans la conjoncture de la tentative réformatrice de Turgot.

On peut dès lors discuter la démarche d'une analyse du discours du côté de l'histoire. Tel que construit par cette analyse du discours partageant bien des traits communs avec la sociolinguistique, le corpus pose une série de difficultés méthodologiques majeures, qui entrent en conflit avec le modèle de scientificité pourtant visé par l'analyse du discours du côté de l'histoire. Échantillon supposé représentatif d'une réalité discursive, le corpus, homogénéisé sur la base des conditions de production du discours, n'est qu'une mise en série d'énoncés ancrant irréversiblement l'analyse dans les *jugements de savoir* de l'analyste. Outre le fait de limiter le potentiel de renouvellement des savoirs historiographiques, cette méthodologie pose la question du statut de l'analyse menée : toute analyse menée à partir d'un tel corpus n'est-elle pas vouée à n'être, pour reprendre l'expression de F. Mazière (2005 : 20), qu'une « simple illustration-justification » du savoir de l'historien ?

En raison de sa clôture sur lui-même, le corpus conduit à une analyse où le sens est éclairé à partir d'un hors-corpus. Cette rupture entre un corpus, dont la taille est d'ailleurs fortement restreinte, et le hors-corpus est fortement contestable. Plus, l'analyse de ce corpus homogénéisé à partir de la notion de

conditions de production, pré-supposant des entités discursives stables, se réduit à mettre brutalement en relation les faits linguistiques avec les positionnements sociaux, risque tautologique dénoncé par J. Guilhaumou et D. Maldidier dès 1979. De surcroît, le geste d'arrachement d'énoncés à leurs contextes est aussi arrachement à leurs co-textes. L'analyste étudie des énoncés et non une série de *discours suivis* (Harris, 1969 : 11), d'où une neutralisation de la textualité, dès lors inaccessible. L'analyse du discours du côté de l'histoire recourant, dans sa version forte, à la méthode harrissienne des énoncés, accentue d'autant plus cette « dénaturation » de la dimension matérielle du discours : appliquer des règles d'équivalence grammaticale permettant d'obtenir un ensemble de phrases transformées, revient à construire des régularités là où le discours n'était qu'aspérité, à négliger ses matérialités originelles pour lui substituer une construction de l'analyste. Il reste toutefois à préciser que la méthodologie harrissienne, exigeant de solides connaissances linguistiques, ne semble néanmoins pas avoir connu un fort retentissement chez les historiens. En revanche, ces derniers semblent avoir été plus réceptifs aux apports que pouvait leur offrir l'outillage lexicométrique.

Analyse du discours du côté de l'histoire et lexicométrie

L'analyse du discours du côté de l'histoire trouve dans la lexicométrie, également développée dans les années 1970, une méthodologie qui lui permet de prétendre aux exigences d'exhaustivité et de systématisme d'un traitement de la *surface discursive* d'un texte.

Ensemble de méthodes statistiques permettant de quantifier le vocabulaire, la lexicométrie, par la rigueur et la fiabilité des décomptes liées à une automatisation du traitement, offre à l'analyse du discours les moyens d'une approche scientifique du texte. Plus, la lexicométrie constitue ce dispositif méthodologique à fort potentiel heuristique permettant à l'analyste d'appréhender un volume de données conséquent au travers d'une lecture objectivée des saillances du corpus :

Pour la lexicométrie, l'objectivité commence quand les projections tendent à faire place à des formalisations. On ne fixe d'avance ni mots-témoins, ni grilles, ni schèmes, ni hiérarchies. On prend tout (exhaustivité des relevés), de la même manière (uniformité du dépouillement) et selon un critère unique (invariance d'une unité de base anonyme) (Demonet, Geffroy, Tournier et al, 1978 : 19).

Se limitant, en ses premiers temps, aux calculs de la valeur absolue des occurrences d'une forme et à leur soumission sous forme d'un index hiérarchique ou alphabétique, la lexicométrie se dote très rapidement d'une gamme élargie d'outils pour observer la matérialité textuelle. Des améliorations comme le calcul du poids relatif des formes par rapport à l'ensemble des occurrences du corpus et le calcul des spécificités lexicales d'un texte ou d'une partie du corpus facilitent la comparaison des textes ou des partitions d'un corpus contrastif. L'analyse factorielle, permettant d'obtenir des vues synthétiques de grands ensembles de données, montre pour sa part sa capacité à renouveler l'état des savoirs historiographiques. Notamment utilisée par A. Prost (1974a) dans le cadre de son étude du vocabulaire des proclamations électorales de 1881 et 1889, l'analyse factorielle des correspondances révèle les engagements électoraux voisins de J. Ferry et L. Gambetta malgré leur rivalité, mais également l'appartenance des boulangistes de 1889 à un centre droit démocratique et national malgré leur origine souvent radicale. La collaboration de l'analyse du discours et de la lexicométrie, apportant de probants résultats quant à la mise au jour de stratégies discursives, s'est également accompagnée d'améliorations méthodologiques conséquentes. Une des manifestations les plus évidentes de la fécondité de cette alliance est l'application au corpus du *Père Duschene* de la méthodologie d'analyse des « segments répétés » définis comme une suite récurrente de mots, élaborée par A. Salem (1984, 1987). Envisagé comme un dispositif d'observation de l'évolution du mouvement cordelier, le corpus du *Père Duschene* est composé de 96 numéros du journal pamphlétaire rédigé par Hébert, publiés du 17 juillet 1793 au 13 mars 1794. Dans un premier temps, J. Guilhaumou (1986) procède à une analyse lexicométrique à partir des formes

simples du corpus : les résultats de cette première expérimentation sont mitigés, en cela que la méthode de l'analyse factorielle des correspondances pointe un certain nombre de phénomènes intéressants mais ne livre aucun renseignement sur l'évolution chronologique au sein du corpus. Celle-ci se révèle grâce à l'utilisation conjointe de la méthode de l'inventaire des segments répétés et celle d'une analyse factorielle des correspondances basées sur un regroupement des numéros du *Père Duchesne* en mois. Une fois les 96 numéros du *Père Duchesne* regroupés selon leur mois de parution, l'analyse factorielle des correspondances laisse clairement apparaître une opposition entre le début et la fin de parution du journal. Le segment répété « tous les » se révèle fortement structurant de cette opposition sur le premier facteur de l'axe de l'analyse factorielle. Cette saillance en apparence banale invite l'historien à porter son attention sur les mots d'ordre du mouvement révolutionnaire, dont les multiformes « tous les » et « il faut » constituent les indices. Une analyse aboutie de ces saillances remontées du texte lui-même permet à J. Guilhaumou d'apporter un éclairage neuf sur la chronologie du mouvement cordelier, et notamment de préciser :

L'historiographie de la Révolution française situe le point culminant du mouvement cordelier pendant l'automne 1793. Nous ne partageons pas cette analyse. La prise en compte de la position des Cordeliers dans l'espace des mots d'ordre, où le *Père Duchesne* occupe une place centrale, permet d'affirmer le rôle hégémonique de cette tendance du mouvement jacobin dès l'été 1793 (Guilhaumou, 1986 : 35).

1.2.2.2. La démarche configurationnelle

Moins de dix ans après les débuts de l'analyse du discours du côté de l'histoire, celle-ci évolue vers une démarche configurationnelle et se constitue en « discipline herméneutique à part entière » (Guilhaumou, 2002a : 32).

Aux années 1980, correspond un contexte de déstabilisation et de refondation de l'analyse du discours lié à l'émergence des concepts d'*intertexte* (Bakhtine, 1977), d'*interdiscours* (Pêcheux, 1975) et d'*hétérogénéité* (Authier, 1979), dont prend acte le colloque interdisciplinaire « Matérialités discursives »

organisé par B. Conein, J.-J. Courtine, F. Gadet, J.-M. Marandin et M. Pêcheux, tenu à Paris X-Nanterre en avril 1980. Est alors formulé le constat indépassable des matérialités discursives surgissant « précisément de ce qui, entre l'histoire, la langue et l'inconscient, résulte comme hétérogénéité irréductible : un ressassement de paroles entendues, rapportées ou transcrites, un fourmillement d'écrits citant des paroles et d'autres écrits » (Pêcheux, 1981 : 15). Dans le même temps qu'il assigne à l'analyse du discours une vocation interdisciplinaire, ce constat remet en cause l'acte de lecture des faits de discours.

Dans ce contexte, le tournant que prend l'analyse du discours du côté de l'histoire dans les années 1980 se concrétise en premier lieu par une prise de distance avec la clôture et l'homogénéité du corpus. À la construction d'un corpus homogène et clos, succède la constitution de dispositifs d'énoncés hétérogènes et pluriels, marquant un tournant décisif dans le mouvement et le geste de lecture de l'analyse du discours du côté de l'histoire, désormais ouvert à la textualité. *A contrario* de la pratique de découpage d'énoncés au sein d'une archive dont la matérialité n'était pas interrogée, l'analyse du discours, prorogant le programme archéologique de M. Foucault (1969), se tourne vers la description des « règles de constitution des objets, de formation des concepts et des positions de sujets » (Foucault, 1994b : 162) à partir de configurations d'énoncés d'archive. L'archive, dans cette perspective, loin de se réduire à *l'ensemble des textes laissés par une société*, correspond à un *lieu de configuration* indissociable d'un geste de lecture « de facture herméneutique, donc foncièrement anti-constructiviste, c'est-à-dire basé sur les capacités réflexives, donc interprétatives, des acteurs « naturels » de l'histoire » (Guilhaumou, 2002a). La prise en compte de la réflexivité et la capacité interprétative des acteurs de l'histoire est synonyme d'un élargissement conséquent du champ d'investigation de l'analyse du discours. Elle marque la fin du primat du discours « politique » produit dans des cadres institutionnels ou fortement contraints comme objet d'analyse et une ouverture au discours des

« gens ordinaires ». L'analyse du discours se confronte dès lors à la diversité de l'archive, allant du « cri du peuple émeutier rapporté en justice » au « traité d'économie politique » en passant par la « correspondance des intendants », les affiches, les rapports de police, les pétitions, les correspondances entre clubs et sociétés populaires (Guilhaumou, Maldidier, 1986 : 43), etc.

La démarche configurationnelle se traduit par un bouleversement des temps de l'analyse : la construction du corpus n'est plus le temps inaugural de l'analyse et cède la place à un geste de lecture herméneutique de description des configurations d'archives significatives d'un sujet historique, d'un objet discursif ou de l'émergence d'un concept à l'intérieur d'un trajet thématique et de ses moments de corpus. Pour illustrer en quoi cette démarche permet de saisir une dimension de l'événement inédite et renouvelle l'accès au sens, nous nous reporterons à l'étude de J. Guilhaumou et D. Maldidier (1986) intitulée « Coordination et discours : « du pain et X » à l'époque de la Révolution française », dont F. Mazière (2005 : 115-117) souligne le caractère programmatique. Cette étude problématise et illustre de façon exemplaire une approche de la *matérialité de la langue dans la discursivité de l'archive*. Procédant à un parcours de la thématique des subsistances durant le XVIII^e siècle, J. Guilhaumou et D. Maldidier (1986, 1994) mettent au jour la récurrence originale de la coordination « du pain et X ». Dans un premier temps, les auteurs effectuent une analyse systématique de la structure à partir d'un corpus clos, réglé sur la base de la coordination de deux groupes nominaux dont l'un des segments est « du pain ». L'analyse grammaticale rigoureuse menée sur cet ensemble de données se heurte dès lors à la neutralisation de l'hétérogénéité des locuteurs et des genres du discours induite par la construction du corpus :

[...] nous saisissons un effet de construction sans pouvoir en rendre compte. Les classements grammaticaux ignorent ce qui se joue autour de nos coordinations. Ils sont aveugles au rejet ou à l'acceptation par tel ou tel protagoniste ; ils regroupent ce qui, dans le discours, s'oppose (Guilhaumou, Maldidier, 1994 : 138).

Dans un deuxième temps, J. Guilhaumou et D. Maldidier sont ainsi amenés à délaisser le corpus initial et son homogénéité problématique pour se tourner vers des dispositifs d'énoncés hétérogènes, aptes à rendre compte des enjeux discursifs de la structure syntaxique dans une perspective diachronique. Les historiens du discours montrent ainsi, dans la dispersion de l'archive, que la coordination, élément syntaxique, participe à l'émergence de nouveaux mots d'ordre dans les discours révolutionnaires et de nouveaux référents.

Alors que l'analyse du discours du côté de l'histoire, dans sa première version, éclairait la matérialité discursive par les conditions de production et induisait une coupure problématique entre le texte et son contexte, l'analyse configurationnelle du discours recentre le contexte dans l'énoncé :

[...] la dualité texte-contexte n'est guère pertinente dans l'analyse de l'événement discursif. En effet, le geste configurant, au regard de la multiplicité des énoncés, n'opère pas une mise en relation entre des faits textuels internalisés et des contextes externalisés, il permet par la lecture d'archive, de concrétiser l'existence discursive de possibles inhérent au substrat humain [...] Ainsi, une fois inséré, par la lecture d'archives, dans une configuration signifiante, l'énoncé est à lui-même son propre contexte (Guilhaumou, 2006a : 124).

Le corpus réflexif : le contexte dans le corpus

Les avancées récentes des capacités de stockage et de traitement des outils informatiques offrent le socle d'un renouvellement de l'approche du corpus, ainsi que l'exemplifie la thèse de l'historien D. Mayaffre (2000).

Le sujet de recherche de cet historien, c'est-à-dire le discours de gauche et de droite dans l'entre-deux-guerres, peut être considéré comme relativement classique (voire Courdresses, 1971). Son originalité est à trouver du côté de la démarche initiée tant à l'égard du corpus construit que de l'analyse *logométrique* adoptée, qui s'autorise le recours à d'autres entrées que celle de la surface graphique du vocabulaire. Réservant la présentation détaillée de l'approche logométrique à une étape ultérieure de notre travail, nous insisterons ici essentiellement sur les caractéristiques du corpus construit par D. Mayaffre et ses conséquences sur le geste d'interprétation.

Si le corpus construit répond de façon traditionnelle en lexicométrie aux critères de constitution que sont l'homogénéité, la contrastivité et la diachronicité, il ne constitue pas un échantillon restreint du discours à analyser. Ce corpus repose au contraire sur une sélection quasi exhaustive du discours parlementaire (allocutions, articles de presse, etc.) de Thorez, Blum, Tardieu et Flandin sur la période 1929 et 1939. L'ensemble représente ainsi quelque 832 discours pour un volume vertigineux de 1.570.868 mots. De par ses critères quantitatifs et qualitatifs, le corpus se caractérise par une intégration au corpus lui-même de son contexte. Contrairement au corpus traditionnel, de volume restreint, de la lexicométrie, impliquant le recours à un hors-texte au cours du geste interprétatif, un corpus réflexif comme celui construit par D. Mayaffre contient ses propres ressources interprétatives, c'est-à-dire son propre co(n)texte :

Nous entendons par réflexivité du corpus le fait que ses constituants (articles de presse, discours politiques, pièces de théâtre ; de manière plus générale, sous-parties) renvoient les uns aux autres pour former un réseau sémantique performant dans un tout (le corpus) cohérent et auto-suffisant (Mayaffre, 2002a).

Il s'agit donc d'un correctif majeur à la rupture méthodologique contestable entre le texte et son contexte qui était inhérente au traditionnel corpus lexicométrique clos et échantillonné.

1.2.2.3. Vers l'histoire linguistique des concepts

Si l'histoire des concepts est moins constituée en France qu'elle ne l'est en Angleterre et en Allemagne, une *histoire linguistique des usages conceptuels*, pour laquelle œuvre notamment J. Guilhaumou (2001) est aujourd'hui un chantier en construction. Elle tire ses racines dans les travaux du laboratoire de lexicologie politique de l'ENS Saint-Cloud et de l'analyse du discours du côté de l'histoire. Selon J. Guilhaumou :

De notre point de vue, il s'agit d'un domaine de recherche qui s'intéresse au contexte historique de signification des concepts majeurs du vocabulaire socio-politique du Moyen-Âge à nos jours, sans pour autant le réduire à de simples

considérations lexicologiques dans la mesure où ces concepts sont appréhendés au sein de relations sémantiques multiples et plus largement d'une interaction sociale à caractère langagier. L'histoire des concepts prend donc en compte les significations et les usages d'un langage spécifique dans une situation spécifique à l'intérieur desquelles les concepts sont développés par des auteurs, des acteurs et des orateurs spécifiques (Guilhaumou, 2006a : 46-47).

Pour présenter ce champ en construction qu'est l'histoire linguistique des usages conceptuels, nous mentionnerons tout d'abord le projet de constitution d'un *Dictionnaire des usages socio-politiques du français (1770-1815)*, initié par l'équipe « 18^e et Révolution française ». Sous la direction de J. Guilhaumou, de R. Monnier et de M.-F. Piguet, historiens, linguistes et lexicologues présentent dans ce dictionnaire les résultats de leur analyse sur l'usage de formes discursives sous la Révolution Française. Un projet unifie l'hétérogénéité des thématiques des huit volumes à ce jour parus de cette entreprise dictionnaire : éclairer la compréhension historique des stratégies politiques par l'étude des formes discursives analysées à la lumière de leurs conditions langagières d'émergence et de diffusion :

Dans les termes de l'analyse de discours, une fois inséré par la lecture d'archives, dans une configuration discursive, l'énoncé ainsi attesté est appréhendé dans sa mise en acte, par la concrétisation d'usages de notions, d'arguments dans un continuum entre le contexte, les ressources et les thèmes du discours socio-politique (Guilhaumou, 2000 : 114).

De par le grand soin qu'elle accorde à resituer la dimension agissante et performative des formes discursives sur leur contexte, l'orientation de l'histoire linguistique des usages conceptuels n'est pas sans évoquer l'histoire pragmatique du discours de Q. Skinner.

Le dictionnaire des usages socio-politiques (1770-1815)

Huit volumes composent à ce jour le *Dictionnaire des usages socio-politiques (1770-1815)*. Les volumes I et IV sont ainsi consacrés à la classe des désignants socio-politiques caractéristiques du discours révolutionnaire, tels qu'*aristocrates, sans-culottes, citoyen/citoyenne, endormeur*, etc. Le fascicule II traite pour sa part des *notions-concepts* : cette catégorie regroupe les termes

présidant à la fondation républicaine devenus normes ou valeurs philosophico-politiques comme *république*, *liberté*, *tyrannie*. L'originalité de l'approche de ces formes réside dans leur appréhension au sein d'un trajet discursif de sorte à révéler leur dimension langagière trop souvent négligée. L'analyse de la « terreur à l'ordre du jour » établie par J. Guilhaumou (1987) à partir d'un dépouillement systématique de la presse révolutionnaire, des Archives Parlementaires et de la correspondance des représentants en mission, est à ce titre exemplaire. De sa formation à son institutionnalisation à partir du 5 septembre 1793 et à sa propagation, l'expression « terreur à l'ordre du jour » fait l'objet d'une étude minutieuse basée sur l'examen de son contexte linguistique : si en diachronie, les énoncés injonctifs cèdent la place aux énoncés définitoires, J. Guilhaumou met en lumière l'intense charge performative de l'expression, par laquelle celle-ci fait l'histoire. Le troisième volume du dictionnaire s'attache pour sa part à présenter quelques-unes des entreprises encyclopédiques et dictionnaires de l'époque révolutionnaire, présentant la particularité d'offrir un niveau réflexif par rapport à l'usage spontané de la langue. Dans ce cadre, S. Branca (1988) se propose d'observer les stratégies des dictionnaires et d'examiner dans quelle mesure ceux-ci témoignent d'un infléchissement des modèles sémantiques durant la Révolution. Elle confronte à cet effet le traitement des formes *citoyen*, *aristocrate* et *insurrection* dans trois dictionnaires classiques et trois *dictionnaires socio-politiques* (Laplace & Tournier, 1970). Alors que le fascicule V est consacré à « Langue, occitan, usages », le sixième volume aborde pour sa part les *notions pratiques*, définies comme des formes « dont la diversification dans l'événement même leur donne une dimension pratique définissable à partir des seules ressources textuelles attestées » (Guilhaumou, 2006a : 76). Confrontées aux notions pratiques, les contributions de ce fascicule se situent au plus proche de l'histoire sémantique, en cela qu'elles s'attachent à décrire la façon « dont les notions s'organisent en faisant réflexivement sens dans leur contexte propre, qui structure en outre les

expérimentations et les attentes des acteurs concernés par l'usage de ces notions pratiques » (Guilhaumou, 2006a : 77). Dans la continuité de ce volume, le fascicule VII problématise pour sa part la distinction entre notions pratiques et notions théoriques à partir de leur articulation aux pratiques sociales qui les justifient. Focalisé sur l'analyse des concepts de l'économie politique émergente comme *production*, *socialisme*, *travail*, ce volume est unifié par le postulat selon lequel les notions théoriques sont des concepts dont le sens varie en fonction des discours qui les énoncent et des contextes sociaux et politiques dans lesquels ils s'inscrivent. Enfin, le dernier fascicule, paru en 2006, est consacré aux notions pratiques, et plus particulièrement à celles, fondatrices, de *patrie* et *patriotisme*, qui sont cernées avec rigueur dans un mouvement de mise en rapport avec leurs conditions langagières de formation.

Vers un *Dictionnaire des mots socio-politiques de la période contemporaine*

Le deuxième projet à travers lequel s'esquisse le développement d'une histoire linguistique des concepts est un chantier visant à l'établissement d'un *Dictionnaire des mots socio-politiques* pour la période contemporaine. Selon J. Guilhaumou (2006a : 72), l'ouvrage collectif *In/égalité/s* dirigé par P. Fiala (1999) représente un « prototype thématique » de ce chantier à caractère programmatique. Les contributions de cet ouvrage retracent l'évolution d'un terme central dans l'espace public, depuis le siècle des Lumières jusqu'à la période contemporaine, à savoir celui d'*égalité* et des formes morphologiques pendantes telles qu'*inégalités* et *égaux* :

Il s'agit d'analyser dans la durée, dans la diversité des usages, des supports et des genres, l'évolution morphologique et les transformations sémantiques d'un ensemble de termes formant le réseau lexical des termes de l'égalité et de l'inégalité, pour repérer sous quelles formes linguistiques et à travers quels conflits, les changements lexicaux sont intervenus et quelles traces les termes gardent aujourd'hui de leurs usages antérieurs (Fiala, 1999 : 11).

Les onze contributions réunies dans ce cadre relèvent d'une analyse lexicale et discursive de corpus textuels très diversifiés, tant au niveau de la diachronie

traitée que du registre dont ils relèvent. L'ouvrage se distingue donc des pratiques lexicologiques classiques en visant une description des usages dans des discours dont la diversification est à relier à la nature elle-même composite du discours socio-politique. P.E. Muller (1999) procède ainsi à l'examen des emplois d'*égalité* dans le discours socialiste autour de 1900, à partir d'un corpus d'articles de presse rédigés par Jaurès, publiés dans le volume *Etudes socialistes* (1901), d'un ensemble de textes de J. Guesde et d'une brochure sur le collectivisme, des ouvrages de G. Clémenceau *L'iniquité* et *Vers la réparation*, ainsi qu'un sous-corpus d'affiches électorales (1881-1910) représentatif de l'ensemble des origines politiques. Pour sa part, S. Bonnafous (1999) examine le lexique de l'égalité en rapport avec la question immigrée à partir d'un corpus composé d'articles de presse politique publiés entre 1974 et 1984, d'une revue éditée par la Caisse Nationale des Allocations Familiales depuis 1947, et un sous-corpus d'articles du *Monde* de 1995 à 1996.

1.2.3. L'analyse textuelle du discours

Après avoir évoqué le déplacement opéré du côté des historiens du discours dans les années 1980, nous souhaitons évoquer un autre renouvellement de l'analyse du discours, initié par J.-M. Adam, posant les bases d'une analyse textuelle du discours (2005). Tout en prorogeant la tradition d'une analyse du discours, notamment dans sa vocation interdisciplinaire, l'analyse textuelle du discours se propose de dépasser ses déficits de définition, d'établissement et d'interprétation du texte pointés par F. Rastier (2001 : 102), G.-E. Sarfati (2003 : 432) et F. Cossuta (2004 : 197). C'est ce tournant décisif dans la praxis de l'analyse du discours dont nous traiterons ici.

Un des apports primordiaux de l'analyse textuelle du discours est de proposer une remise en cause de l'opposition problématique des notions de *texte* et de *discours* affirmée dans l'école française de l'analyse du discours (Guespin, 1971 : 3). Cette dernière, en raison « d'une disparité théorique

initiale » (Sarfati, 2003 : 432), se caractérise en effet par un déficit textuel, qui est à relier au primat accordé à la contextualité du sens. C'est bien la notion de contexte, d'ailleurs assimilée aux conditions de production de l'énoncé, qui structure en analyse du discours l'opposition texte vs discours, figée par la formulation suivante :

Discours= texte + condition de production
Texte= discours – condition de production (Adam, 1990 : 23)

Réduisant le texte à un discours décontextualisé, cette formule désarticule le textuel du discursif et justifie la séparation des champs disciplinaires de la linguistique textuelle et de l'analyse du discours.

L'analyse textuelle du discours se fonde sur l'abandon de cette formulation et institue les notions de texte et de discours comme deux objets distincts pour des raisons méthodologiques, mais complémentaires :

[...] il s'agit de dire que ces deux concepts [texte et discours] se chevauchent et se recourent avant tout en fonction de la perspective d'analyse choisie (Adam, 2006a : 23).

Traduisant une différence de statut du texte et du discours, le syntagme *d'analyse textuelle du discours*, convoquant le discours pour objet, vise, dans ce cadre, à marquer l'inscription de la « linguistique textuelle comme un sous-domaine du champ plus vaste de l'analyse des pratiques discursives » (Adam, 2005a : 19). Pour aller plus en avant dans la (re)définition de la relation entre texte et discours, il convient d'examiner le travail opéré par J.-M. Adam sur la notion de contexte.

Alors que l'opposition texte/discours reposait sur une réduction du contexte aux conditions de production de l'énoncé, J.-M. Adam (2005) développe une conception élargie du contexte, matérialisée par l'option de la graphie « co(n)texte ». Ensemble hétérogène et diversifié de données relevant, certes, de la situation extralinguistique, à laquelle nous n'avons jamais accès que comme « (re)constructions par des sujets parlants et/ou par des analystes » comme le précise d'ailleurs J.-M Adam (2006a : 23), le contexte est

également l'environnement linguistique. La première co(n)textualisation, la plus immédiate, est ainsi celle opérée par le co-texte, à entendre comme les énoncés entourant le texte à sa gauche et à sa droite, dans l'espace matériel du document. Le co-texte ne regroupe pas uniquement les éléments linguistiques analysés par l'analyse transphrastique tels que les connecteurs argumentatifs, les organisateurs textuels, mais également les textes qui constituent le voisinage matériel du texte considéré. Le contexte d'un énoncé, selon J.-M. Adam, recouvre également les connaissances encyclopédiques des sujets, leurs préconstruits culturels et les lieux argumentatifs.

Il convient de s'arrêter ici sur ce qui fait lien entre le contexte et le cotexte tel que définis dans cette perspective. S'inspirant notamment de G. Kleiber (1994 : 19), l'auteur pose que le contexte, qu'il soit linguistique ou extralinguistique, n'est pas une donnée extérieure au sujet mais qu'il renvoie à un processus mémoriel. La notion de *mémoire (inter)discursive* (2006a) est ainsi cette catégorie permettant de relier tout énoncé à d'autres textes, qu'ils correspondent à une autre partie du texte ou à un texte antérieur :

[La mémoire discursive] n'est pas tant alimentée en permanence par des événements de la situation extralinguistique que par les énoncés portant sur ces événements et constituant eux-mêmes des événements (Adam, 2005a : 27).

Partant, le contexte s'étend aux rapports qu'un énoncé entretient avec d'autres séries de textes, ce qui revient à poser l'interdiscursivité et l'intertextualité comme paliers contextuels :

[...] le contexte d'un texte donné est dans les genres présents dans l'interdiscours dont il est issu et dans l'intertextualité qu'il mobilise ; le contexte d'un texte est dans les variations historiques de ses éditions, dans les énoncés péricertextuels et les co-textes qui l'entourent matériellement ainsi que dans les gloses successives dont il a été l'objet par son auteur ou des commentaires (métatextes) (Adam, 2006a : 28).

La méthodologie d'analyse textuelle du discours relève donc d'une reconstruction de la mémoire (inter)discursive d'un texte empirique, dans un mouvement de contextualisation régi par un principe d'accessibilité, donnant la

priorité au co(n)texte linguistique disponible en mémoire par rapport aux données extralinguistiques. Dans ce cadre, le programme d'une analyse textuelle du discours devient l'étude non pas d'un texte fermé sur lui-même mais d'un ensemble de textes, reliés par leur généricité ou par leur intertextualité au texte dont la convocation est aussi nécessaire que celle du co-texte linguistique immédiat.

Ce programme à dimension herméneutique (Adam, 2006a : 32) a pour pendant la mise en œuvre d'un établissement philologique et comparé du texte afin d'éviter « toute fermeture illusoire du texte sur lui-même » (Adam, 1999 : 143). Dès lors que l'analyse textuelle du discours se focalise sur le régime de circulation du texte et qu'elle l'envisage comme donnée matérielle, elle fait l'expérience du « leurre de l'évidence naturelle du texte » (Adam, 2005b : 70). Pas plus qu'il n'est UN texte du discours de Montréal du Général de Gaulle, il n'existe UN texte du conte de la Belle au Bois dormant : seules différentes textualisations sont disponibles. S'ouvrant chacune à des re-contextualisations différentes, ces textualisations sont autant de faits de discours singuliers où se donnent à lire des variations qui sont des accès privilégiés au sens. Il revient ainsi à l'analyse textuelle du discours d'éclairer ses variations à la lumière d'une analyse contrastive et différentielle, dont dépend l'historicisation du texte comme fait de discours :

Un texte n'est un fait de discours que par sa mise en circulation avec l'*interdiscours* d'une formation socio-discursive, elle-même définie comme lieu de circulation de textes (*intertextualité* propre à la mémoire discursive d'un groupe) et de catégories génériques (*interdiscursivité* des genres et des sous-genres) (Adam, 2005b : 28).

C'est ainsi parce qu'elle réinjecte la dimension historique du texte que l'établissement philologique du texte opère un rôle décisif dans la réunion des sciences du texte et de l'analyse du discours.

1.2.4. La dynamisation de notre champ de recherche est conditionnée par la constitution de vastes ensembles de données textuelles d'intérêt historique

1.2.4.1. Peu d'expériences inaugurales prennent la PQR3R pour objet

Les travaux tournés vers le discours et ses matérialités, qu'ils s'inscrivent en histoire ou en sciences du langage, ne se sont jusqu'à présent guère appuyés sur la presse de la Troisième République, qui reste dans cette perspective, une ressource à construire et à explorer.

L'analyse du discours du côté de l'histoire s'est essentiellement concentrée sur le XVIII^e siècle (Maldidier & Robin, 1974 ; Robin, 1975 ; Guilhaumou, 1986, 1987, 1992) et la période de l'entre-deux-guerres (Marcellesi, 1976 ; Peschanski, 1989 ; Mayaffre, 2000), voire la période de mai 1968 (Demonet *et al.*, 1975 ; Maldidier & Robin, 1977). Les corpus sur lesquels se basent les analyses dans ce cadre sont majoritairement constitués de textes de congrès syndicaux (Marcellesi, 1970, 1971) et d'allocutions d'hommes politiques (Courdesses, 1971).

À de notables exceptions telles que l'exploration de la collection du *Père Duchesne*, la presse reste ainsi une source documentaire peu explorée, ou ne donne lieu qu'à des corpus au volume très restreint, procédant de la sélection d'articles traitant d'un thème en synchronie dans un nombre limité de journaux (Maldidier, 1971 ; Maldidier & Robin, 1977). D'indéniables difficultés de constitution de corpus de journaux président à cet état de fait. Les réflexions d'A. Prost (1974b: 117-124) fournissent sur ce point un témoignage de premier ordre. L'historien a pour projet d'analyser le discours produit par le « mouvement combattant » sur le fonctionnement des institutions républicaines entre les deux guerres, dont il justifie l'intérêt historique en mettant l'accent sur la perception originale du phénomène politique qu'elle permet :

[...] l'étude du mouvement combattant nous ouvre l'accès à l'opinion de ces classes moyennes qui, en d'autres lieux, ont assuré le succès des régimes autoritaires. Leur opinion sur la crise des institutions mérite donc une attention toute particulière. [...] L'opinion des milieux d'anciens combattants sur les problèmes constitutionnels – le conseil économique et social ou la dissolution automatique – s'accompagne de considérations sur la politique en général qui traduisent une certaine façon de vivre l'événement, de ressentir le régime, de se situer par rapport à la politique (Prost, 1974b : 118-119).

Pour dispositif d'observation de ce discours, A. Prost choisi de s'appuyer sur la presse du mouvement combattant. Dans ce but, il sélectionne ainsi un certain nombre de journaux locaux et nationaux de diverses tendances dans une perspective contrastive. Il est alors confronté aux difficultés d'accès à des documents dont la consultation ne pouvait être réalisée que sur place et dont la photocopie était interdite. Le microfilm constituait alors la solution de levée de la contrainte matérielle d'accès au texte, mais son coût et le temps nécessaire à sa création (« 3 à 4 mois entre la demande et la réception du travail » (Prost, 1974b : 123)) étaient incompatibles avec les conditions de la recherche de l'historien. S'insurgeant contre les obstacles qui entravent la recherche de l'historien désireux de travailler sur la presse et condamnant dans ce cadre l'action de la Bibliothèque Nationale¹³, A. Prost met en relief les incidences de ces conditions d'accès au document sur la qualité du corpus réuni et par suite, de celle de son analyse. Alors qu'il aurait voulu constituer un corpus de textes intégraux, l'historien est tout d'abord contraint de se rabattre sur « une solution de fortune » aux antipodes des exigences d'une démarche scientifique : un corpus somme toute restreint d'énoncés sélectionnés à partir de mots-pivots. De surcroît, A. Prost se voit également obligé d'abandonner son projet initial d'une analyse lexicométrique du discours du mouvement combattant pour se tourner vers l'analyse d'une sélection de termes et de leurs réseaux d'oppositions et de substitutions des formes grammaticales, constituant « dans

¹³ « Que la principale bibliothèque de recherche française ne soit pas en mesure de rendre aux chercheurs un service plus rapide et moins onéreux est un scandale dont la responsabilité incombe à une gestion bureaucratique étroite et inadaptée » (Prost, 1974b: 123).

le cas présent tout ce que l'historien a pu emprunter au linguiste » (Prost, 1974b : 124).

Au travers de ce témoignage, transparaissent les difficultés auxquelles tout projet d'étude sur la presse était confronté, et comment celles-ci ont contribué à retarder le développement de l'exploration de cette ressource.

1.2.4.2. Retour sur des entreprises pionnières de constitution de base textuelle de presse ancienne

S'il n'existe à ce jour aucune entreprise autre que la nôtre d'acquisition planifiée de données de PQR3R, certains laboratoires de recherche, profitant de l'amélioration des conditions technologiques, ont initié des projets de constitution de corpus de presse depuis les années 2000.

L'Écho de la Fabrique

En France, le projet le plus notable est celui mené depuis 2004 par l'ENS-LSH, visant à l'édition critique d'un « corpus » de journaux ouvriers lyonnais des années 1831-1835, en partenariat avec la Bibliothèque Municipale de Lyon, du Musée Gadagne d'Histoire de Lyon et des Marionnettes du Monde, responsables de la mise à disposition des fonds nécessaires à l'étude de la presse lyonnaise.

Placé sous la responsabilité de Ludovic Frobert, ce projet interdisciplinaire (histoire, économie, analyse du discours du côté de l'histoire, etc.) se matérialise par l'édition commentée de chaque numéro hebdomadaire d'un journal dont l'intérêt historique est des plus grands : *L'Écho de la Fabrique*. Premier journal des canuts lyonnais, publié du 30 octobre 1831 au 4 mai 1834, ce journal porte les revendications des ouvriers de la soie et d'autres corps ouvriers lyonnais dès la veille de la première insurrection lyonnaise en octobre 1831. La feuille défend notamment l'idée d'un tarif minimum (la Mercuriale) pour les tisserands en soie exploités par les marchands et réclame la constitution d'une jurisprudence. Après l'insurrection lyonnaise d'avril 1834, *L'Écho de la*

Fabrique se voit censuré par un procès condamnant le contenu politique de sept de ses numéros.

Outre ce premier titre, le projet de *L'Écho de la Fabrique* édite également un ensemble de feuilles de la petite presse lyonnaise ouvrière des années 1831-1835 (à laquelle vont mettre fin les lois de septembre 1835) comme *L'Écho des Travailleurs*, mais aussi *La Tribune Prolétaire* (1834-1835), *L'Union des Travailleurs* (1835), *Le Nouvel Écho de la Fabrique* (1835), *L'Indicateur* (1834-1835), etc.

Depuis l'été 2006, plusieurs chercheurs (France, Canada, États-Unis) mènent des recherches *sur* ou *avec* les ressources numérisées de *L'Écho de la Fabrique*. Un recueil d'articles critiques est en cours, rassemblant de nombreux travaux tels que ceux de l'économiste L. Frobert, de l'historien J. Popkin sur le journaliste prolétaire Marius Chastaing, de J. Beecher et du sociologue P. Mercklé sur le fouriérisme des canuts, de l'analyste du discours M. Tournier sur les labels ouvriers dans le journal des canuts, de S. Mombert, examinant pour sa part les rubriques littéraires de *L'Écho de la fabrique*, et enfin de J. Guilhaumou. Ce dernier livre dans un récent article (Guilhaumou, 2007) une analyse lexico-sémantique d'un corpus réflexif issu de la base de *L'Écho de la Fabrique*, délimité aux interventions du premier rédacteur du quotidien, A. Vidal, durant la période 1832-1833. Exploitant les outils d'analyse assistée par informatique intégrés à la base (contexte, concordances, co-occurrence), l'historien du discours met au jour le trajet discursif de *peuple* à *prolétaire*, posant les conditions langagières de l'existence de la classe ouvrière, et au cours duquel se dessine l'émergence d'une consistance positive du second terme.

L'entreprise *Lancaster Newsbooks Corpus*

Sur le plan international, une autre entreprise de constitution d'une base textuelle de presse a retenu toute notre attention : celle du *Lancaster Newsbooks Corpus*¹⁴ initiée par le département de linguistique de l'Université de Lancaster en 2001.

Placée sous la responsabilité de Tony McEnery et Andrew Hardie, dont les recherches s'inscrivent en linguistique de corpus, l'entreprise vise à constituer une base de données en mode plein-texte à partir d'un ensemble de 7000 gazettes anglaises (*newsbooks*) de la seconde moitié du XVII^e siècle, correspondant au fonds numérisé dit « Thomason Tracts »¹⁵ de la *British Academy*. Ces gazettes, formes de proto-journalisme à la parution irrégulière, rapportent les nouvelles du Parlement anglais, décrivent les campagnes militaires et se font l'écho de tous les faits et gestes des personnalités politiques. Outre une source précieuse pour l'histoire du journalisme, ces gazettes constituent, de par leur contexte de parution, une archive dont l'intérêt historique et linguistique est précieux. La deuxième moitié du XVII^e siècle est en effet une période d'intense agitation politique, - de la Guerre Civile anglaise à la restauration de la monarchie sous Charles II, en passant par le Commonwealth et le protectorat d'Olivier Cromwell -, sur lequel ces gazettes livrent un témoignage précieux.

À ce jour, deux études en linguistique de corpus particulièrement avancées sont menées à partir des ressources de la base. La première, menée par P. Clough, R. Gaizauskas et S. Piao à partir d'un corpus de 800 000 mots des gazettes publiées de décembre 1653 à la fin mai 1654, se propose d'observer les phénomènes d'intertextualité, de reprise textuelle ainsi que de variance (*text re-use*) caractéristiques du proto-journalisme. Le traitement de ce sujet d'étude revêt d'autant plus d'intérêt qu'à la fin du XVII^e siècle en Angleterre, les rédacteurs travaillaient simultanément pour plusieurs gazettes,

¹⁴ <http://www.ling.lancs.ac.uk/newsbooks>

¹⁵ <http://www.bl.uk/collections/thomason.html>

tandis que les gazettes étaient plagiées ou plus ou moins réécrites pour être imprimées dans une autre gazette, voire contrefaites. Pour mettre en lumière les phénomènes de reprise, allant de la citation au plagiat, P. Clough, R. Gaizauskas et S. Piao ont ainsi développé des algorithmes et un logiciel visant à mesurer la similarité lexicale des textes du corpus considéré.

Un deuxième axe de recherche du département de linguistique de Lancaster alimenté par la mise à disposition des ressources numérisées de presse ancienne relève quant à lui d'une étude de l'argot et des métaphores dans une ère de moralité puritaine. L'étude se fonde sur un corpus de quelque 200 000 mots, correspondant à la parution du *Mercurius Fumigosus* de 1654 à 1655, gazette alliant rapports de débauche, humour cru et poèmes fantastiques voire fantasques. À travers l'utilisation de fonctionnalités documentaires (concordances, requête lexicale unique ou multiple), est ainsi menée une recherche sur les métaphores et les attestations des termes d'argot à l'apogée de la période puritaine.

Des entreprises telles que *L'Écho de la Fabrique* ou du *Lancaster Newsbooks Corpus* illustrent les grandes directions du renouvellement des savoirs historiographiques permises par la constitution de grandes bases textuelles de presse ancienne. De telles entreprises exigent cependant la mise en œuvre de lourdes opérations d'acquisition des données, que nous souhaitons à présent examiner.

1.3. Construire une base textuelle de presse ancienne : aspects méthodologico-techniques

Jusqu'au début du XXI^e siècle, les fonds de PQR3R n'étaient disponibles que sur support papier et/ou microfilms, empêchant la mise en œuvre, comme nous l'avons évoqué précédemment de toute étude de grande ampleur.

Bien conscients de l'intérêt historique de cette presse, les EA 2181 LASELDI et 3187 ATST, réunissant leurs efforts dans le pôle *Archive, Bases, Corpus* de la

MSHE Claude-Nicolas Ledoux, ont initié, dès 2003, soit deux ans avant le plan de numérisation de la presse du XIX^e siècle mené par la Bibliothèque Nationale de France, un programme de constitution d'une base plein-texte répondant aux exigences de la communauté scientifique.

Après avoir présenté les objectifs de ce projet pionnier en la matière, visant à l'obtention de données textuelles établies et normalisées, nous ferons état des difficultés induites par une chaîne de traitement de la presse ancienne, qui sont autant de défis techniques et méthodologiques à relever pour les sciences des textes.

1.3.1. Constitution de données textuelles

1.3.1.1. Renforcer le statut scientifique de la presse ancienne comme ressource

Signaler qu'après l'écriture et l'imprimerie, le numérique figure une troisième révolution technique modifiant profondément notre rapport au texte, est aujourd'hui un lieu commun. Tout un chacun, dans son quotidien, expérimente l'évidente rupture instaurée par rapport au manuscrit et à l'imprimé : celle de la levée, jusqu'alors inédite, de la contrainte physique pesant sur le texte comme le rappelle R. Chartier et G. Cavallo :

L'électronique, qui permet la communication des textes à distance, annule la distinction, jusqu'ici ineffaçable, entre le lieu du texte et le lieu du lecteur. [...] Détaché de ses matérialités et de ses localisations anciennes, le texte en sa représentation électronique peut théoriquement atteindre n'importe quel lecteur en n'importe quel lieu (Chartier & Cavallo, 2001 : 39).

Les apports de la numérisation de la presse ancienne ne sauraient néanmoins se réduire à la simple facilitation des conditions d'accès à des documents auparavant confinés dans leurs lieux de conservation. Conséquence de la brutalité du changement introduit par une évolution technique majeure, la numérisation est trop fréquemment réduite à une « dématérialisation » du texte, concept erroné sur le plan épistémologique qui reflète et alimente l'illusion d'un

texte libéré par le numérique de toute matérialité. Cristallisant le manque de théorisation, ou du moins l'adoption d'une position critique tardive vis-à-vis du numérique, le concept de dématérialisation, s'imposant systématiquement dès lors qu'il est question d'archivage numérique, masque ce que figure la conversion de l'analogique au digital : un changement de régime matériel (Jeanneret, Y., 2000, 2003) à travers lequel se réaffirme la matérialité du texte et son lien au document.

Le numérique met - au contraire - à mal la représentation d'un texte existant en lui-même, hors de toute matérialité et nous rappelle l'importance de ses conditions de production, de distribution et de réception. Ce faisant, l'ère du numérique nous invite à repenser le texte et ses frontières. Numériser la presse ancienne nous confronte directement à l'inscription d'un texte sur un document dont la structuration fait sens et met en exergue les liens du texte à son contexte, à son péri-texte et au paratexte. Sur le plan théorique, le numérique nous convie donc à reconsidérer le statut de ces éléments qui font déjà partie du texte ; sur le plan méthodologique, il interroge les moyens par lesquels préserver l'intégralité des propriétés du texte tel qu'inscrit sur son *document* d'origine. Conserver les propriétés significatives et constituantes du texte, de la saillance typographique à l'emplacement topographique en passant par la coquille, tel est le dessein d'un établissement scientifique des données textuelles. Alors que le mode image offre la garantie d'une préservation des propriétés du texte, le mode plein-texte permet quant à lui de dépasser le seuil du texte brut, état du texte auquel ce dernier ne peut être réduit, pour ouvrir notre approche du texte à la pluralité de ses dimensions. L'établissement scientifique du texte, en tant qu'ensemble des opérations de restitution du texte visant à offrir un socle à son analyse critique, a dans ce cadre pour tâche de proposer un accès aux différents états du texte qui le composent. L'annotation linguistique portant sur les niveaux grammatical et morphosyntaxique constitue ainsi un chantier primordial de l'établissement scientifique des données. Pour répondre aux exigences de l'analyse textuelle des discours, ces annotations

doivent viser à une fiabilité optimale et ne peuvent donc être opérées de façon purement automatique en raison du taux d'erreurs qui est inhérent à tout automate. Il s'agit donc de mettre en œuvre une annotation qui soit en premier lieu, *contrôlée* afin de garantir une qualité optimale, mais également *historicisée* (Viprey, 2005b ; Viprey & Lethier, 2008a). L'annotation, qu'on envisage ici la lemmatisation ou la mise en place de codes grammaticaux, implique en effet inévitablement un acte d'interprétation dont il convient de signaler l'origine. On touche ici au principe fondamental de tout établissement du texte à prétention scientifique : *mise en ordre valorisante* (Viprey, 2005b : 56), l'établissement se doit de rendre compte de son action sur le texte pour qu'elle puisse pleinement constituer une valeur ajoutée.

Enfin, la constitution de données établies en mode plein-texte prend tout son sens en raison de la modification majeure du geste de lecture qui se joue du support papier au support numérique. En premier lieu, ce changement de régime de matérialité a pour pendant la mise en œuvre d'une lecture assistée par informatique synonyme d'une scientificité accrue. En effet, les outils d'analyse statistique textuelle (Lebart & Salem, 1994) permettent à l'analyse de prétendre à l'exhaustivité et à la systématisme de l'indexation et des relevés, articulés à des vues synthétiques au fort potentiel heuristique d'ensembles de données qu'une lecture traditionnelle ne saurait appréhender. En second lieu, ces mêmes outils substituent à une lecture exclusivement linéaire et syntagmatique du texte, une analyse paradigmatique, *tabulaire* et *réticulaire* pour reprendre les mots de J.-M. Viprey (2005b : 59). Cette ouverture aux dimensions textuelles autres que séquentielles initie donc une approche privilégiée de la textualité.

La numérisation de la presse ancienne ne se justifie dès lors pas principalement sur la base d'un renouvellement de ses conditions d'accès mais sur celle du renouvellement du statut scientifique de cette source et de ses lectures. La constitution de données en mode plein-texte, établies selon un

rigoureux protocole philologique, est cette étape décisive et nécessaire dont dépend la mise en œuvre d'explorations approfondies de la PQR3R.

1.3.1.2. Assurer portabilité et pérennité aux données textuelles

À l'exigence de la constitution de données finement établies à laquelle doit répondre une archive de base ancienne, s'ajoute l'impératif de garantir la portabilité et la pérennité des données en procédant à leur normalisation. Le contexte scientifique actuel pointe en effet la nécessité de (re)fonder la politique de création et de gestion des données textuelles afin de mettre fin à leur déshérence (Rigot, 2006).

Envisagées comme un dispositif d'observation adapté à la singularité d'un projet de recherche, les données textuelles se révèlent trop souvent à « usage unique » et déperissent, au terme de l'analyse pour laquelle elles ont été construites, dans les archives d'un laboratoire. Cette absence de gestion des données est paradoxale et regrettable à plusieurs titres. En premier lieu, elle entre en contradiction avec le soin méticuleux apporté à la constitution des données, qui est pourtant une phase de la recherche centrale, tant sur le plan méthodologique qu'à l'égard du temps qui lui est consacré. Tout chercheur rompu à la constitution de corpus confirmera bien volontiers que l'acquisition des données textuelles représente un investissement souvent plus important que celui de l'analyse à proprement parler. Enfin, le manque de réflexion vis-à-vis de l'archivage des données et de leur partage contredit la visée scientifique des sciences des textes : la dissémination des données, clivées d'un point de vue géographique et disciplinaire, est un obstacle à la mise en œuvre de projets de recherche de grande ampleur et à la comparaison des résultats par la confrontation des méthodes et des points de vue.

Parce qu'elle conditionne les avancées de l'analyse des données textuelles, tant sur le plan méthodologique que sur le plan des réalisations, la mise en commun des données textuelles est aujourd'hui un axe central de recherches des laboratoires d'analyse du discours. Parmi les axes du chantier

Textopol, mené au Céditec (EA 3119, Université de Paris XII) à l'initiative de P. Fiala et J.-M. Leblanc, visant notamment au développement de protocoles lexicométriques, la base *Textopol*¹⁶ met ainsi en ligne un ensemble important de textes socio-politiques contemporains réunis en corpus, constitués par les enseignants-chercheurs et doctorants de ce laboratoire de 2000 à 2006, ainsi que les données textuelles issues des travaux du Laboratoire de lexicologie politique et de lexicométrie de l'ENS/St Cloud depuis 1967. Autre archive exemplaire participant au développement des outils et de la recherche empirique en logométrie, la base de données de discours politiques *PoliText*¹⁷ (UMR 6039 *Bases, Corpus et Langage*), initiée par D. Mayaffre, met à disposition des discours de locuteurs politiques sur la période 1789-2002.

La mise en ligne des données textuelles, si elle permet leur décloisonnement, ne suffit pas à elle seule à permettre leur réutilisation. Celle-ci est tout d'abord conditionnée par la description fine du corpus constitué. Or, le chercheur omet trop souvent, dans l'immédiateté de la constitution de son corpus, d'explicitier les conditions de constitution des données, ainsi que les choix méthodologiques et théoriques sous-jacents à leur prétraitement (indexation, annotation, etc.). Le partage des données repose ainsi nécessairement sur leur description fine mais implique également, et de façon plus problématique, que soient garanties leur lisibilité et leur portabilité. Or, jusqu'à présent, les données textuelles primaires sont massivement stockées et éditées dans des formats propriétaires, qui impliquent de disposer de la même suite de logiciels de traitement de texte pour consulter les données tout en s'assurant de leur intégrité. Sans nous attarder, ici, sur les problèmes éthiques posés par les formats propriétaires, nous mettrons en relief le risque quant à la pérennité des données qu'ils représentent. D'un changement de code à l'autre, les données de l'utilisateur peuvent être altérées ou inutilisables.

¹⁶ Accessible à l'adresse : <http://textopol.free.fr>

¹⁷ Accessible à l'adresse : <http://www.unice.fr/ILF-CNRS/politext/>

La nécessité d'un format d'échange commun s'impose également dans le cadre de l'interopérabilité entre logiciels d'analyse de données textuelles. Afin de permettre une chaîne de traitement collaborative et de pouvoir soumettre un même état de corpus aux fonctionnalités respectives des logiciels d'analyse de données textuelles, il convient en effet de disposer d'un format d'échange commun permettant de conserver les différents niveaux d'annotation du corpus. C'est tout le sens des travaux actuellement menés dans le cadre du *Réseau pour l'échange de ressources et de méthodologies en analyse de texte assistée par ordinateur* ATONET, regroupant les concepteurs des logiciels *Alceste* (M. Reinert), *Astartex* (J.-M. Viprey), *Dtm* (L. Lebart), *Lexico3* (A. Salem *et al.*), *Sato* (F. Daoust) et *Webllex* (S. Heiden). Le format XML-TEI s'impose dans ce cadre comme format-pivot (Daoust & Marcoux, 2006).

1.3.1.3. Le XML-TEI

À l'heure de l'affranchissement des formats propriétaires, XML (*eXtensible Markup Language*), promu par le *World Wide Web Consortium* (W3C), s'impose comme le format ouvert de référence. Associé aux recommandations du consortium TEI (*Text Encoding Initiative*), ce format offre le socle technique et méthodologique d'un protocole philologique rigoureux, garantissant la pérennité et la portabilité des données.

Langage général issu d'une simplification du SGML, XML est un format d'édition de textes numériques ayant pour caractéristique de séparer le contenu logique du document de sa représentation visuelle. Selon ses créateurs, son principal avantage tient à son extensibilité, autorisant l'utilisateur à définir les balises des éléments, contrairement à HTML dont le stock de balises est globalement statique. En XML, l'utilisateur peut inventer une quasi-infinité de noms de balises pour satisfaire à des besoins particuliers. Le XML constitue de surcroît un balisage dont la syntaxe est relativement simple et qui peut être rapidement maîtrisée par des néophytes. Les principes syntaxiques minimaux sont restreints : les caractères réservés < et > délimitent les balises qui

structurent le document ; chaque balise peut comporter des attributs qui définissent son interprétation ; la structure d'un document est hiérarchique : chaque document doit être contenu dans un élément racine qui regroupe l'ensemble du document et tout élément doit s'inscrire dans un élément de niveau supérieur.

L'extensibilité du *XML* induit cependant que l'ensemble du codage soit décrit dans un document externe de type *DTD* (*Document Type Definition*) ou de type *XML Schéma*. Si ces deux types de documents sont des descriptions de format de document *XML* permettant d'en définir la structure, *DTD* est inspiré du *SGML* et *XML Schéma* constitue un document *XML*, d'où certaines différences importantes pour l'utilisateur. C'est ce que remarquent G. Williams et C. Ropers, pour qui l'utilisation du *XML Schéma* est jugée préférable à une *DTD* :

[...] le schéma représente le mieux la sérendipité dont ont su faire preuve les concepteurs d'XML. En effet, disposer d'une DTD offre une certaine forme d'assurance assez proche de celle offerte par un processus de compilation. Elle permet de s'assurer que le document *XML* créé est non seulement *bien formé* (qu'il respecte les règles générales d'XML) mais aussi *valide* (qu'il ne présente aucune erreur de logique interne). En revanche, la DTD ne donne pas toujours une information très précise quant à l'endroit où se trouve l'erreur, et pas la moindre information quant à la manière de corriger l'erreur en question. Tout ceci revient à dire que l'utilisateur qui dispose d'une DTD doit aussi avoir une grande expérience de la conception de documents *XML* s'il veut avoir une chance d'éradiquer d'éventuelles erreurs. Le schéma, quant à lui, applique le proverbe « il vaut mieux prévenir que guérir » en se servant de la documentation explicative pour faire des propositions à l'utilisateur au fur et à mesure de la création du document. Si l'utilisateur s'en tient aux propositions qui lui sont faites, il ne court pas de risque de faire une erreur en chemin et n'aura donc rien à corriger. Par conséquent, disposer d'un schéma permet à l'utilisateur de se concentrer sur l'utilisation des données plutôt que sur l'apprentissage avancé d'XML (Williams & Ropers, 2007 : 31).

Cette remarque met en relief un obstacle essentiel dans l'exploitation des possibilités du *XML* : la construction de *DTD* et/ou de *XML Schéma* sont des tâches complexes et longues, exigeant des compétences informatiques approfondies, comme en témoignent les propos de G. Williams et C. Ropers :

[...] on devra au moins s'accorder sur le fait que les deux comportent le même désavantage : même si *XML* est plus simple que *SGML* en ceci qu'il permet de ne pas spécifier certains paramètres pour lesquels les pré-processeurs

disposent d'information par défaut, les documents explicatifs dont on a besoin restent complexes et longs à construire (Williams & Ropers, 2007 : 31).

Répondre aux besoins de la communauté des sciences humaines en matière d'encodage des corpus et lui proposer un modèle général d'expression de la structure d'un texte, quels que soient sa langue et son genre : tel est le dessein de l'initiative TEI¹⁸ (*Text Encoding Initiative*) réunie en Consortium dès 1987. Celui-ci s'efforce de fournir un mode d'encodage du texte selon un ensemble de principes, dits de « Poughkeepsie »¹⁹, selon lesquels les recommandations d'encodage devraient :

- être suffisamment précises pour représenter les propriétés des textes intéressants pour les chercheurs ;
- être simples, claires et concrètes ;
- être utilisables facilement par les chercheurs et ne pas nécessiter l'utilisation de logiciels spécifiques ;
- permettre une définition rigoureuse des textes en vue de traitements efficaces ;
- être modifiables par l'utilisateur ;
- respecter les normes en vigueur ou sur le point d'être adoptées²⁰.

Les directives (*guidelines*) de la TEI s'imposent aujourd'hui comme un standard international²¹ permettant l'échange et le partage des données textuelles s'appuyant sur un format normalisé et indépendant des logiciels et des plateformes. Bien que les principes édictés plus haut restent les mêmes, les recommandations de la TEI évoluent au rythme des outils informatiques disponibles et des besoins de la communauté scientifique : de 1987 à 2005, ce

¹⁸ La TEI est soutenue par l'*Association for Computers and the Humanities*, l'*Association for Computational Linguistics* et l'*Association for Literary and Linguistic Computing*. Le projet a été en partie financé par le National Endowment for the Humanities américain, la DG XIII de la CEE, la fondation Andrew W. Mellon et le Social Science and Humanities Research Council du Canada. Les *Recommendations* ont été publiées en mai 1994, après six ans de travaux auxquels ont participé un groupe de chercheurs de toute nationalité et de toute discipline, sous la direction de Lou Burnard Université d'Oxford, et Christopher M. Sperbreg-McQueen, Université d'Illinois à Chicago.

¹⁹ En référence à l'endroit où s'est tenue la conférence.

²⁰ http://www.tei-c.org/Guidelines/Customization/Lite/teiu5_fr.html#toc_SEC1

²¹ On dénombre une utilisation de la TEI dans tous les pays du monde tandis que ce standard ne concerne pas uniquement les chercheurs en sciences humaines, mais toutes les industries du texte. À titre d'exemple, signalons que la TEI est ainsi le standard qui a été utilisé par le producteur européen de CD-Roms Chadwick-Healey, pour sa base de données de *Poésie anglaise*.

ne sont pas moins de cinq versions de la TEI qui ont ainsi été livrées à la communauté.

Insistons sur l'ambition du projet et la complexité de sa réalisation, devant modaliser l'extrême diversité des objets des sciences humaines à encoder (corpus oraux, dictionnaires, prose scientifique, édition critique, etc.) tout en respectant leurs caractéristiques. Il faut souligner fortement que la TEI vise plus à offrir des propositions d'encodage de données pour en faciliter le partage qu'à instaurer une norme. Pour s'adapter à la singularité des documents traités tout en modélisant leur structure, la TEI prévoit ainsi 450 éléments, classés en 22 modules, dont un ensemble de balises obligatoires (*core tag sets*), différents modules de bases (*base tag sets*) et certaines balises additionnelles.

Si la TEI se révèle un outil puissant, l'appropriation de ses nombreuses et complexes recommandations reste extrêmement fastidieuse. Pour ne pas constituer paradoxalement un obstacle à la constitution de données normalisées dont il cherche à faciliter le partage, le consortium TEI a mis en place en 2005 une version P5 (*proposal 5*) articulée autour du langage XML. Dans cette version, l'encodeur, après avoir sélectionné les balises proposées, crée la grammaire de son document et valide ainsi ses données XML par rapport à cette grammaire dans la syntaxe de son choix (DTD, *Relax NG*, XML *Schéma*). Toujours dans la perspective de faciliter l'encodage d'un corpus en XML - TEI, la version 5 introduit de surcroît un nouveau fichier XML utilisant une grammaire dérivée de la TEI, nommé ODD (*One Document Does it all*) ainsi qu'un outil en ligne convivial pour générer une DTD (*Roma*). Sous le fichier ODD, l'encodeur peut dès lors choisir les modules à utiliser et construire ainsi sa grammaire XML, tout en ayant la possibilité de procéder à certaines variations par rapport à la TEI « canonique » : suppression des éléments non pertinents, ajout d'attributs, ajout d'un nouvel élément, etc. Contrairement aux possibilités offertes dans les versions précédentes de la TEI, ODD conserve la mémoire de chacun des changements opérés, documentés au sein même du document. Parce qu'ODD

est pleinement un format pivot, l'interopérabilité des données textuelles est donc mieux assurée.

Il nous semble délicat d'aller plus avant dans l'état de l'art des techniques d'encodage des textes en raison du caractère extrêmement évolutif de celui-ci. Il nous semble important de signaler qu'au début de nos recherches doctorales, qui correspond précisément à l'introduction de la version 5 de la TEI, la littérature de présentation destinée aux chercheurs de sciences humaines non rompus aux technologies informatiques était quasi-inexistante. À l'exception notable de S. Loiseau (2002) établissant, sous forme de rapport, les différents protocoles de construction de données normalisées et établies, il est intéressant de relever que les problématiques de la normalisation et de l'établissement des données, au cœur du programme d'une nouvelle *philologie numérique* notamment appelée par F. Rastier, et ce, dès 2001, ne sont réellement devenues centrales dans la communauté de l'Analyse Textuelle des Données qu'à partir de 2005. En témoigne l'accroissement soudain de la littérature et des événements scientifiques²² qui leur sont consacrés. Dès 2005, en effet, un nombre de plus en plus important d'articles, de sections d'ouvrages (Habert, 2005 : 123-135), voire de numéros de revue (Malrieu, 2007), amorcent une réflexion sur les enjeux d'une représentation fine et normalisée du texte et présentent des bilans d'expériences d'exploitation de tels corpus.

1.3.2. La numérisation de la presse ancienne

Le terme de numérisation est aujourd'hui employé dans de très nombreux contextes et de façon quelquefois inadaptée : parmi d'autres abus de langage, pensons par exemple à la fameuse « numérisation » des thèses, alors que cette opération ne consiste qu'en une simple mise en ligne de documents préalablement archivés dont la forme première est numérique. Or, dans un sens strict, la numérisation est un processus de conversion d'une information

²² Parmi lesquels la tenue de la journée d'Etudes « Corpus et questionnements du littéraire » au Modyco, organisée par D. Malrieu les 15 et 16 novembre 2005

analogique en une information digitale exploitable par l'informatique. De façon générale, la numérisation est un processus technique déterminé en premier lieu par les usages attendus en fin de processus. Selon qu'elle vise à l'obtention d'une base image ou d'une base texte, la numérisation consiste en une chaîne d'opérations plus ou moins complexe. Le matériel à utiliser, les choix techniques varieront en fonction du type de données attendu en fin de processus, mais aussi en fonction de la nature et du volume des documents à traiter.

1.3.2.1. La numérisation en mode image

Parce qu'elle est celle qui nécessite la plus simple chaîne d'opérations, nous débuterons notre exposé des techniques de numérisation par la présentation d'une entreprise visant à convertir un document primaire en mode image. Dans le cas de la presse ancienne, la numérisation peut s'effectuer à partir de la collection papier du quotidien ou d'un microfilm.

Dans le cas d'une numérisation à partir d'un support papier, la nature et les caractéristiques matérielles de la presse ancienne conduiront à renoncer à l'utilisation de scanners bureautiques, également dits « scanners à plat ». Ceux-ci, à cause de leur principe de conception, nécessitent en effet que la surface à numériser soit dirigée vers le bas et posée contre la vitre : d'où une fragilisation des supports-papier qui est bien sûr incompatible avec les normes de conservation pesant sur les documents anciens. De surcroît, la taille des journaux excède le format toléré par ce type d'appareils, qui correspond à un A3. L'utilisation de scanners de livres, récemment apparus sur le marché des scanners spécialisés, s'impose dès lors. Répondant pour leur part aux exigences de la numérisation de documents reliés de grand format tels que la presse ancienne, ces scanners à « livre ouvert » ont un principe de fonctionnement non-destructif : le document à numériser doit être ouvert dans son sens de lecture, le texte dirigé vers le haut. Deux classes d'appareils sont actuellement disponibles. La première regroupe un ensemble de modèles dits

de table, d'un prix très abordable, qui sont conçus pour la numérisation de documents récents, où les caractères sont bien contrastés, et dont le format n'excède pas A3. La deuxième classe de scanners de livres, destinée au secteur des bibliothèques, tolère pour sa part des documents reliés (jusqu'à 50 cm d'épaisseur) d'un format allant jusqu'à A1. Ces scanners, pensés pour répondre aux besoins de plans de numérisation de grande ampleur, disposent d'un ensemble d'accessoires ergonomiques, tels que les commandes au pied. Ils intègrent un plan de travail, une balance à plateau et une vitre pour assurer une surface de numérisation plane, en vue d'optimiser la qualité des données obtenues.

Les grands plans de numérisation s'effectuent pour leur part à partir des microformes, lorsque ce mode de conservation est disponible, en raison de la haute productivité des scanners de microformes. Ces appareils haut de gamme sont en effet caractérisés par une automatisation très aboutie, qui limite l'intervention de l'opérateur au paramétrage initial, alors qu'ils offrent les mêmes possibilités de résolution que les scanners de livres. Cependant, les microformes témoignent souvent d'une qualité médiocre limitant la qualité du résultat de la numérisation.

Concernant le choix de résolution et son mode (trait, niveaux de gris, couleur) à adopter, ceux-ci seront dépendants des usages visés par les concepteurs du projet. Sur ce point, il convient de préciser que numériser dans une résolution maximale n'est pas systématiquement un choix judicieux : plus la résolution est grande, plus les fichiers images seront lourds, difficiles à gérer et coûteux en terme de stockage, mais aussi longs à afficher ou à télécharger lors de la consultation. À l'inverse, choisir une résolution faible et un format compressé nuira à la finesse du grain de l'image. Une entreprise de numérisation déployée en vue de la diffusion (en ligne ou sur place) des documents privilégiera ainsi une résolution moyenne, tandis qu'on optera pour une résolution maximale dans le cas où le numérique est envisagé comme solution de conservation.

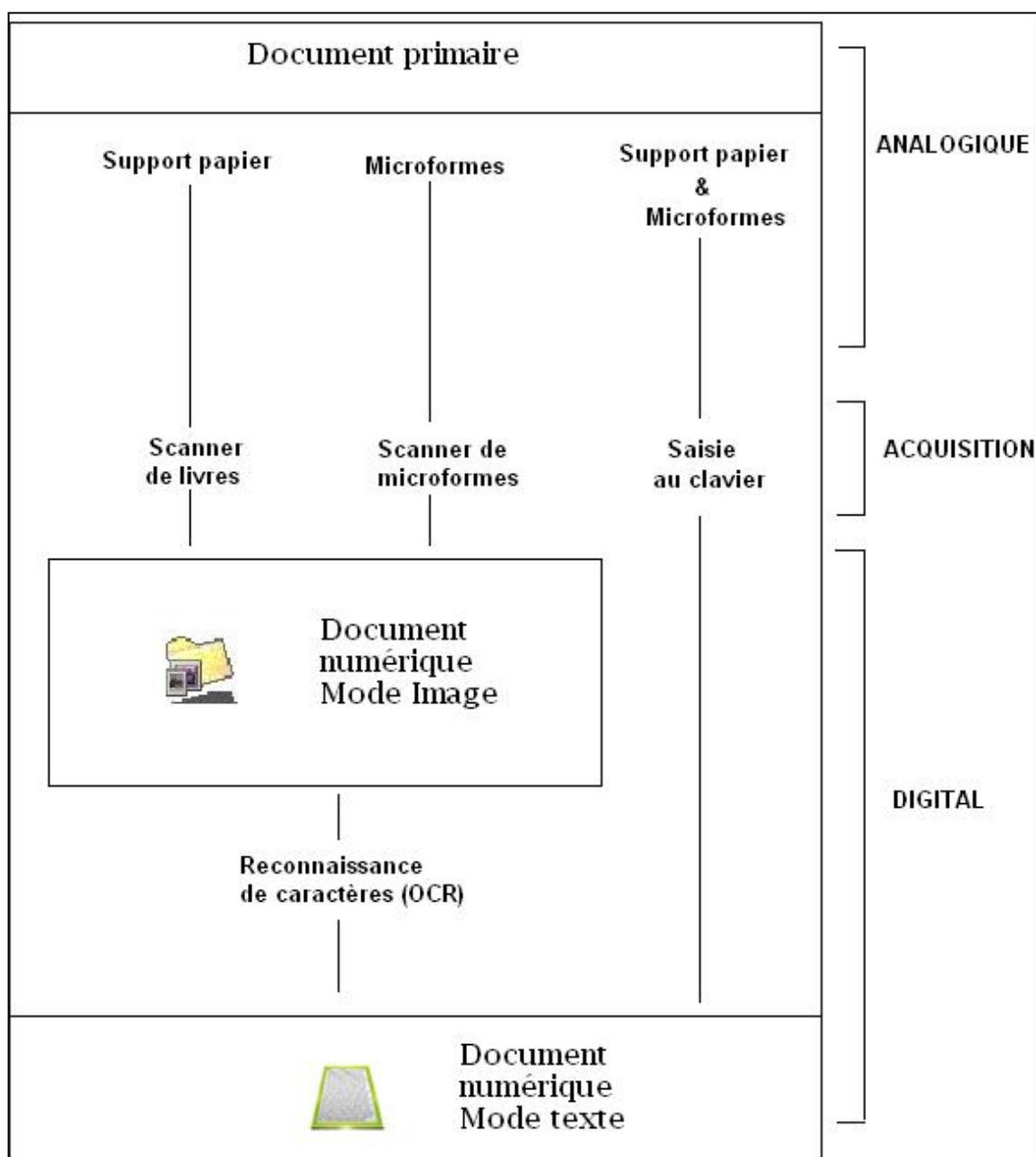


Figure I-1 : Schéma des techniques de numérisation de la presse ancienne.

Le mode image est privilégié par les bibliothèques

Motivés par un impératif majeur de conservation et de diffusion, les plans de numérisation de la presse régionale lancés en 2007 parallèlement à la numérisation des grands titres nationaux privilégient la construction de

réalisations numériques du mode image, à l'instar des choix opérés pour la presse ardéchoise (1891-1989), la presse locale de Montpellier (XIX^e- XX^e siècle)²³, celle de Maine-et-Loire (XVIII^e-XX^e siècle)²⁴, de la Champagne-Ardenne et de l'Isère²⁵.

Les choix de format de stockage et de résolution varient d'une bibliothèque à l'autre, suivant que le numérique est envisagé prioritairement comme un mode de conservation ou comme un mode de diffusion. De nos recherches, il ressort que la plupart des responsables de la numérisation ont néanmoins opté pour les formats JPEG, JFIF²⁶ ou PDF, avec une résolution de 200 ou 300 dpi. Ces formats compressés et allégés associés à une résolution moyenne répondent aux objectifs de diffusion et de valorisation de documents fragilisés ; ils seront néanmoins inutilisables *a posteriori* pour une numérisation plein texte. Rares sont les entreprises qui ont opté pour le format TIFF (*Tagged Image File Format*) : c'est néanmoins le choix opéré par les bibliothèques chargées de la numérisation de la presse du Morbihan et de l'Ain, et de la Bibliothèque Municipale d'Étude et de Conservation de Besançon (BMEC). Une proportion importante des entreprises de numérisation de la presse régionale ne prévoit, à court terme du moins, qu'une consultation locale des données images, stockées sur CD-ROM, à l'intérieur même de la bibliothèque. C'est notamment le cas réalisé par la Bibliothèque Municipale d'Epernay (Champagne-Ardenne), ayant procédé à la numérisation du *Polygraphe* (1843-1844)²⁷ et *L'Observateur sparnacien* (1903-1904), dont la version numérique est consultable uniquement sur place. La plupart des plans de numérisation

²³ http://www.numerique.culture.fr/mpf/pub-fr/document.html?id=FR-DC-B341726101_005&base=dcollection&hppage=10&totaldocs=74&pagenome=resultats.html&querybase=dcollection&qid=miKl-q&from1=rechgeo

²⁴ <http://www.numerique.culture.fr/mpf/pub-fr/document.html?base=dcollection&id=FR-DC-ZZZ-8712c8b0&hppage=10&totaldocs=100&pagenome=resultats.html&querybase=dcollection&from1=rechgeo>

²⁵ http://www.archinoe.fr/romans-archives/index.php?option=com_content&task=view&id=45&Itemid=51

²⁶ JPEG File Interchange Format

²⁷ http://www.numerique.culture.fr/mpf/pub-fr/document.html?base=dcollection&id=FR-DC-B512306201_5e83ce43

prévoyant une diffusion en ligne n'ont pas encore abouti. Parmi les programmes achevés de numérisation et de mise en ligne, on citera le cas des Archives Communales de Romans-sur-Isère, qui mettent à disposition une partie de la collection du journal politique hebdomadaire *Jacquemart* (1880-1905)²⁸.

De façon très majoritaire, l'acquisition des données est externalisée par le recours aux services d'opérateurs privés spécialisés²⁹.

1.3.2.2. La numérisation en mode texte

Lorsqu'un projet de numérisation vise à l'obtention des données en mode plein-texte³⁰, la chaîne de numérisation induit un nombre d'étapes supplémentaires et suppose des contraintes techniques pour l'acquisition des images très nettement supérieures à celles évoquées précédemment.

Comme on peut l'observer sur le schéma I-1, la numérisation en mode plein-texte implique alors une étape automatisée d'océrisation (*Optical Character Recognition*) exigeant que la qualité des fichiers image soit optimale. Il est donc capital d'opter en amont de cette étape pour une résolution de bonne qualité, en général égale ou supérieure à 300 dpi pour les documents anciens, sous peine d'obtenir des données en mode plein-texte inutilisables

Principes et états des savoirs faire en matière d'océrisation

Développés dans les années 1950 (Jacquesson & Rivier, 1999 : 85), les logiciels OCR déduisent la valeur Ascii des signes typographiques à partir de configurations de points. Si, dans leurs premières versions, ces logiciels ne pouvaient reconnaître que des textes composés dans des polices de caractères spécialement conçues à cet effet (OCR-A et OCR-B), ces logiciels sont aujourd'hui capables de lire la plupart des typographies. Une fois les caractères

²⁸http://www.archinoe.fr/romans/visu_affiche.php?PHPSID=d49610fb8c7da6e469175d700fbc687c¶m=visu&page=1

²⁹ Des entreprises telles que NTA (Puy-en-Velay) ou Azentis (Saint-Ouen) numérisent les documents anciens au tarif de 0,18€ la vue (à partir de 20 000 vues). Ainsi, pour 98157 vues du *Petit Comtois* (1887-1944), la BMEC a dépensé un budget de 18053,07€.

³⁰ Signalons ici qu'il existe d'autres solutions pour l'acquisition des données en mode plein-texte telles que la saisie clavier sur laquelle nous reviendrons ultérieurement, et la dictée orale.

extraits, qu'on nomme dans la littérature spécialisée « primitives élémentaires », des moteurs de reconnaissances statistiques et/ou structurels sont convoqués pour améliorer la qualité de la reconnaissance³¹.

Les résultats de ce type de procédure se révèlent excellents pour les documents modernes et les documents où la richesse du vocabulaire est limitée. *A contrario*, les résultats de l'OCR demeurent extrêmement fluctuants dès lors que le document d'origine présente certaines imperfections entraînant une perte de contraste et/ ou un vocabulaire diversifié. C'est ce que constatent A. Jacquesson et A. Rivier (1999), se risquant à océriser un texte relevant du discours littéraire datant du 18^e siècle³² :

La lecture des caractères est plus souvent ambiguë pour le programme, et les espacements ne sont pas toujours correctement interprétés. Le logiciel, de plus, s'appuie sur des dictionnaires pour confirmer sa traduction, à la défaveur des formes orthographiques vieilles. D'autres écueils restreignent l'intérêt de ce procédé : la présence de tableaux ou de schémas, les taches interférant avec le texte, des caractères qui se touchent, des fonds colorés ou très jaunis (Jacquesson & Rivier, 1999 : 86).

En raison de l'irrégularité des caractères imputables aux typographies anciennes, de l'état général du document, le taux de reconnaissance optique d'un document ancien est très faible. Conçus pour reconnaître des formes de langue moderne, les moteurs de reconnaissance sont incapables d'améliorer ce taux voire l'amoindrissent en substituant aux graphies anciennes non enregistrées dans leurs bibliothèques de forme des graphies du XX^e siècle.

Obstacles à l'océrisation de la presse ancienne

Au cours de nos recherches, nous n'avons trouvé aucun rapport mené sur les difficultés d'acquisition de données en mode plein texte posées plus spécifiquement par l'océrisation de la presse régionale de la Troisième

³¹ Le principe général de fonctionnement de l'OCR valable pour la majorité des logiciels est linéaire. Cependant, il convient de signaler les expérimentations d'Ogier et Al (2000), Ramel et al (1999) visant à réaliser des boucles dans la chaîne de traitement simultanément à la détection d'ambiguïtés.

³² A savoir un extrait de *Relations de différents voyages dans les Alpes du Faucigny*, 1776, p. 62, cité dans Jacquesson & Rivier (1999 : 87).

République ou plus largement de presse ancienne. Aussi avons-nous mené une étude de la qualité des données en mode texte de la seule entreprise ayant prétendu à cet objectif et présentant aujourd'hui les fruits de son plan de numérisation de la presse du XIX^e siècle : la Bibliothèque Nationale de France et son site *Gallica*.

Consciente de l'urgence d'un mode de conservation durable et des enjeux d'une diffusion élargie et de la valorisation des collections de presse ancienne, la BNF a lancé en 2005 un plan pluriannuel courant jusqu'à 2010 dont les objectifs sont de conserver le patrimoine écrit, enrichir les services à distance et faciliter l'accès aux documents libres de droits. Ce programme a pour visée de numériser une trentaine de titres de presse nationale, du début de leur parution jusqu'aux années 1940, conformément au respect du droit d'auteur. À l'exception de la collection de l'*Ouest-Eclair*, les titres numérisés par la BNF sont tous des grands quotidiens nationaux, parmi lesquels *Le Figaro*, *Le Temps*, *La Croix*, *L'Humanité*, *Le Journal des Débats*, etc. L'ensemble des documents à numériser représente près de 3,2 millions de pages selon les informations diffusées sur *Gallica*. À l'ampleur de ce projet correspond un budget tout aussi colossal : la BNF bénéficie d'un financement de 3, 5 millions d'euros, grâce au concours financier du Sénat et de la collaboration des quotidiens nationaux et d'*Ouest-France*.

Un cahier des charges très bien conçu...

Tout comme les BMEC, la BNF externalise la chaîne de numérisation en recourant aux services de sous-traitants pour l'acquisition des données proprement dite, leur océrisation et leur structuration. À partir des chartes de numérisation³³ et des cahiers des charges publiés conjointement aux offres de marché, ainsi que des renseignements diffusés par la BNF à l'attention des professionnels, il nous a été possible de retracer les principales modalités

³³ La description de l'étape ici présentée a été construite à partir du recueil d'informations mises en ligne par la BNF, notamment : http://www.bnf.fr/PAGES/infopro/numerisation/pdf/charte_ocr.pdf

techniques de la chaîne de numérisation effectuée et de constater la cohérence et la pertinence des choix opérés.

Pour ce qui est de l'acquisition des données, la BNF livre les supports-papiers à ses prestataires qui se chargent de les scanner. La BNF se charge au préalable de démonter les volumes reliés des numéros des quotidiens, de les dépoussiérer et de restaurer certaines parties dégradées des journaux. Le cahier des charges précise que jusqu'en 2008, l'acquisition des images s'est opérée exclusivement en niveaux de gris dans une résolution de 300 dpi. Offrant des résultats jugés très insatisfaisants, cette qualité a été abandonnée en 2009 au profit d'une résolution supérieure de 400 dpi³⁴. Les données sont ensuite enregistrées au format TIFF (*Tagged Image File Format*), format flexible non-compressé. Nous pouvons donc constater que l'étape d'acquisition des données a été conçue de sorte à obtenir la meilleure qualité image possible.

Dans un deuxième temps, les données images sont converties en données texte via une étape d'océrisation. Cette opération étant assurée par des prestataires, nous n'avons malheureusement pu accéder à la moindre documentation la concernant (choix du logiciel, type de moteurs de reconnaissance, paramètres adoptés, etc). Parallèlement à la conversion des données images en données texte, le prestataire est chargé de segmenter la page selon le format ALTO 5 (*Analyzed Layout and Text Object*). Basé sur le langage XML, ce format vise à conserver les coordonnées du contenu textuel et iconographique dans l'image et à permettre ainsi la corrélation des coordonnées d'un bloc image et d'un élément textuel interrogé. Le plan de la page est alors encodé selon cinq éléments génériques :

- Le *TopMargin* correspond ainsi à la zone de l'ours³⁵ du journal,

³⁴ Cf. http://www.bnf.fr/PAGES/infopro/numerisation/pdf/charte_numerisation.pdf

³⁵ Partie supérieure du journal recensant l'éditeur et le directeur de la publication, le nom et l'adresse de l'imprimeur, le nom du rédacteur en chef, etc.

- le *BottomMargin* correspond à la zone inférieure de la page où ne figure aucune zone de texte,
- le *LeftMargin* et le *RightMargin* correspondent pour leur part respectivement aux marges gauches et droites du journal où ne se trouve aucun élément textuel.
- Enfin, le *PrintSpace* désigne la zone de texte.

Une fois la page ainsi segmentée, la zone de texte (ou *PrintSpace*) est ensuite identifiée et partitionnée en quatre éléments *BlockGroup*. L'élément *TextBlock*, décrivant le bloc de texte, permet de regrouper les lignes de texte en un ensemble cohérent. L'entité *illustration* identifie pour sa part un élément iconique. Toutes les autres unités graphiques, comme les tableaux, sont quant à elles identifiées comme *GraphicalElement*. Enfin, l'entité *ComposedBlock* identifie l'imbrication de divers éléments (*TextBlock*, *illustration*, *GraphicalElement*).

À chaque unité documentaire sont associées des métadonnées (ou « notices ») en format *Dublin Core simple*³⁶. Ces notices sont ensuite intégrées au catalogue de la BNF, et par suite, à celui de *Gallica*, de sorte que l'utilisateur puisse en aval procéder à une recherche dans les champs des notices bibliographiques et sur le texte des documents numérisés. Précisons qu'à notre connaissance l'enrichissement des données textuelles *via* une TEI simplifiée que la BNF met en œuvre pour certains documents imprimés³⁷ n'a pas été appliqué aux imprimés périodiques.

En somme, le procédé technique développé par la BNF, visant à l'obtention de données plein-texte structurées et identifiées, se caractérise par des choix tout à fait pertinents d'acquisition des données image.

³⁶ Schéma générique de métadonnées permettant la description de ressources numériques et leurs mises en relation avec d'autres ressources. <http://dublincore.org/>

³⁷ Tels que la *Revue de Synthèse*.

Cf. http://www.bnf.fr/PAGES/infopro/numerisation/num_projets.htm#synthese

... pour des résultats médiocres

Quelle est maintenant la qualité des données plein-texte résultant de ce processus ? Telle est la question à laquelle nous allons tenter de répondre ici. Dans le dessein d'évaluer la qualité des données en mode texte, nous avons sélectionné dans *Gallica*³⁸ devenue *Gallica 2*³⁹, un numéro du titre *La Presse*, à savoir l'édition du 1^{er} août 1883, parce que cette date correspondait à la parution du premier numéro du *Petit Comtois*. La Une de ce numéro, sur laquelle nous avons concentré notre attention, ne présente aucune dégradation particulière. Comme on peut le constater sur la figure ci-dessous, la page numérisée en niveaux de gris est marquée par certaines pliures très localisées, sans que celles-ci compromettent de grands emfans de texte. Il s'agit donc d'un document de relativement bonne qualité, qui est d'ailleurs assez représentative de l'état de la collection de *La Presse* du même mois.

Considérons à présent la qualité du mode texte proposé par *Gallica*. Le site précise que « le taux de reconnaissance » de cette page atteint les 82%. Ce calcul repose sur l'attribution, par le logiciel d'OCR, pour chaque mot d'une valeur de fiabilité (*word confidence*, abrégé <WC>) pouvant aller de 0 à 10. Le taux de qualité de chaque page est ainsi calculé à partir de la somme des <WC> pour chaque page divisé par le nombre de mots. Ce qu'il est essentiel de noter, c'est qu'un taux de 82% qu'un néophyte pourrait juger *a priori* confortable est tout à fait insatisfaisant. Rappelons à la suite de H. Emptoz, F. Lebourgeois, V. Eglin et Y. Leydier (2003 : 109) qu'« un taux de reconnaissance de 99% se traduit par plus d'une erreur toutes les deux lignes ».

Pour évaluer de façon grossière la qualité de l'océrisation de la page présentée ci-dessus et examiner le nombre d'erreurs correspondant à un taux de reconnaissance de 82%, nous avons sélectionné de façon aléatoire un passage de cette Une : en l'occurrence, le début de la cinquième colonne.

³⁸ <http://gallica.bnf.fr/>

³⁹ <http://gallica2.bnf.fr/> , qui dès février 2009, sera gallica.fr

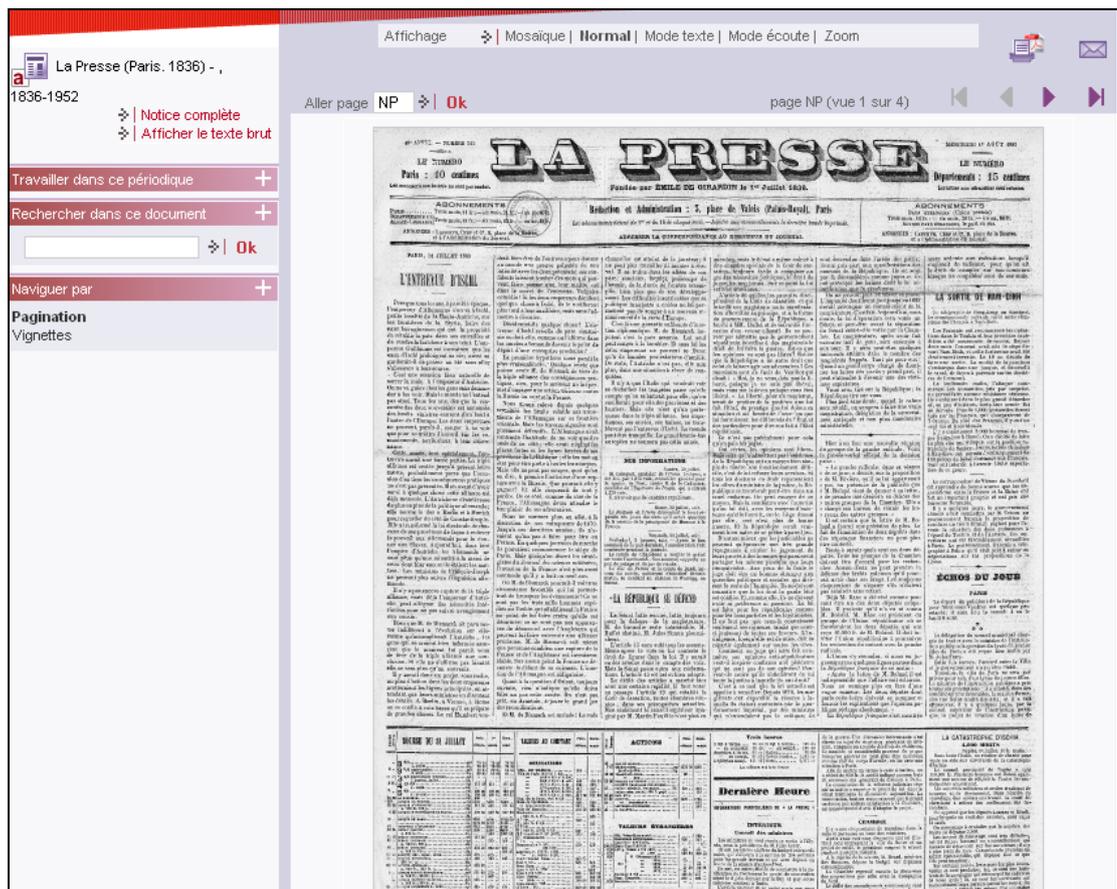


Figure I-2 : Capture d'écran. Interface de consultation de Gallica⁴⁰.

Pour la clarté de l'exposition, nous avons mis en parallèle ci-dessous ce fragment de la page sélectionné avec son équivalent plein texte fourni par Gallica, en ne procédant à aucune modification de mise en forme et de correction des données textuelles. Notre intervention sur le texte s'est limitée à colorer de rouge les erreurs de reconnaissance du texte relevées par nos soins.

D'emblée, on constate sur la figure I-3 que les erreurs colorées par nos soins sont nombreuses. Certaines erreurs sont imputables à la faible reconnaissance de l'accentuation des formes graphiques ; d'autres, à la suppression d'espaces figurant pourtant sur le support-papier, tandis qu'une partie des taches du document d'origine sont manifestement analysées comme

⁴⁰ <http://gallica.bnf.fr/ark:/12148/bpt6k544545w>

de la ponctuation. Une autre classe d'erreurs concerne la reconnaissance des caractères mêmes, aboutissant à la non-reconnaissance de mots entiers. À titre d'exemple, on notera que la lettre « t » est confondue avec un « l » ou avec un « i » et la lettre « n » est assimilée à la consonne « l », la lettre « u » est décomposée comme la répétition de deux lettres « i », etc. En définitive, l'exploitabilité des données océrisées est gravement compromise.

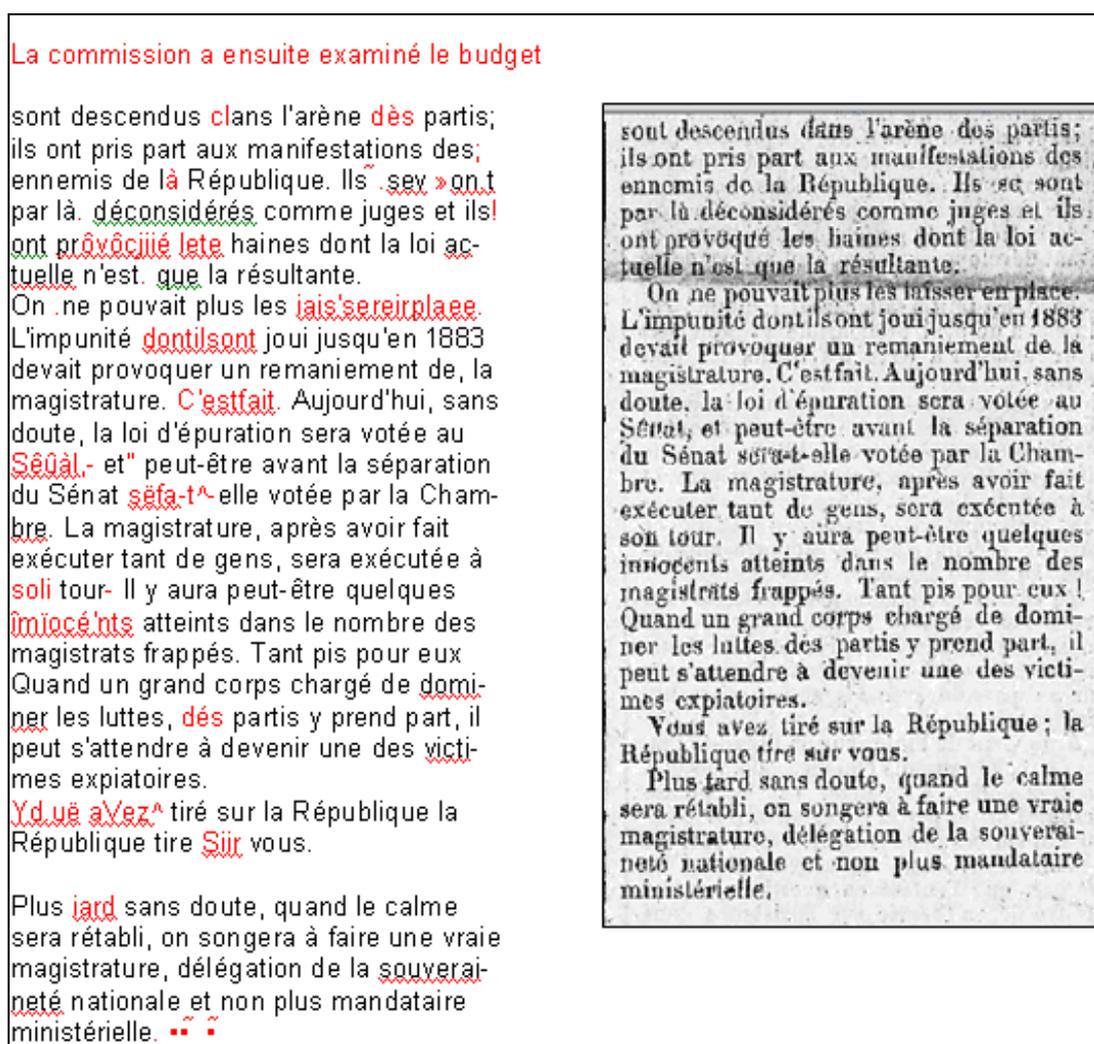


Figure I-3 : à gauche, le texte (OCR brut) ; à droite, l'image correspondante.

On aura également remarqué sur la figure ci-dessus une première séquence de caractères (mise en rouge par nos soins) n'apparaissant pas dans

la partie image. D'où provient cette suite de caractères non perceptible sur le bloc image supposé correspondant ? Il convient de considérer l'emplacement topographique du bloc image extrait. Si l'on (re)contextualise le bloc image dans l'espace de la page, on observe qu'il est localisé en haut de la cinquième colonne du numéro de *La Presse*, comme le montre la figure ci-dessous :

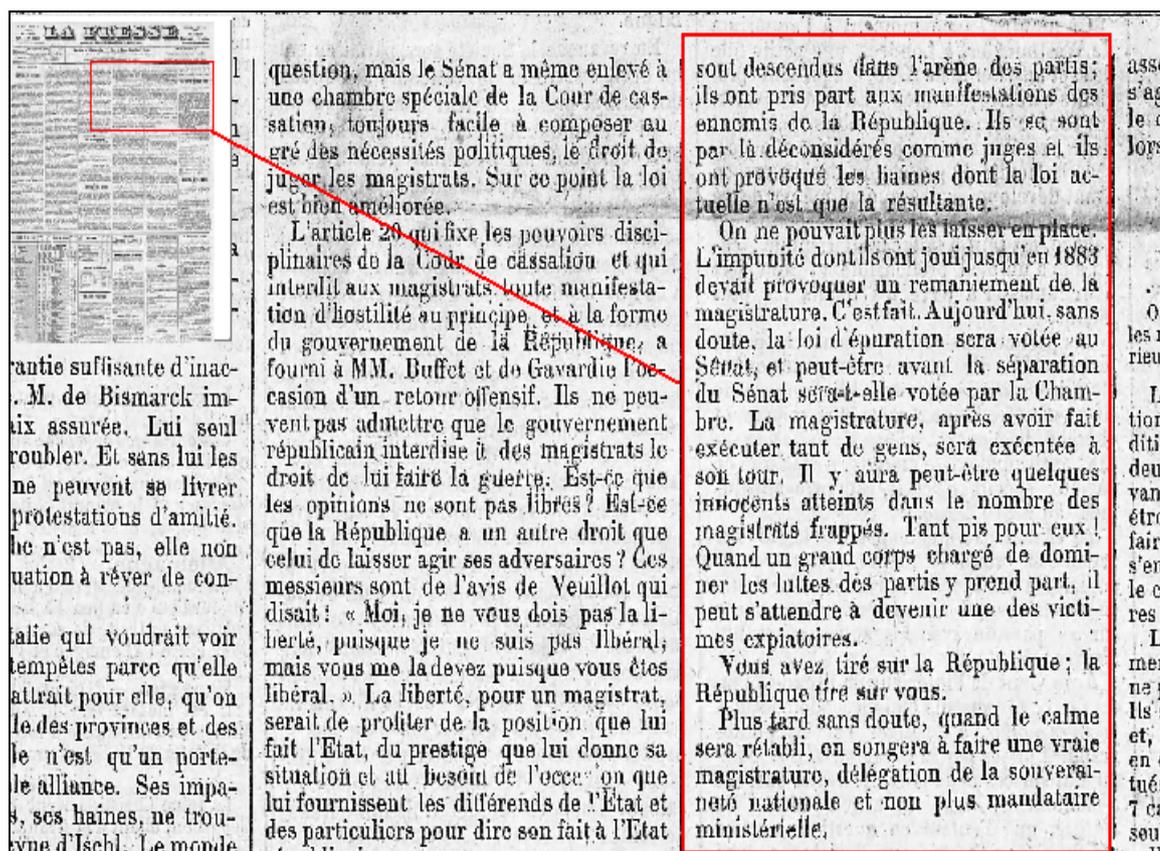


Figure I-4 : Zoom sur le bloc image du haut de la 5^e colonne de la page 1 du numéro de *La Presse* (1883-08-01).

L'énoncé « la commission a ensuite examiné le budget » n'apparaît pas dans l'article qui occupe le haut des trois premières colonnes de *La Presse*, comme l'atteste la figure I-4. Il se localise au contraire dans la dernière ligne de la quatrième colonne du rez-de-chaussée, ce qu'illustre la figure I-5. Cet exemple, qui pourrait paraître anodin, pointe un défaut fondamental des données en mode plein-texte de *Gallica*. Celles-ci sont constituées à partir d'une lecture

verticale du document image, sans considération relative à la structuration de la page de journal. Pour le logiciel d'océrisation, il y a continuité entre le texte du rez-de-chaussée d'une colonne A et le bandeau du journal situé en colonne B. D'où le fait que les données textuelles se succèdent sans respect de la progression du texte.

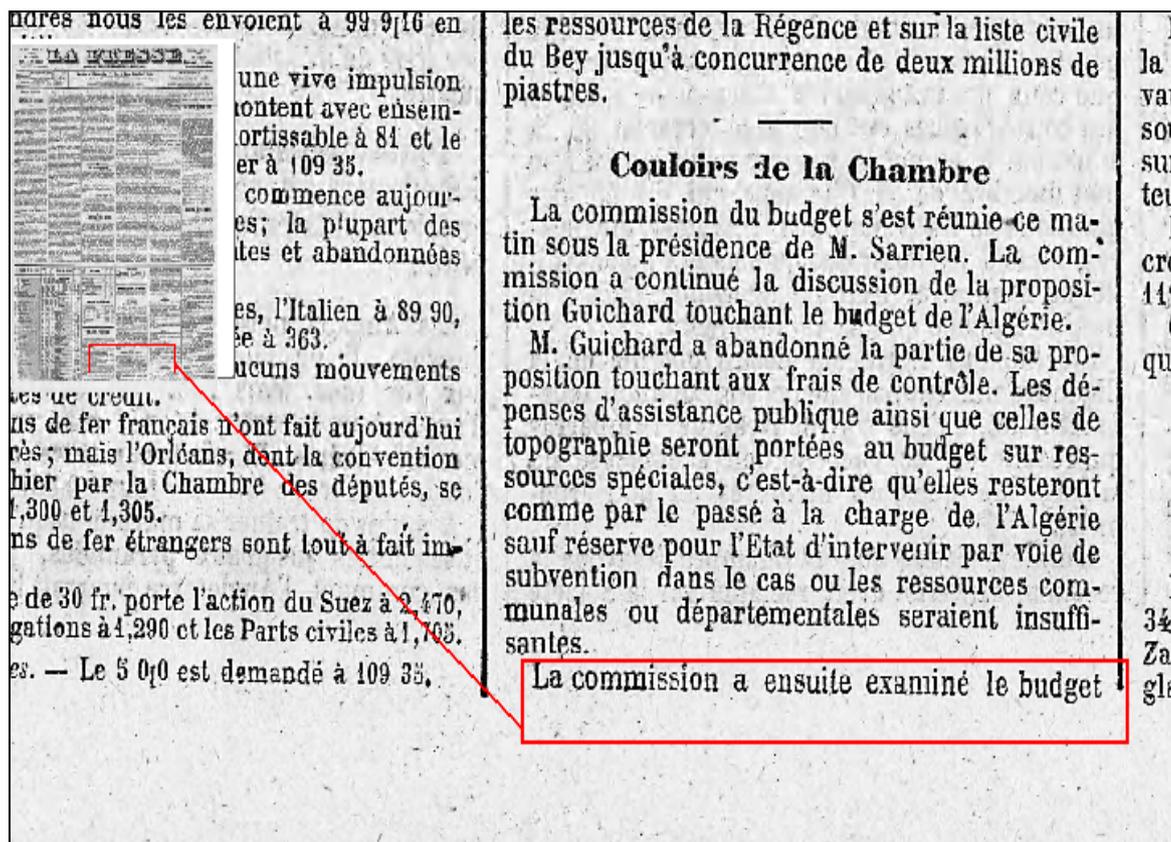


Figure I-5 : Zoom sur le rez-de-chaussée de la 4^e colonne de la page 1 du numéro de *La Presse* (1883-08-01).

En somme, au faible taux de reconnaissance des caractères constaté, s'ajoute un problème de non-respect de l'unité et de la cohésion des unités rédactionnelles, qui aggrave la qualité très insuffisante des données textuelles océrisées présentées dans *Gallica*. L'intérêt d'une océrisation de cette qualité

se limite dès lors à assurer l'indexation des documents et à permettre, par la suite, la requête lexicale en arrière-plan des PDF⁴¹.

C'est d'ailleurs l'objectif auquel se restreignent les rares projets de numérisation de la presse régionale de la Troisième République menés par les bibliothèques municipales procédant à l'océrisation des données. Ainsi les campagnes de numérisation de la presse régionale de l'Ain⁴² (19^e-20^e siècle), de la presse Ardéchoise (1891-1989)⁴³, la presse locale de la Meuse (1828-1940), ou encore la collection du *Journal de Toulouse et du Midi Socialiste (1908-1944)* ont opéré en externe une numérisation⁴⁴ en vue d'offrir aux utilisateurs la possibilité de requête lexicale en arrière-plan du PDF et non de fournir une transcription fidèle du document en mode texte. Les ressources d'arrière-plan sont d'ailleurs inaccessibles. La réussite des requêtes d'occurrences lexicales, se traduisant par une mise en surbrillance sur le PDF du terme convoqué, reste dépendante du taux de reconnaissance de l'OCR. Pour idée, si l'appel d'offre établi par la Bibliothèque Municipale de Toulouse spécifiait un taux de reconnaissance optique des caractères de 90%, les résultats de l'océrisation se sont révélés proches de 60%⁴⁵ en raison de la dégradation des documents.

⁴¹ À ce sujet, il est à signaler que l'ergonomie de l'interface en vue d'une recherche d'occurrence(s) est très faible dans *Gallica* : par exemple, la recherche d'une occurrence se limite à un seul titre et ne peut porter sur l'ensemble des collections de presse ancienne. Les résultats apparaissent alors sous forme de listes répertoriant chaque numéro présentant une (ou plusieurs) occurrence(s) du mot recherché. Il sera à la charge de l'utilisateur d'afficher le texte brut et de le parcourir dans son intégralité pour exporter les contextes contenant l'occurrence interrogée.

⁴² <http://memoire.arald.org/fr/presse.php>

⁴³ <http://www.ardeche.fr/education-culture-jeunesse-sports/archives-departementales/recherche/evenement-presse-ardechoise>

⁴⁴ Il est intéressant d'observer les différents choix techniques opérés dans ce cadre : la description de la campagne de numérisation de la Meuse précise ainsi que les documents à océriser sont des fichiers images de type JPEG issus de la numérisation de micro-films. Ce choix, qui nous laisse très sceptique quant à la viabilité même du projet, pointe le manque de concertation entre l'institution qu'est la Bibliothèque Municipale et les experts de la numérisation.

⁴⁵ Indication fournie par F. Courtial, Conservateur, BMVR de Toulouse.

De ce focus sur les résultats de la BNF en matière d'océrisation, il ressort que seul un contrôle manuel des données textuelles issues de l'océrisation garantit une transcription fidèle du document primaire. Condition *sine qua non* de l'exploitabilité des données océrisées, cette phase de correction représente néanmoins une opération longue et fastidieuse, qui engendre des coûts de main-d'œuvre potentiellement très lourds. Les résultats peu probants de l'océrisation appliquée aux documents anciens diminuent ainsi très nettement l'attractivité de cette procédure par rapport à la saisie manuelle. Certaines entreprises européennes proposent ainsi un service de saisie clavier opéré par des typistes dans les pays en voie de développement. Ce type de délocalisation, tirant intérêt des médiocres conditions de rémunération ainsi que de la souplesse (voire de l'inexistence) du droit du travail pour proposer des tarifs à tous égards « compétitifs »⁴⁶, soulève bien évidemment de sérieux problèmes éthiques. Sur le plan scientifique, on peut légitimement mettre en cause la qualité d'une acquisition de données textuelles qui ne serait pas opérée par des usagers experts aptes à réaliser et à valider des choix philologiques. La problématique des limites de l'océrisation appliquée à la presse ancienne et son corollaire, le développement d'un marché de saisie directe délocalisé, nous semblent mettre en relief l'urgence du développement de la recherche en matière d'acquisition de données textuelles, dont les enjeux scientifiques se doublent d'une dimension éthique et politique, comme nous le verrons *infra*.

⁴⁶ Une entreprise telle que Sun-groupe SARL, basée dans le canton de Neuchâtel, à Pékin et dans le Hénan (République Populaire de Chine), se targuant « d'allier la précision, le sérieux et la confiance que l'on témoigne à la Suisse avec la qualité, la persévérance ainsi que la performance de la Chine dans la même entreprise » (<http://www.sun-groupe.com/>) propose un tarif de 0,75 euro pour 1000 caractères transcrits. Pour ordre de comparaison, la saisie directe des données textuelles *Newsbooks* menée par le Département de linguistique de l'Université de Lancaster était rémunérée à hauteur de 10 GBP (soit un peu plus de 11 euros) pour mille mots.

Conclusion

De ce retour sur l'approche de la presse régionale de la Troisième République en sciences humaines, sur les mutations épistémologiques concourant à mettre la matérialité textuelle au centre du processus de connaissance, et enfin sur les savoir-faire relatifs à l'acquisition des données, ressort l'impétueuse nécessité d'un chantier de construction d'une base textuelle de presse régionale de la Troisième République.

S'il est un chantier des plus décisifs, c'est tout d'abord en raison du rôle déterminant qu'il tient dans la dynamisation des lectures d'une source peu explorée. Conditionnant l'amélioration de l'état des connaissances des formes génériques d'une presse évitée par l'analyse du discours et les sciences de la communication, la PQR3R est une ressource dont les apports sont des plus précieux pour le renouvellement des savoirs historiographiques relatifs à une période fondatrice appréhendée à l'aune de ses concepts socio-politiques. Répondant aux attentes d'une histoire linguistique des concepts désireuse de s'appuyer sur une archive diversifiée, la presse régionale de la Troisième République constitue un dispositif majeur d'accès aux concepts circulant dans l'espace public (Habermas, 1978). Complétant les indications fournies par les ressources habituellement convoquées (monographies, dictionnaires, textes philosophiques et/ou politiques) en vue d'observer les concepts, lieux et enjeux de la lutte des forces socio-politiques en présence, la PQR3R se pose comme cette source apte à témoigner de la formation de l'opinion sous la Troisième République.

Les enjeux de la constitution d'une base de PQR3R sont en second lieu méthodologico-techniques. Partant du constat de bon sens selon lequel « nous n'appliquons jamais à nos données que les traitements qu'elles peuvent supporter (Prost, 1974 : 123-124) » et de celui des avancées de l'analyse textuelle du discours et de la logométrie, convergeant vers un établissement philologique des textes, les avancées des sciences du texte sont aujourd'hui

dépendantes d'une réflexion sur les techniques d'acquisition des données textuelles. Dans un contexte technologique extrêmement évolutif, la constitution d'une base textuelle de presse ancienne figure ainsi un laboratoire d'expérimentations méthodologiques et techniques aux objectifs pluriels. Il s'agit, d'une part, de lever les obstacles pesant sur l'acquisition de données de presse ancienne en mode plein-texte, en vue de rendre la constitution de telles bases compatibles avec les exigences du contexte de recherche actuel où le temps et les moyens financiers sont chichement comptés. Il s'agit d'autre part, et conjointement, de proposer des solutions méthodologiques pour l'établissement et la normalisation des données textuelles, afin d'en garantir la pertinence, la portabilité et la pérennité.

Chapitre II.

Constitution du dispositif d'exploration du *Petit Comtois*

Notre ambition est d'autoriser, à l'heure du numérique, un renouveau des lectures d'une presse ancienne susceptible de livrer un précieux témoignage sur un siècle fondateur. Il s'agit donc de constituer une *base textuelle* de presse quotidienne régionale de la Troisième République qui réponde parallèlement à un besoin prioritaire des sciences des textes : celui de disposer de ressources textuelles non plus construites pour les besoins d'un projet isolé, sujettes à la déshérence et à l'obsolescence, mais au contraire des ressources textuelles partageables et finement établies dans une perspective philologique. De la disponibilité de telles données dépendent non seulement un renouveau des savoirs historiographiques, mais également la mise à l'épreuve des outils méthodologiques des sciences des textes, ainsi que l'intensification du dialogue critique dans cette communauté, que nous entendons comme autant d'exigences scientifiques.

Les caractéristiques de la presse ancienne, le volume des données à traiter et surtout l'objectif de disposer de données plein-texte, ont induit un protocole expérimental complexe de numérisation, auquel nous consacrerons le premier temps de ce chapitre. Il s'agira ici de rendre compte des difficultés rencontrées lors de notre entreprise d'acquisition de données en mode plein-texte structurées, établies et normalisées ainsi que de pointer les pistes méthodologiques élaborées et programmées, issues de nos expérimentations. *A contrario* d'une conception dominante de la numérisation comme « mal

nécessaire » à la recherche dont elle ne saurait relever du fait de son intérêt scientifique mineur, – comme le résume M. A. Chabin, remarquant que la numérisation « n'est que la première étape de l'important : le processus de recherche et non le produit de ce processus » (Chabin, 1999 : 9) – , nous montrerons en quoi la numérisation fait partie intégrante de la recherche. Dispositif central pour le développement des sciences du texte, elle est un temps de choix techniques et herméneutiques déterminants, qui conditionnent en retour les possibilités de lecture du texte et le partage de ces ressources.

Nous nous attacherons dans un second temps à décrire le corpus étudié, en insistant sur le mouvement par lequel nous avons constitué en *corpus* les données de la *base*. Si corpus et base se sont très provisoirement confondus dans l'étape concrète d'acquisition des données, notre corpus procède d'une sélection dans l'archive de données textuelles organisées en vue de constituer un dispositif d'observation du discours d'un quotidien républicain de province durant les débuts de la Troisième République (1883-1903). À la présentation de ce corpus, succédera une description de ses caractéristiques.

2.1. Du support papier à la base textuelle

2.1.1. Brève présentation du projet de base textuelle PQR3R

Avant de décrire le processus d'acquisition des données de la base PQR3R, il convient d'évoquer brièvement les origines de ce projet. D'emblée, il nous semble intéressant, sinon révélateur du rôle de l'acquisition et l'établissement des données textuelles dans la réunion des sciences du texte et de l'analyse du discours, de remarquer que l'idée d'une base textuelle de PQR3R a émergé du dialogue entre deux laboratoires : le Centre Jacques Petit (Littérature Française, section 9 du CNU) et le Laseldi (Linguistique, section 7).

Initié en 2003, cet ambitieux projet s'est matérialisé en 2005 par une campagne « inaugurale » de numérisation du quotidien *Le Petit Comtois* (1883-1944), dont nous dresserons ici le bilan. Cette première campagne d'acquisition

répondait à un double objectif : d'une part, construire les premières données textuelles *intégrales, établies* du point de vue philologique et *normalisées* de la base PQR3R, desquelles serait extrait le corpus étudié dans notre thèse ; d'autre part, offrir des propositions pour l'optimisation des traitements techniques impliqués par la constitution d'une base textuelle. L'ampleur de ce projet a nécessité l'association d'un grand nombre de partenaires, tels que la Bibliothèque Municipale d'Etude et de Conservation de Besançon et sa tutelle la Ville de Besançon, la MSHE Claude-Nicolas Ledoux et le Conseil régional de Franche-Comté. Ces derniers ont subvenu aux investissements financiers liés à l'achat de matériel (logiciels, scanner, poste fixe, etc.), ainsi qu'aux vacations des opérateurs. Ils nous ont également permis de nous consacrer pleinement à ce projet à l'aide d'une allocation de recherche.

La chaîne de numérisation de la collection, schématisée ci-dessous, s'est déroulée en cinq temps principaux : d'abord l'acquisition optique des données sous mode image, la restauration des fichiers ainsi obtenus, leur conversion en mode plein texte par océrisation, puis la correction des données textuelles et leur structuration ; enfin, leur conversion en format standard selon les recommandations de la TEI. Signalons que pour cette dernière étape, nous avons bénéficié des conseils précieux du réseau ATONET (*Réseau pour l'échange de ressources et de méthodologies en analyse de texte assistée par ordinateur*), qui nous a fait partager son expérience de construction de données normalisées. Au cours d'un stage en août 2008, nous avons en effet eu l'occasion de suivre une école d'été d'*Analyse de Textes par Ordinateur* (ATO), pour reprendre la terminologie en vigueur chez nos collègues québécois, et de profiter des précieux conseils de F. Daoust qui nous a accordé de son temps pour mettre en place un processus de normalisation des données du *Petit Comtois*. Nous avons également bénéficié de la proximité du Pr. Lou Burnard, l'un des animateurs du consortium TEI, membre du Conseil Scientifique de notre MSHE, qui a utilisé nos données pour expérimenter ses développements les plus récents.

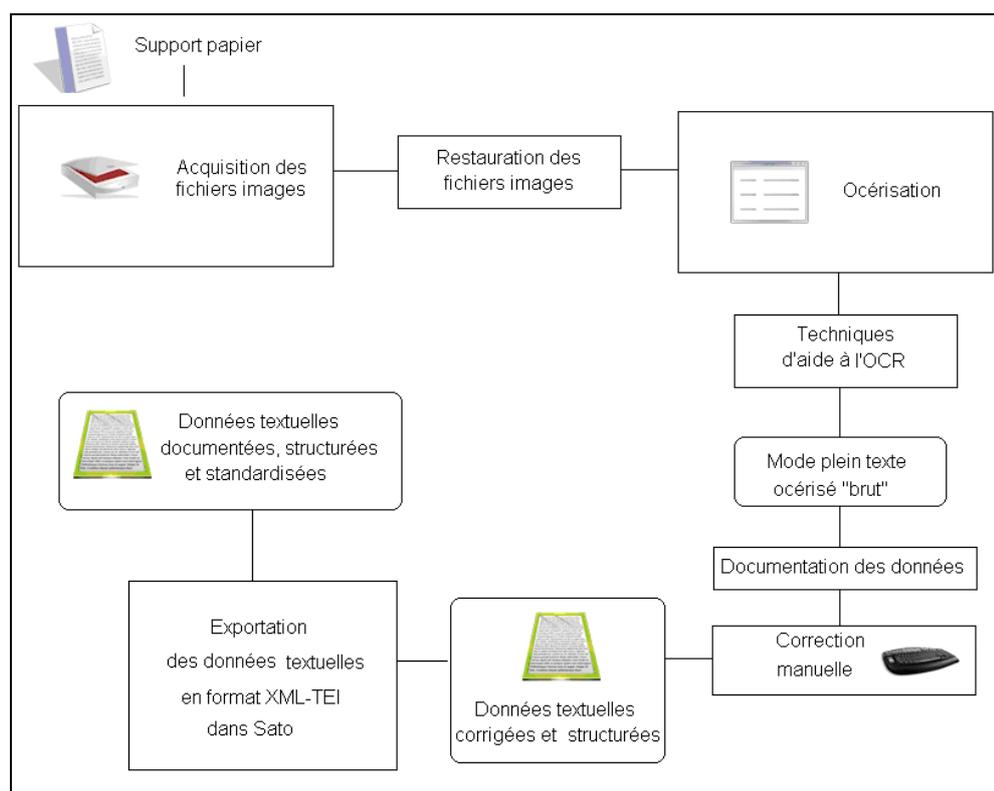


Figure II-1 : La numérisation du *Petit Comtois*.

2.1.2. Acquisition des données image

Le premier module d'acquisition des données a consisté à convertir un document papier en une image numérique par le balayage optique du support. Conformément aux accords du partenariat mis en place entre le pôle *Archive, Bases, Corpus* de la MSHE Ledoux et la Bibliothèque Municipale d'Étude et de Conservation de Besançon (désormais BMEC), la réalisation de cette étape a été confiée aux opérateurs de cette dernière, sur la base d'un cahier des charges établi par l'équipe universitaire. Les objectifs patrimoniaux et scientifiques de chacun des partenaires ont donné lieu à l'achat d'un matériel adapté, satisfaisant à la qualité des ressources numérisées attendue tout comme aux normes de non-destructivité des collections papiers. Le projet de constitution de la base plein-texte a contribué à poser comme condition *sine qua non* la qualité des images numérisées obtenues au cours de ce module.

Les objectifs scientifiques visés par le pôle *Archive, Bases, Corpus* de la MSHE Ledoux ne se limitant pas à une simple indexation du contenu du quotidien, la qualité de la prise d'image se devait en effet d'être la plus grande possible pour satisfaire aux exigences du logiciel de reconnaissance de caractères utilisé. De ce côté, le choix du matériel à adopter a ainsi été guidé par un souci de performance, tant au niveau des possibilités de résolution du scanner que de sa capacité à améliorer le contraste des images obtenues à partir de feuillets particulièrement bruités. À cette exigence de performance, se sont ajoutées celles de tolérance du format des volumes reliés du *Petit Comtois* et de respect des normes de conservation portant sur les collections.

Il a été décidé, suivant ces impératifs discutés collégialement, de procéder à l'achat d'un poste de numérisation *Digibook*, répondant au cahier des charges précédemment précisé. Acceptant des formats pouvant aller jusqu'à DIN A1 (A X A2), *Digibook* ne produit pas de chaleur sur le papier, et répond de fait aux normes de conservation instituées. Ses performances de numérisation sont quant à elles très intéressantes, puisque la résolution (500 dpi) peut être optimisée par différents réglages (hauteur de la caméra, mise au point, diaphragme manuel). De surcroît, le temps de lecture optique, tout comme le temps de transfert vers le disque dur local est très court. Ce scanner haute-productivité a été reçu et mis en place en décembre 2005 à la BMEC. Les premiers essais sur les numéros papier du *Petit Comtois* ont aussitôt été effectués par les opérateurs de la BMEC, qui se sont empressés de nous présenter les résultats obtenus. Après examen de ces premiers fichiers PDF, numérisés dans un niveau de résolution de 200 dpi et de 500 dpi, nous avons décidé (1) que la numérisation des supports-papiers devrait être réalisée en 300 dpi, niveau de résolution optimisant la poursuite de la chaîne de numérisation, par l'équilibre satisfaisant entre informations et bruits qu'il présentait ; (2) d'enregistrer et de stocker les images sous mode TIFF (*Tagged Image Format File*). Ce format présente l'avantage de limiter le risque d'une

compression destructive, contrairement à un mode JPEG (*Joint Photographic Expert Group*).

L'acquisition des données a débuté en Janvier 2006 et s'est étendue sur une période de 20 mois. Nous prenions toutes les deux semaines réception des fichiers numérisés du quotidien. Chaque lot de fichier image était dès lors traité selon les étapes de la chaîne de numérisation décrite plus haut. Huit mois environ après le début des opérations, au terme des premiers essais de réalisation de la chaîne complète d'acquisition des données en mode plein-texte, force était de constater que la conversion de chaque numéro du *Petit Comtois* en données textuelles exploitables ne pouvait être opérée dans un laps de temps inférieur à deux heures. La nature chronophage de ces opérations entrant en conflit avec les conditions actuelles de la recherche doctorale, nous avons donc été contraints de revoir à la baisse l'ampleur de la série textuelle que nous projetions d'analyser et, somme toute, la globalité de notre projet de recherche. Ce faisant, l'expérimentation des difficultés de constitution de données numérisées de presse n'a fait que confirmer la nécessité d'élaborer des chaînes de numérisation et de correction abouties et productives pour la création de grandes bases textuelles. Les conditions empiriques d'acquisition des données textuelles ont ainsi directement induit certains choix méthodologiques quant à la constitution de notre corpus de recherche. Deux solutions s'offraient à nous : numériser en continu les numéros du *Petit Comtois* et à terme, avoir la possibilité d'étudier le discours du quotidien d'août 1883 au début de l'année 1885, ou privilégier la diachronie en procédant à un sondage d'un certain nombre de numéros du quotidien par année de parution. Nous avons choisi de privilégier cette dernière solution et de procéder par sondage : nous verrons plus loin comment nous justifions précisément la constitution de cet échantillonnage. Pour l'instant, nous préférons expliquer les obstacles que notre projet expérimental a rencontrés.

La difficulté majeure contrariant la réalisation d'un projet de numérisation de presse ancienne est avant tout la qualité médiocre des supports physiques à

traiter. Aussi convient-il de décrire brièvement les caractéristiques de ces derniers. Signalons d'emblée que chaque exemplaire papier du *Petit Comtois* est composé de quatre pages¹, dont la qualité est sensiblement identique à celle du numéro reproduit ci-dessous. Cet état médiocre du support-papier est le résultat conjoint d'une consultation humaine soutenue et de l'œuvre du temps sur un papier de faible qualité. On notera ainsi sur toutes les pages du numéro de nombreuses taches circulaires, vestiges d'empreintes graisseuses. De surcroît, la surface du document est généralement marquée d'un pli horizontal du journal, ayant pour conséquence la déformation et l'illisibilité de nombreux caractères. Fréquemment, on rencontre sur la page du quotidien une pliure verticale extrêmement prononcée, résultat du mode de conservation des supports papier : qu'une infime partie d'un support papier soit froissée et, sous le poids du volume, un empan entier du texte devient indéchiffrable. L'illisibilité du texte peut également être causée par la méthode d'impression du journal : certaines bandes de papier semblent avoir été froissées pendant le passage dans la rotative.

Enfin, la lisibilité des caractères de certains passages est compromise par la transparence du *verso* sur le *recto*, et ce, tout spécialement dans la troisième page du quotidien, à travers laquelle apparaissent les caractères de la quatrième page. Les publicités, avis et annonces, essentiellement concentrés dans la surface de cette page au long de la parution du quotidien, sont mis en forme par un système d'encrage extrêmement prononcé, qui a pour effet de transpercer le papier de qualité médiocre utilisé pour l'impression du quotidien.

¹ Hors numéro exceptionnels.

LE PETIT COMTOIS

JOURNAL RÉPUBLICAIN DÉMOCRATIQUE QUOTIDIEN

RÉDACTEUR EN CHEF, JULES GROS

ABONNEMENTS		RÉDACTION ET ADMINISTRATION		INSERTIONS	
France et colonies	3 mois 8 fr. 6 mois 15 fr. 1 an 30 fr.	BESANCON — 7, Square Saint-Amour, 7 — BESANCON	Annouces	La ligne 30 c.	
Pays étrangers	3 mois 10 fr. 6 mois 18 fr. 1 an 35 fr.	Les abonnements partent des 1 ^{er} et 16 de chaque mois	Réclames	50 c.	
Édition hebdomadaire	3 fr.		Fautes diverses	4 fr.	

CHEMIN DE FER. — Service d'été. — L'heure de la gare est en retard de 16 minutes sur celle de la ville.

Départ de Besançon : Pour Dijon : Matin, 5 h. 13, 8 h. 07; Pour Belfort : Matin, 8 h. 12; Pour Besançon : Matin, 7 h. 14, 8 h. 22, 9 h. 30; Soir, 2 h. 25, 5 h. 20, 9 h. 46; Soir, 2 h. 25, 6 h. 22.

Arrivée à Besançon : De Dijon : Matin, 5 h. 45, 7 h. 21, 9 h. 28; De Belfort : Matin, 7 h. 53; De Besançon : Matin, 7 h. 53, 9 h. 17; Soir, 1 h. 47, 4 h. 37, 8 h. 06.

De Vesoul : Matin, 7 h. 11, 9 h. 17; Soir, 1 h. 56, 4 h. 46, 8 h. 15.

De Gray : Matin, 7 h. 12; Soir, 1 h. 52, 4 h. 42, 8 h. 11.

AVIS

L'Administration et la Rédaction du Petit Comtois croient n'avoir rien négligé pour assurer au lecteur un service d'informations intéressantes, complètes et rapides.

Nous donnerons la fin des séances un jour avant le journal de Paris et plusieurs heures avant les feuilles locales.

Nous nous sommes assurés de la collaboration d'un grand nombre de personnes très informées dans les départements limitrophes, et nous serons en situation de donner au lecteur le mouvement politique de la ville, pour toute la région et dès les premières heures du jour.

Nous comptons, du reste, améliorer avec le temps le service particulier des informations; aussi prions-nous le public pour lequel nous travaillons de se montrer indulgent pour les imperfections toujours inséparables du débet.

Le choix des feuilletons sera de notre part l'objet d'un soin tout particulier. Aujourd'hui, notre but est d'éveiller le sentiment critique du lecteur, en opposant deux œuvres d'un caractère très différent.

L'empoisonneuse, roman éminemment dramatique, d'un auteur très apprécié à Paris, séduira par le mouvement et l'étrangeté des situations; Fromont jeune et Risler aîné, l'auteur célèbre d'*Alphonse Beauclerc*, sera prisé surtout par les gouvernés, qui demandent à l'écrivain les qualités de l'observateur.

EXPLIQUONS-NOUS

On nous demande qui nous sommes. Nous allons le dire : En 1880, c'est-à-dire à une époque où ceux qui nous accablent aujourd'hui de jeter la division dans le parti républicain gardaient un silence diplomatique, nous

FEUILLETON DU PETIT COMTOIS
Du 1^{er} août 1883.

L'EMPOISONNEUSE

Par PIERRE NINOUS
Première partie
LE DIABRE DE ROQUEBERRE

LA DERNIÈRE HEURE

Le 29 décembre 1863, vers sept heures du soir, la petite ville de Roqueberre, dans l'Armagnac, était en proie à une animation qui contrastait singulièrement avec sa tranquillité de chaque jour.

La plupart des habitants, sur le pas de leur porte, chuchotaient entre eux.

Les domestiques des grandes maisons couraient vers le quartier bas de la ville, bousculant leurs maîtres, qui se dirigeaient en toute hâte du même côté; la cloche d'agonie de la vieille église envoyait lentement dans les airs ses tintements égaux et tristes; il était évident que quelque chose de considérable se mourait à Roqueberre.

nous sommes rencontrés un certain nombre d'hommes résolus à dire leur opinion. Nous pensions que l'empire était vermoulu, et, en tout cas, vicié par son origine. Si nous avions pu prévoir que ce que nous combattons s'effondrerait dans le sanglant gâchis de 1870, nous aurions mis une sourdine à nos voix, parce que, hier comme en ce moment, nous plaçons la patrie bien avant nos préférences personnelles.

Aujourd'hui, nous nous retrouvons les mêmes hommes, à très peu de chose près, non moins résolus à faire connaître notre sentiment, et ne réclamant en cela que le droit inaliénable qui appartient à tous les citoyens. Or, notre sentiment, le voici :

Il y a tout juste quarante ans que M. de Montalembert, prenant ouvertement la défense des jésuites au nom de la liberté, contre M. Villemain et l'Université, s'écriait dans un discours resté célèbre : « Nous sommes les fils des croisés, et nous ne reculerons pas devant les fils de Voltaire. »

Ce mot était peut-être bien placé dans la bouche d'un descendant des vieilles familles aristocratiques. Mais il n'engageait pas davantage le pays que ne l'engagent plus tard les quelques douzaines d'illuminés qui consacreront la France au Sacré-Cœur.

Nous sommes, nous, les nouvelles couches sociales dont a parlé Gambetta. Loin de renier notre origine, nous bravons les dédains que cette parole a suscités, car c'est à ces couches nouvelles, suivant nous, que l'avenir appartient. Nous pensons, en effet, que les anciennes aristocraties ont irrémédiablement perdu le pouvoir pour n'avoir su, depuis bientôt un siècle, tenter que de stériles efforts de retard vers le passé.

Donc nous ne renions rien de notre ancienne politique. S'il était nécessaire, nous dénoncerions encore l'insanité de ces démocraties césariennes, ou l'on prétend dominer le peuple en adulant sans cesse ses instincts, bons ou mauvais. Mais le danger n'est point là aujourd'hui. Il n'est pas davantage dans la monarchie, morte bien longtemps avant M. de Chambord. Les

libériciens de la domination du peuple par l'abrutissement n'ont pas le profil de César. On les retrouverait au besoin dans les oligarchies, de quelque nature qu'elles soient.

Le danger, il est dans l'avortement des réformes sans cesse promises et toujours éludées. Il est dans cette alliance singulière des libéraux tremblants avec les intrançaisants, qui repoussent toutes les améliorations, sous prétexte qu'elles ne sont pas suffisamment radicales, et dans la paralysie qui s'ensuit.

Si la France devait apprendre un jour que le plus grand profit de la Révolution doit être de permettre à MM. Dupont ou Durand de désigner les gardes champêtres au lieu et place du pouvoir exécutif, et de mijoter ainsi leur réélection, elle se jetterait peut-être de nouveau dans les bras de la dictature.

Or, nous ne voulons de dictature d'aucune sorte. Nous ne parvenons pas à comprendre en quoi la liberté est intéressée à ce que la République n'ait pas, elle aussi, en gouvernement fort et dans des conditions au dehors. Il faut savoir aller en Egypte à l'occasion, quand on ne veut pas voir l'étranger contester, par egoïsme un peu national, les droits des nationaux.

CHAMBRE DES DÉPUTÉS

Après épuisement de l'ordre du jour sur la convention avec la Compagnie du Nord, M. Tirard, ministre des finances, a déposé le budget extraordinaire pour 1884.

MM. Pierre Legrand, Strepel, Pilchon et Des Rotours présentent plusieurs amendements relatifs à des questions d'intérêt local.

M. Girard propose un amendement relatif aux pensions pour les employés blessés, les veuves et les enfants des employés tués dans l'exercice de leurs fonctions.

Le rapporteur Rouvier fait observer que la Chambre a déjà repoussé un amendement semblable.

Le ministre fera ce qui est possible pour garantir l'intérêt des employés; mais il est impossible d'insérer des dispositions spéciales dans les conventions.

L'adoption de M. Girard est repoussée par 222 voix contre 172.

M. Graux discute ensuite son amendement ayant pour but de réduire la retenue des employés pour la retraite combinée dans les autres compagnies.

M. Tirard, ministre des finances, réplique que le système des pensions de retraite à la Compagnie du Nord diffère de celui des autres compagnies. Les employés sortis de cette Compagnie conserveront leurs droits. Néanmoins le ministre s'efforcera de donner satisfaction à M. Graux.

M. Graux retire son amendement.

MM. Trépoignant et Douville-Maillefeu discutent des questions de tarifs.

Après les explications de M. le ministre, la convention avec la Compagnie du Nord est votée.

L'ordre du jour appelle ensuite la convention avec la Compagnie du Midi.

M. Faure présente un amendement d'intérêt local qui est repoussé.

M. Achard présente un amendement ayant pour but de réduire de 12 à 10 millions la somme stipulée par la convention pour garantie de dividendes. Il dit que cette garantie est donnée contre exploitation du canal du Midi.

M. le ministre entre dans des explications

le plus considérable, le plus riche, et, chose plus rare, le plus estimé de la contrée.

En effet, maire de la ville et membre du conseil général depuis de longues années, il avait fait un bien immense dans le pays; il l'avait doté de ses frais d'une infinité de choses trop utiles et trop coûteuses pour que le gouvernement voulût les accorder; il avait régulé ses impôts, mis de l'ordre dans les petites finances, pris en main les intérêts de chacun; enfin il avait su conquérir la sympathie de tout le monde, et on lui pouvait à cette heure supposer par une douleur qui n'était pas feinte.

La nuit s'avancait, une froide nuit d'hiver où l'on entendait la pluie battre les murs et le vent siffler dans les arbres.

Dans une grande chambre de l'hôtel de Sauvetat, la famille, le docteur, les vieux domestiques étaient réunis autour du lit du mourant.

Afin que la lumière ne gênât point le malade, une seule lampe éclairait l'immeuble et rendait plus triste et plus lugubre cette scène de deuil.

M. de Sauvetat, affaissé sur ses oreillers, la figure horriblement contractée par les affres de l'agonie, avait les yeux fermés. Son sursaut était pénible, tout agité de soubresauts convulsifs. Sa respiration inégale soulevait les couvertures.

De loin en loin une plainte déchirante s'échappait de ses lèvres.

Sa femme, madame de Sauvetat, assise dans un grand fauteuil au chevet du lit, suivait des yeux les progrès rapides du mal et l'affaiblissement visible du mourant.

Elle pleurait abondamment, et de temps à autre, par quelque exclamation désespérée, elle manifestait son malheur ou suppliait le docteur de sauver son mari.

Au pied du lit, une autre jeune femme, qu'on appelait simplement Martine, était debout, aussi pâle que celui dont on attendait le dernier soupir.

Plus blanche qu'un marbre, aussi immobile et en apparence aussi impassible que lui, elle ne gemait pas, elle ne pleurait pas, mais ses traits d'une beauté admirable, altérés par un désespoir indicible, en faisaient si bien l'image de la suprême douleur ici-bas, qu'en la voyant on se sentait remué jusqu'au fond de l'âme.

« N'y a-t-il donc plus d'espoir ? s'écria tout à coup madame de Sauvetat. O doctor! docteur! sauvez-le, et que Dieu prenne ma vie, s'il le faut, en échange de la sienne. »

— M. de Sauvetat toucha à sa dernière heure, Madame, dit M. Delorme en essayant une large rebelle : il n'y a plus rien à tenter. »

12,500

Figure II-2 : Une du Petit Comtois, (1883-08-01).

2.1.3. Techniques d'aide à l'OCR

Les défauts des fichiers image que nous venons d'évoquer, provenant très minoritairement de la phase de numérisation proprement dite, mais surtout de la qualité du support papier, sont à déplorer dans la quasi-totalité des numéros de 1883 à 1903. Les résultats des premiers tests d'océrisation ont constitué un verdict sans appel en confirmant l'extrême faiblesse du taux de reconnaissance optique des caractères des fichiers image du quotidien.

Si l'on ne peut intervenir sur les supports-papier, il reste qu'il est aujourd'hui possible de procéder à une « restauration », selon la terminologie consacrée, des fichiers image par le biais d'un logiciel de retouche et de traitement d'image comme *Bookrestorer*, créé par la société I2S. Parmi ses multiples fonctions, ce logiciel permet d'opérer un recadrage, d'effacer les marques de doigts, de corriger l'éclairage, de redresser les lignes et d'atténuer l'effet de la courbure. Dans un premier mouvement, nous avons tout d'abord choisi de procéder à une binarisation des fichiers images. Cette fonction de *Bookrestorer* permet d'optimiser la qualité du seuillage du document, en convertissant en noir et blanc les documents captés en niveaux de gris. À cet effet, deux paramètres sont manipulés : celui de filtrage, ayant pour fonction d'éliminer certains détails parasites, et celui de profondeur, agissant sur l'épaisseur des contours des caractères. Le bruit sur le document est ainsi considérablement réduit, ce qui augmente par suite le taux de reconnaissance du logiciel de reconnaissance optique des caractères. L'image II-6 est assez représentative du bruitage sur les feuillets traités : elle comporte de nombreux caractères brisés ou effacés, qu'un faible contraste ne fait qu'accentuer. De plus, son fond non-uniforme imputable à la qualité médiocre du papier s'accompagne d'une transparence du *recto* sur le *verso*.



Figure II-6 : Page avant binarisation.



Figure II-7 : Après binarisation.

La figure II-7 témoigne de l'amélioration apportée par le module de binarisation : les caractères sont plus contrastés ; les caractères brisés, quant à eux, ont été reconstruits grâce au paramètre de profondeur. Enfin, les taches ont été « nettoyées » tandis que le verso n'est plus apparent. Cette fonction, bien que fournissant des résultats impressionnants, présente des limites que nous avons observées au fil de nos expérimentations. La binarisation ne permet tout d'abord pas de pallier les défauts de contraste d'images acquises en faible résolution. Par ailleurs, cette opération doit être ajustée en fonction du support sous peine de perdre une quantité importante d'information (disparition des caractères fins ou peu contrastés, points sur les « i ») ou au contraire de faire

apparaître tous les défauts du support papier (fissures du papier, bavures sur les contours): ces deux extrêmes entraînant bien évidemment une altération du taux de reconnaissance. Ainsi, les feuillets du *Petit Comtois* étant dans des états de conservation variables d'un numéro à l'autre, l'intervention manuelle d'un opérateur de restauration reste nécessaire.

Dans un deuxième temps, nos efforts se sont concentrés vers une solution d'atténuation des courbures résultant du mode de stockage des supports papiers. Grâce à un principe de correction géométrique, il est possible de redresser ces lignes courbes dans *Bookrestorer*. Pour évaluer leurs déformations, le logiciel détecte tous les éléments horizontaux et verticaux présents dans l'image ainsi que les lignes de texte, et calcule le degré d'inclinaison de la page. Cette opération se révèle essentielle en vue du module d'océrisation, sans laquelle plusieurs mots par ligne ne seraient pas reconnus par le logiciel. Les images suivantes qui représentent une page intérieure du *Petit Comtois* illustrent parfaitement les apports du logiciel. Sur la figure II-8, on constate très clairement à la droite de la page des courbures atteignant un degré d'inclinaison tel que certains caractères sont déformés et ne peuvent donc plus être identifiés par le logiciel.

trentaine de personnes ; à l'arrière est une petite cabine fort spacieuse, où trois personnes peuvent vivre très à l'aise.

Ce petit chef-d'œuvre se nomme *Diana*, et nous avons ouï dire qu'il appartenait à deux Parisiens en train de faire un voyage excursionniste jusqu'à Marseille.

Hier, à dix heures du matin, on a transporté à l'hôpital, et de là à Belleuax, une femme âgée d'environ trente-cinq ans. Cette femme était enceinte et était tombée dans la Grande-Rue, en proie à une attaque d'éclampsie. Ces sortes de convulsions se produisent assez souvent chez les femmes sur le point d'accoucher.

Caisse d'épargne de Besançon.

Du 24 au 31 octobre 1883.

Reçu de 447 déposants, dont 99 nouveaux, la somme de 117,437 fr.

Remboursé à 281 déposants la somme de 98,949 fr. 00 c.

Le directeur de service était M. Dubourg.

Montbéliard. — Le *Quatorze Juillet* consacre au vote de confiance un article que nous avons lu avec attention. L'auteur de l'article est d'avis qu'il fallait attendre le dépôt de la demande de crédits qui *aurait fait connaître la carte à payer.*

L'auteur de cet article réduit à une question d'argent le point de savoir si on doit tenir ou renoncer à l'exécution des traités de 1874 et de 1883, et si la France doit avoir une politique coloniale dans l'extrême Orient.

Si la question doit être réduite à ces termes, nous admettons l'abstention de certains députés du Doubs ; mais alors nous nous étonnons que ceux qui font du Tonkin une question uniquement de doit et avoir soient si prompts à voter 25 millions

pour la rétribution des conseillers municipaux.

Haute-Saône

La rentrée du tribunal de Lure a eu lieu samedi 3 novembre comme pour les autres tribunaux. On a beaucoup remarqué la présence à la solennité d'une escorte de gendarmerie. La rentrée a été précédée d'une messe du Saint-Esprit comme à Paris.

THÉÂTRE DE BESANÇON

Aujourd'hui lundi

Représentation de M^{lle} AGAR

LES MÈRES ENNEMIES

On se rappelle l'unanimité de la presse à constater le succès des *Mères ennemies*, le beau drame de M. Catulle Mendès, lors de son apparition ; ce succès ne s'est pas démenti à la reprise qui vient d'en être faite, et M^{lle} AGAR a retrouvé les mêmes applaudissements qui avaient salué l'admirable création d'Elisabeth Boleska. Nous sommes heureux d'apprendre à nos lecteurs que les *Mères ennemies* vont être représentées sur notre théâtre avec une interprétation presque semblable à la représentation d'origine : le rôle principal sera joué par M^{lle} Agar, la créatrice ; les autres rôles, soit par les artistes du même théâtre où a été jouée la pièce, soit par des artistes appartenant aux principales scènes de Paris. Les costumes sont ceux de la création. Nous ne doutons pas que le public ne veuille assister à une aussi intéressante représentation.

M^{lle} AGAR jouera mardi, 6, à Lons-le-Saunier.

Le Gérant : A. LOMONT.

BESANÇON, IMPRIMERIE DE MILLOT FRÈRES.

Figure II-8 : Page avant atténuation des courbures.

Après restauration, le phénomène de courbure, s'il n'est pas totalement éradiqué, apparaît considérablement atténué, comme on peut le constater en observant la figure II-9 :

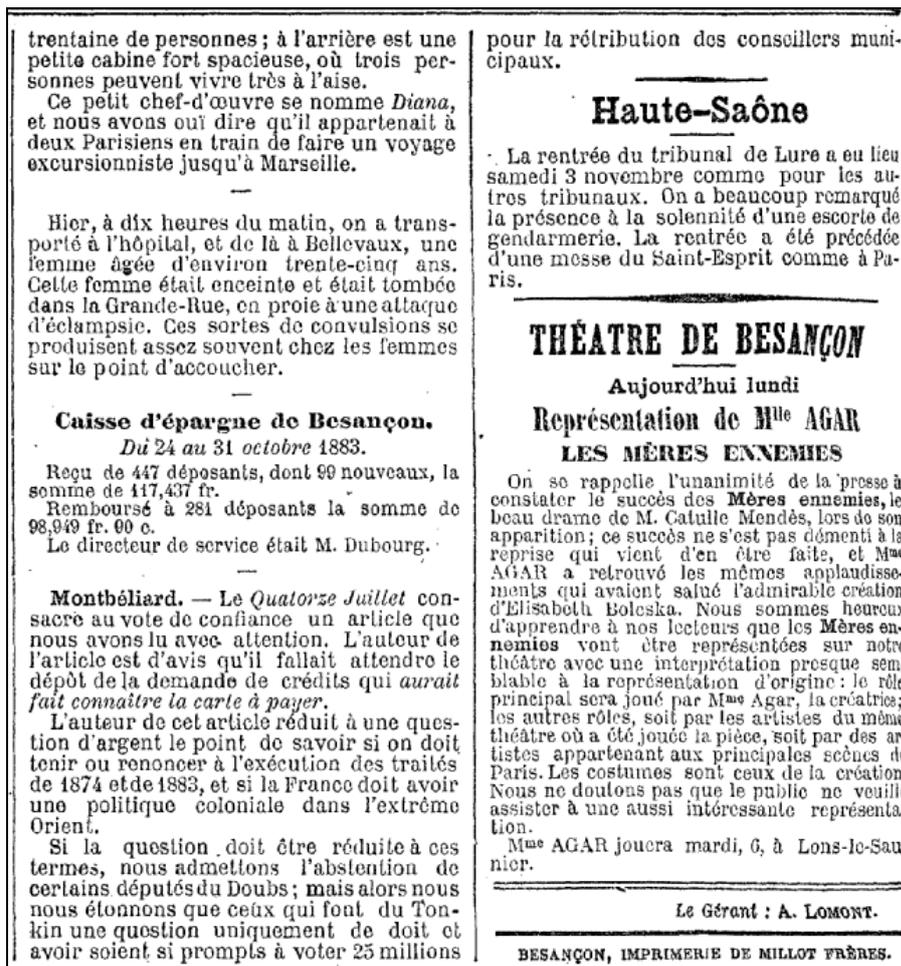


Figure II-9 : Page après atténuation des courbures.

2.1.4. Océrisation et correction automatique, semi-automatique et manuelle

Une fois restaurés, les fichiers ont été océrisés à l'aide d'un logiciel disposant d'un moteur de reconnaissance parmi les plus puissants actuellement commercialisés : *ABBYY FineReader*. Celui-ci, grâce à des bibliothèques de formes, détecte automatiquement les caractères présents dans un document, de sorte à convertir le document image en document texte.

La performance d'un moteur de reconnaissance, si grande soit-elle, ne saurait cependant garantir à elle seule un taux acceptable de reconnaissance des

caractères. Encore une fois, les choix opérés antérieurement (résolution, mode de stockage, prétraitement de l'image), déterminent tout autant la qualité de l'opération d'océrisation que la performance du logiciel de reconnaissance.

En aval de l'acquisition des données, le paramétrage du logiciel d'OCR en fonction des caractéristiques du support à traiter est capital quant à la réussite de l'opération. Les choix techniques qui ont déterminé l'étape d'océrisation des fichiers du *Petit Comtois* ont été les suivants :

2.1.4.1. Ouverture et traitement des images

Une fois l'image ouverte dans *ABBYY FineReader 8.0*, deux solutions se présentent à l'utilisateur:

- Le programme peut automatiquement détecter les zones de texte et d'image. Cette analyse est ensuite utilisée pour déterminer les zones à reconnaître et l'ordre dans lequel s'effectuera la procédure de reconnaissance optique des caractères.
- L'utilisateur peut tracer manuellement les zones de textes à reconnaître.

Nos tests d'océrisation menés à partir d'une analyse automatique du document ont montré que cette fonctionnalité n'était pas adaptée à la mise en page de la presse ancienne. Dans son parcours automatique de la page du *Petit Comtois* prise pour unité de surface de lecture, le logiciel ne distingue pas la rupture entre le haut de la page et son rez-de-chaussée, ce qui a pour effet d'intégrer le texte du feuillet occupant le rez-de-chaussée avec les autres articles présents dans le haut de la page. De surcroît, dans le cas où la colonne est en partie effacée, le programme *ABBYY FineReader 8.0* mène une lecture des caractères sur l'axe horizontal. Pour ces raisons, nous avons opté pour le traçage manuel des zones de texte à lire par le programme, qui permettait de contrôler la lecture du logiciel.

2.1.4.2. Enregistrement du texte reconnu

Guidée par le souci de respecter la matérialité du texte, nous avons procédé aux choix suivants :

- Les données textuelles ont été enregistrées sous mode RTF (*Rich Text Format*) dans le but de conserver les variations typographiques du document (italique, gras, etc.).
- Le programme a été paramétré de sorte qu'il reproduise les polices utilisées mais ne restitue pas les objets selon leur emplacement d'origine. Le texte reconnu est ainsi aligné à gauche.
- En ce qui concerne les paramètres de texte, nous avons choisi de conserver les sauts de ligne et de paragraphes.

2.1.4.3. Options de reconnaissance optique des caractères

Le paramétrage du moteur de reconnaissance d'*ABBYY FineReader 8.0* consiste à renseigner le logiciel sur la langue du document à traiter et sur son type d'impression d'origine. Le programme propose d'adapter sa lecture à seulement deux catégories principales d'impression : le texte tapé à la machine et le texte imprimé sur imprimante matricielle. Ces propositions ne sont bien entendu pas adaptées à la presse ancienne : c'est là une autre preuve de la nécessité d'outils d'océrisation développés par et pour la communauté scientifique travaillant sur des documents anciens. Par défaut, nous avons donc opté pour le premier paramétrage.

2.1.4.4. Qualité de l'océrisation effectuée

Avec ces paramètres, un taux d'erreur conséquent reste à déplorer en bout de chaîne de numérisation. La reconnaissance automatique des caractères est entravée par la diversité des formes qui représentent un même caractère, résultant d'un faisceau convergent de facteurs, parmi lesquels les dégradations dues au temps et à l'acidité du papier, le bruit des images

binarisées, mais aussi les méthodes manuelles de composition typographique perdurant jusqu'à la fin du XIX^e siècle et le taux d'irrégularité de l'impression qui les caractérise.

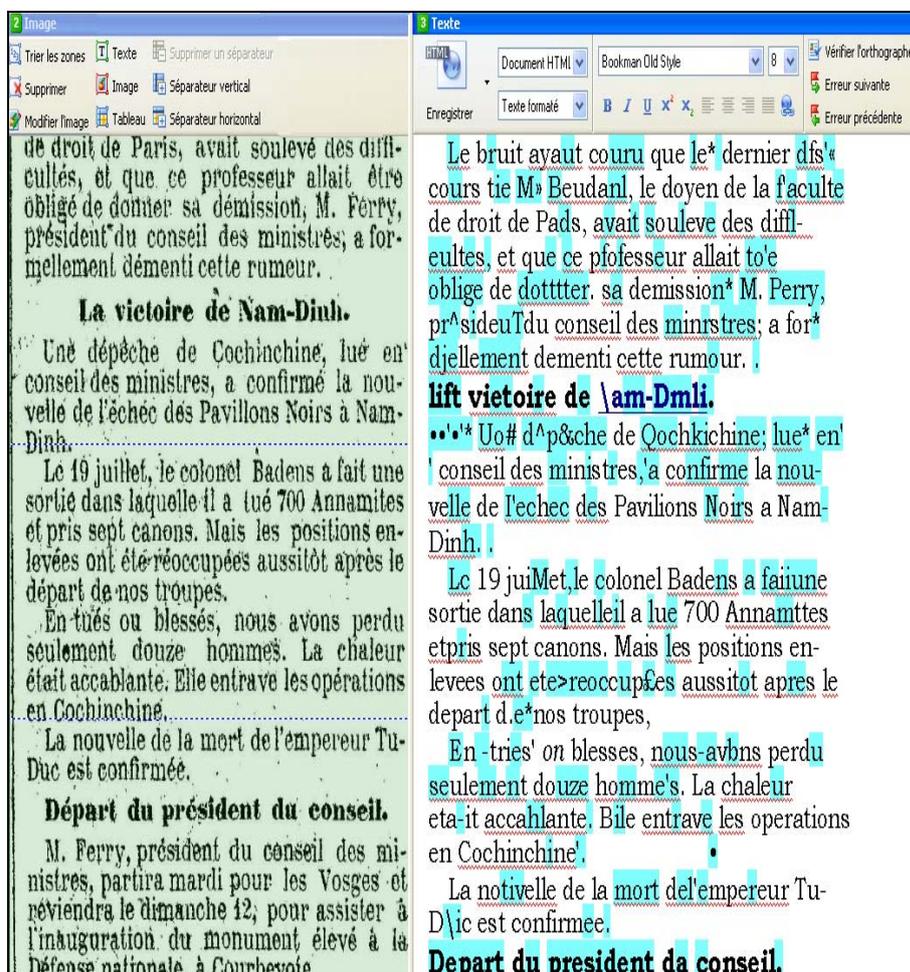


Figure II-10 : Capture d'écran du programme de reconnaissance optique.

Pour illustrer notre propos, nous nous appuyerons sur la description de la capture d'écran du programme de reconnaissance optique reproduite ci-dessus. La partie gauche de l'interface montre la zone image parcourue par le logiciel. Dans la partie droite de l'interface, figure le texte reconnu par le logiciel. Le logiciel met en surbrillance les caractères n'ayant pas fait l'objet d'une reconnaissance certaine par le programme. On notera néanmoins que les « doutes » exprimés par le logiciel ne correspondent pas systématiquement à

une erreur. Inversement, certains caractères non surlignés, et de fait, présentés comme « sûrs », sont pourtant mal identifiés. La qualité du texte issu de l'océrisation est sensiblement identique à celle de l'image, à savoir médiocre. Dans la partie image de l'interface, on aura remarqué des taches, localisées tant à proximité des caractères qu'entre les lignes et les blancs séparateurs, dont l'espacement est d'ailleurs fluctuant. Ces micro-taches sont analysées comme de la ponctuation.

De façon plus problématique, les caractères atrophiés et les ligatures donnent lieu à une reconnaissance erronée. On remarquera ainsi sur la figure II-10, le cas de la suite de deux « l » dans le mot « juillet », identifiés comme un « M ». Dans le tableau suivant, listant les erreurs les plus fréquemment rencontrées, les suites de caractères « de », « ns », « ir » et « ll » attestent la difficulté du programme à restituer ces séquences :

Caractère à reconnaître	Caractère reconnu		Caractère à reconnaître	Caractère reconnu
o	c		t	l
o	e		t	'
n	h		J	l
s	5		ê	ô
f	l		ns	ris
u	v		ir	tr
n	t		de	cle
ll	11		m	ra
l	1		m	rn
d	cl		rc	n

Tableau II-1 : Erreurs de reconnaissance les plus fréquentes.

On remarquera que certains caractères donnent tout particulièrement lieu à des erreurs de reconnaissance. Si sur la figure II-10, le caractère « o » ne fait pas l'objet d'erreurs, on observera que conformément au tableau II-1, la lettre « n »

est mal identifiée. Ainsi, dans le déterminant « Une » sous le titre en gras « La victoire de Nam-Dinh » , la forme arrondie du « n » est analysée comme un « o », et la tache surmontant le « e », fait conclure au logiciel qu'il se trouve en face d'un dièse. Autre problème de lecture très gênant : sur 12 accents présents sur l'image, aucun n'est reconnu correctement par le programme.

2.1.4.5. Vérification et correction orthographique

Pour améliorer le taux de reconnaissance, *ABBYY FineReader* offre une phase de transcription assistée par ordinateur, ou phase d'apprentissage. Cette phase consiste à enrichir la bibliothèque des caractères du logiciel. Dans ce mode de lecture assistée, le logiciel rencontrant une forme inconnue va interroger l'opérateur sur le caractère à associer à cette forme, qui l'identifiera manuellement dans une fenêtre de dialogue. Ainsi, comme très fréquemment dans *Le Petit Comtois*, la courbure du caractère « é » apparaît par exemple atrophiée sur la figure II-11, qui présente une boîte de dialogue de vérification orthographique. Le logiciel assimile initialement cette forme à un « ô » : l'accent aigu est confondu avec un accent circonflexe et la forme circulaire du caractère non marquée par le trait central distinguant le « e » du « o » incitent le logiciel à identifier un « ô » comme l'indique la fenêtre de dialogue suivante :



Figure II-11 : Exemple de boîte de dialogue (1).

Enregistré dans la bibliothèque de formes, ce caractère sera désormais identifié selon nos attentes.

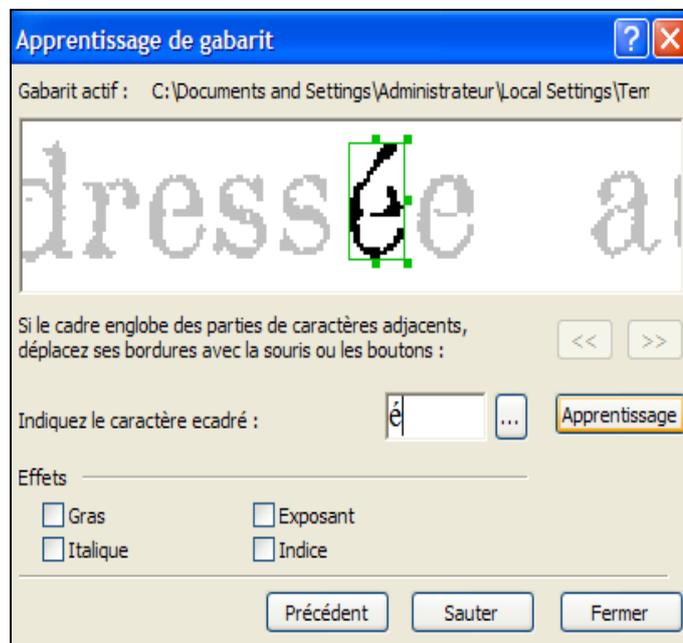


Figure II-12 : Exemple de boîte de dialogue (2).

Afin de limiter le taux d'erreur, nous avons donc construit un gabarit d'apprentissage rassemblant plus de 260 formes à partir de la lecture contrôlée d'une dizaine de numéros du *Petit Comtois*. L'éditeur de gabarit ainsi construit regroupe donc une dizaine de formes pour un caractère donné, classé selon la typographie du caractère (gras, italique, casse, accentuation) :

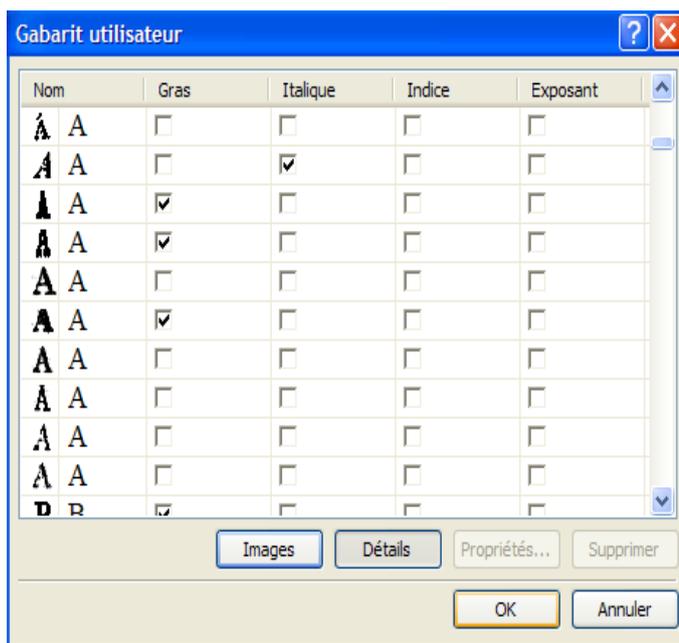


Figure II-13 : Extrait du gabarit utilisateur.

Grâce à la mise en place de ce gabarit utilisateur, les résultats de l'OCR ont été légèrement améliorés, pour un temps de mise en place cependant relativement conséquent et des perspectives limitées. Durant la période sur laquelle se déploie la collection du *Petit Comtois*, de nombreux changements de typographie, de papier, frappés par diverses dégradations (liquide, graisse, etc.), représentent autant de variations d'un même caractère. Or, un gabarit ne peut prétendre à l'exhaustivité sans devenir une lecture totalement assistée, caractère par caractère. Développer la bibliothèque de formes du logiciel de reconnaissance ne constitue donc pas la voie la plus prometteuse d'une optimisation du taux de reconnaissance des caractères.

2.1.4.6. La correction manuelle

À défaut d'outils disponibles permettant de corriger semi-automatiquement les erreurs rencontrées, nous avons consacré la majeure partie du temps de constitution de notre corpus à la correction manuelle du plein-texte. Armée de patience, nous avons ainsi effectué un travail de *scribe-copiste* (Viprey, 2006b : 176) évalué à plus de 1200 heures. Certaines fonctionnalités disponibles dans tous les logiciels de traitement de texte, comme la création de nouvelles entrées au dictionnaire, ont bien entendu été mobilisées. Néanmoins, ces solutions peu ergonomiques ne sont pas suffisantes, tant en raison de leur faible rendement que de leur manque de scientificité, pour la mise en œuvre de la constitution d'une base textuelle de très grand volume telle que le sera, à terme, la base textuelle de presse quotidienne régionale de la Troisième République, archivant des dizaines de titres de presse.

2.1.5. Documenter les données

Pour être exploitable, toute donnée textuelle doit s'accompagner de données la qualifiant. La pertinence et la viabilité d'une base textuelle repose impérativement sur la description de chacune des qualités des unités documentaires qui la composent. En effet, ces données descriptives conditionnent en amont la sélection des données pertinentes en rapport avec le projet de recherche d'un utilisateur-expert, et, en aval, l'exploitation même des données textuelles. Convoquées comme critères de partitionnement, ces données descriptives intégrées à chaque texte fournissent les prérequis d'analyses comparatives entre textes. En lexicométrie, ces données

référentielles renseignant sur les qualités diverses d'un texte comme son auteur, son genre, sa date de parution, etc., sont nommées *métadonnées*² :

Les métadonnées (*meta-data*) représentent la documentation des ressources utilisées. Elles ont deux objectifs complémentaires. En premier lieu éviter les ressources « orphelines », dont on ne sait pas ce qu'elles rassemblent, avec quels choix et quel travail éditorial ; et dont finalement on ignore ce qu'elles signifient exactement (ce sont des données rendues muettes). [...] Le deuxième objectif est de permettre à quelqu'un de savoir qu'une ressource existe et de connaître ses caractéristiques avec suffisamment de précision pour pouvoir la retenir ou non en fonction de ses objectifs propres (Habert, 2005 : 126-127).

Pour identifier chaque unité documentaire, nous avons instauré un ensemble d'attributs discursifs renseignant sur :

- La **source** des données textuelles considérées, recevant pour valeur le titre du quotidien ;
- La **date** des données textuelles, dont la valeur est la date de parution de l'article, rédigée au format aaaa/mm/jj ;
- Le **numéro** des données textuelles, dont la valeur précise le numéro de parution du journal duquel sont extraites les données ;
- Leur **auteur** ;
- La **rubrique**, recevant pour attribut l'intitulé de la rubrique qui héberge les données ;
- Ainsi que le **titre**, dont l'attribut renvoie au péritexte de l'article.

Nous avons jugé ces premières indications insuffisantes pour restituer les caractéristiques pertinentes des articles de presse. Partant du postulat que l'aire scripturale (Peytard, 1975) du journal est un support signifiant, nous avons établi des lignes de renseignements relatives à l'emplacement de l'article dans le quotidien. Le choix de la page où est ancré un article, sa surface, son emplacement dans la page, comptent parmi les ressorts de la mise en valeur

² Remarquons ici que les métadonnées, axées sur le discursif, telles qu'entendues ici ne sont pas développées en vue d'applications documentaires comme les notices catalographiques classiques.

de l'information et représentent potentiellement des éléments pertinents pour l'utilisateur expert de la base. En vue de décrire l'emplacement topographique de l'article, nous avons modélisé la page de journal en la segmentant en colonnes, puis attribué un code à chacune d'elles : « 1 » pour la première colonne, « 2 » pour la deuxième colonne, « 3 » pour la troisième colonne, etc. Pour permettre une localisation de l'article, il a fallu développer un système de coordonnées. Nous avons divisé la page du journal selon son axe vertical en trois zones : le haut de la page (« H »), le milieu de la page (« M ») et enfin, le bas de page (« B ») correspondant au rez-de-chaussée. La figure II-14 donne une représentation de cette segmentation.

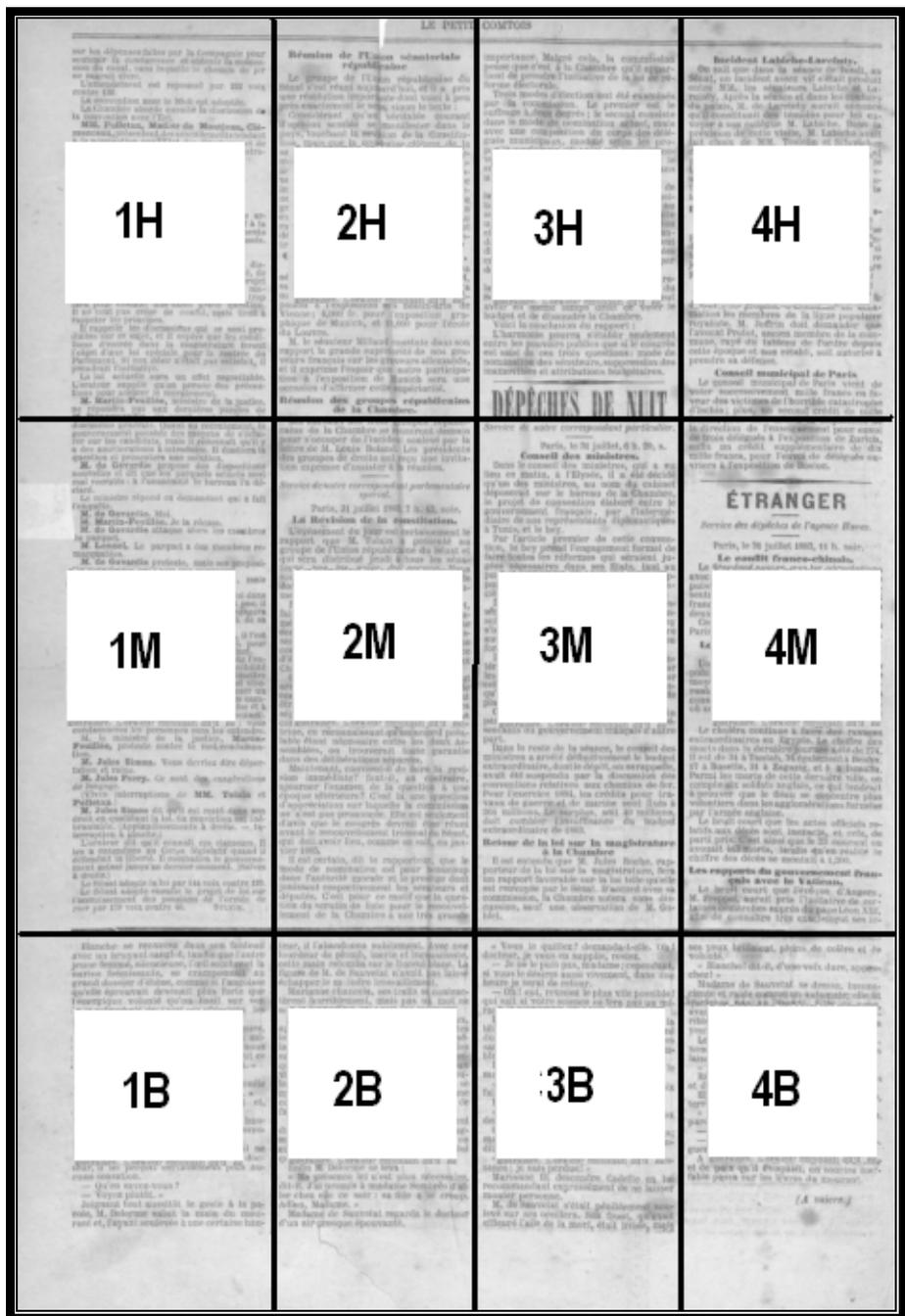


Figure II-14 : Code topographique.

Enfin, l'espace occupé par chaque article est parallèlement évalué par une quantification en nombre de colonnes couvertes. Ces indications donnent ainsi lieu à deux attributs supplémentaires de métadonnées, à savoir :

- L'**emplacement** de l'article, dont l'attribut attendu est le code de ses coordonnées
- La **largeur**, visant à informer de la surface occupée par l'article en termes de colonnes

Chaque article, envisagé comme une unité documentaire, est ainsi systématiquement précédé des métadonnées suivantes :

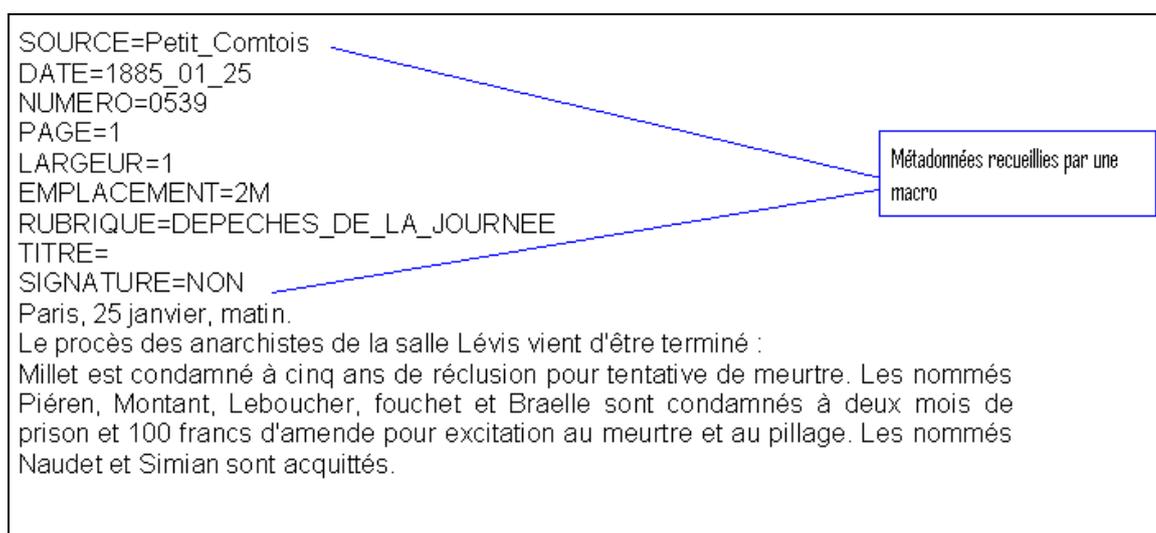


Figure II-15 : Exemple de métadonnées.

Sur le plan technique, ces métadonnées ont été recueillies par macro-instruction³, méthode choisie pour obtenir un gain de temps et de normalisation de chaque bloc de métadonnées. Dans un premier temps, la saisie des variables a souffert d'un manque de normalisation (variations graphiques non pertinentes et parfois confusions plus graves) qui a pu compromettre partiellement le recours aux métadonnées dans notre travail.

³ Cette macro a été construite par Nicole Salzard, ingénieure informatique au LASELDI, que je remercie chaleureusement pour son aide précieuse.

CHRONIQUE_REGIONALE- HAUTE_SAONE
CHRONIQUE_REGIONALE- HAUTE- SAONE
CHRONIQUE_REGIONALE- HAUTE -SAONE
CHRONIQUE_REGIONALE- Haute_Saône
CHRONIQUE_REGIONALE- Haute- Saône
CHRONIQUE_REGIONALE_HAUTE_SAONE
CHRONIQUE_REGIONALE- HAUTE- SÂONE
CHRONIQUE_REGIONALE- HAUTE- SAONE.
CHRONIQUE_RÉGIONALE- HAUTE-SÂONE

Figure II-16 : Exemple des variations de saisie dans le champ des rubriques pour un même intitulé.

Ces expérimentations ont montré la nécessité de remettre à plat le processus d'acquisition des données. La deuxième campagne de numérisation du *Petit Comtois*, à la différence de la première, a donc été conçue de sorte (1) que la définition et la description des unités documentaires soient confiées à un opérateur dont les qualifications étaient celles d'un ingénieur d'étude (2) que la procédure de saisie des informations descriptives soit semi-automatisée (contrôle par une liste déroulante).

2.1.6. Établissement et normalisation des données textuelles : vers une philologie numérique

L'établissement scientifique des données textuelles et leur normalisation sont les deux versants d'un chantier qui n'a effectivement débuté qu'en 2008 et qui se trouve aujourd'hui encore être en construction. Cette mise en œuvre tardive de cet empan de nos recherches se justifie tout d'abord par le contexte technologique très évolutif auquel est soumise la constitution des données du *Petit Comtois* : comme nous l'avons souligné précédemment (chapitre I), le développement de la pratique de normalisation des données textuelles, s'étant imposé assez récemment, a tardé à s'accompagner d'une littérature technique traitant pas à pas des étapes à adopter pour normaliser un corpus. C'est lors d'un stage en août 2008 au centre ATO de l'UQAM que nous nous sommes

familiarisée avec la philosophie XML-TEI, dont nous évoquerons à présent les points forts, avant d'exposer le protocole de normalisation des données élaboré lors de notre travail collaboratif avec F. Daoust au centre ATO.

2.1.6.1. XML-TEI et philologie numérique

Le XML-TEI marque un tournant dans la praxis des sciences du texte, dans la manière de construire leur objet tout d'abord, mais aussi de penser leur action sur le texte. Ce sont ces points sur lesquels nous voudrions insister à présent.

Documenter les données pour faciliter leur échange et matérialiser leur statut d'objet construit

En premier lieu, les recommandations TEI impliquent d'intégrer à tout document des informations descriptives sur son contexte de production, une fiche bibliographique. Ces informations descriptives, susceptibles de devenir un critère de sélection de corpus ou sous-corpus, sont présentées dans l'entête TEI du document, signalée par la balise <TEIHeader>. Selon les recommandations minimales de la TEI, l'entête précisant les références bibliographiques du document doit renseigner sur son titre, sur l'autorité responsable de sa création⁴ (Université ; Laboratoire ; chercheur ; etc.), sur les décisions d'encodage adoptées (présentation de la signification des balises, niveau de finesse de l'établissement, etc.) ainsi que sur l'histoire des traitements, stipulant, s'il y a lieu, les différentes révisions du document. Tout commentaire sur les choix de codage et l'historique de révision est précieux pour une méthodologie éminemment itérative comme l'analyse de données textuelles, où le codage est très fréquemment modifié en cours d'analyse. La fiche signalétique peut se compléter en fonction des options de son responsable : on

⁴ Les éléments contenus dans <titleStmt> renseignent ainsi sur le nom de l'institution assurant le financement du projet (<funder>), le nom du chercheur principal responsable de la création d'un texte (<principal>).

pourra notamment donner de plus amples renseignements sur les choix éditoriaux (soit par un paragraphe présentant l'intégralité des choix, ou de préférence, par renvoi à un ouvrage méthodologique), ou encore préciser les termes de la disponibilité du document, la langue employée, la taille du document électronique stocké, ou préciser toute information venant s'ajouter à celles non traitées dans les balises préexistantes. Il s'agit d'un bouleversement de la *praxis* du texte sans précédent depuis l'imprimerie. Ces recommandations sont l'occasion de mettre en lumière et de matérialiser le fait que :

[...] les textes ne sont pas des données, mais des constructions problématiques issues de diverses procédures (Rastier, 2001 : 82).

En intégrant les choix du responsable de la construction du corpus et en proposant de renseigner sur l'historique de construction de ce dispositif, les recommandations TEI matérialisent l'acte de *textualisation* opéré par le chercheur sur les données.

Rendre compte de la complexité du texte

L'effort d'établissement du texte se prolonge par l'emploi des possibilités d'annotations pour rendre compte de la matérialité textuelle. Le codage XML-TEI peut ainsi être invoqué pour expliciter la structure formelle d'un texte. Or, les informations sur la structure du texte, de la division en pages jusqu'à la division en paragraphes, de la segmentation en chapitres jusqu'à la segmentation en mots, constituent autant d'éléments participant à la construction du sens. Mais le XML-TEI se met également au service de la philologie numérique en lui donnant les moyens techniques de tisser ensemble les différents états d'un texte, qu'il s'agisse de variations *auctoriales* ou de variations *éditoriales* (Adam, 2005b : 72). Ce faisant, le XML-TEI fournit au chercheur une solution de représentation du texte apte à figurer l'instabilité de ce dernier.

Partager les annotations du texte

Alors que l'analyse des données textuelles prend non seulement pour objet la surface graphique du texte, mais aussi le texte lemmatisé ou étiqueté, il

est désormais impératif de pouvoir partager les différents codages opérés au cours d'une étude. Nous envisageons les différents types d'étiquetage d'un texte opérés au fil de l'analyse comme des états du texte à part entière, matérialisant pour leur part l'acte de lecture. Qu'il s'agisse d'un étiquetage morphosyntaxique (Kastberg, 2006), phonétique ou encore sémantique (Véronis, 2004b), tout codage du texte est une valeur ajoutée au texte, qui, à ce titre, justifie d'être capitalisée. Ces niveaux d'analyse témoignent d'un accès renouvelé à la pluralité des dimensions du texte qui n'est plus aujourd'hui incompatible avec le respect de sa matérialité « première » ou seconde. Le XML-TEI permet la mise en commun de ces enrichissements du texte, établis au prix de longues heures de travail, mais aussi leur confrontation, condition *sine qua non* de la scientificité des Sciences des textes. Les possibilités d'annotation offertes par le XML-TEI nous intéressent tout particulièrement en cela qu'elles permettent d'insérer, dans un même document, des éléments co(n)textuels. Or, cette annotation, qu'on pourrait qualifier de référentielle ou d'encyclopédique, est un besoin impérieux pour une communauté scientifique attelée à l'étude de textes dont la valeur historique a souvent pour corollaire un contexte disjoint de celui de l'analyse. Songeons également à toutes les difficultés posées par l'étude des textes sollicitant intensivement l'intertexte et/ou ne construisant explicitement leur système de référents, dont la connaissance est pourtant requise pour la lecture, entendue comme acte de compréhension et d'interprétation. Il nous semble important de remarquer sur ce point, à la suite de J.-M. Viprey (2005), que la philologie numérique renouvelle les conditions de consultation des annotations encyclopédiques, décuplées par les capacités ergonomiques des outils numériques comme l'hypertexte :

Toute intrusion encyclopédique, quelles que soient sa pertinence et/ou sa naïveté, toute insertion philologique, pourtant si nécessaire dans les perspectives de plus en plus profondément diachroniques que comporte la lecture intégrale du texte littéraire français, interviennent comme des déchirures dans le discours-lecture. L'ergonomisation, l'optionnalisation, la hiérarchisation des annotations ne sont pas un luxe, mais un élément essentiel du retour au texte, du confort intellectuel, du contact des œuvres (Viprey, 2005b : 58).

En somme, le codage XML-TEI, outil philologique par excellence, ouvre la voie à une *herméneutique partagée*, en cela que le partage des lectures qu'il autorise a pour pendant une intensification des confrontations de proposition de sens et des méthodologies d'analyses. Si le potentiel heuristique d'un encodage XML-TEI est certain, il convient d'examiner à présent les conditions techniques de sa mise en œuvre et d'évaluer le degré de finesse d'établissement du texte optimal auquel un projet de constitution d'une archive de grande ampleur peut prétendre.

2.1.6.2. Encodage des données du *Petit Comtois* en XML-TEI

En vue d'obtenir les données du *Petit Comtois* en XML, nous avons dans un premier temps converti l'ensemble des fichiers RTF en code HTML, format issu de la norme SGML.

Lors de nos premières expérimentations, il s'est avéré que les codes HTML générés par *Microsoft Word* étaient de qualité très inférieure à ceux produits par *Open Office*, même si ces derniers contenaient encore un certain nombre de balises inutiles. Les codes HTML d'*Open Office* signalent notamment à chaque saut de paragraphe la fonte utilisée, et ce, même si la police ne varie pas. Certaines balises, pourtant potentiellement pertinentes, se révèlent porteuses de renseignements erronés. C'est le constat que nous avons pu faire au sujet des types de polices, qu'il aurait été appréciable de conserver, mais qui constituaient une indication peu fiable dans la mesure où les fontes du XIX^e siècle n'avaient pas été correctement reconnues par le logiciel d'OCR. De la même façon, nous avons pu observer que les changements de taille de fonte signalés dans les codes HTML ne représentaient pas fidèlement les variations observées sur le document d'origine. Il nous semble là encore devoir imputer cette classe d'erreurs au logiciel d'océrisation. La variation de taille de police est un élément participant au sens du texte : elle est fréquemment utilisée dans *Le Petit Comtois* pour signaler des phénomènes de citation comme on peut l'observer sur la figure II-17. Aussi, nous avons pris le parti d'en garder trace.

Ce choix se justifie d'autant plus que l'insertion des variations de taille de police dans le fichier HTML, outre ceci qu'elle peut être automatisée, ne nécessite pas un investissement très important et peut être accomplie très facilement dans une étape de vérification des données. On trouvera en annexe le code HTML de l'unité documentaire présentée en figure II-17.



Figure II-17 : Unité rédactionnelle (1884-01-16), p.2.

La transformation des données HTML en format XML a été opérée via le logiciel accessible en ligne *Sato* (Système d'Analyse de Texte par Ordinateur, Daoust, 1996). Dans le cadre du réseau ATONET, œuvrant pour une mise en commun de ressources et de méthodes dans le domaine de l'analyse de données textuelles assistées par informatique, les développements récents de *Sato* en

font un terrain d'expérimentations pour la capitalisation des données, annotées ou non, et l'interopérabilité entre les différents logiciels d'analyse de données textuelles (*Alceste, Astartex, DTM, Lexico3, Weblex*). Afin de permettre le transfert de données textuelles d'un logiciel à l'autre sans perte de leur enrichissement, *Sato* se présente comme une interface d'analyse où l'utilisateur peut procéder à la conversion de ces données en un *format pivot*, basé sur le XML et les recommandations de la TEI (Daoust & Marcoux, 2006 : 329).

Nous avons ainsi déposé tous les fichiers HTML de la base sur la plateforme *Sato*. Par l'utilisation d'un filtre rédigé en PERL⁵, nous avons ensuite extrait les informations pertinentes, sur la structure du texte, c'est-à-dire les sauts de paragraphes, les saillances typographiques (gras et italique), et les tailles de police, que nous avons préalablement corrigés dans les fichiers HTML manuellement lorsque nécessaire. À la suite du passage de ce filtre, l'unité documentaire reproduite ci-dessus se présente comme suit :

⁵ Ce filtre a été rédigé par F. Daoust, que je remercie très chaleureusement pour son aide extrêmement précieuse.

*size="11pt" *{
ARTICLE :
SOURCE : Petit Comtois
DATE : 1884-01-16
NUMERO : 0168
PAGE : 2
LARGEUR : 1
EMPLACEMENT : 1M
RUBRIQUE : intérieur
TITRE :
SIGNATURE : NON
TEXTE :
}
*page= 1884-01-16_P 2_ 1M
*partie=rubrique intérieur
*partie=titre
*partie=texte

*size="11pt" On sait que l'extrême gauche a délégué cinq de ses membres pour rechercher les propositions intéressant les questions sociales qui pourraient être soumises à la Chambre pour remédier à la crise ouvrière résultant du chômage.

*size="11pt" M. Jules Guesde, a apprécié cette résolution dans le *size="11pt"
style+="ital" _Cri du peuple. *size="11pt" L'article du pontife du collectivisme révolutionnaire est intitulé : *size="11pt" Taisez-vous!* _*style="ital"

*size="11pt" En voici, quelques extraits, que nous livrons aux méditations des intransigeants.

*size="10pt" C'est ça qui va donner du pain à ceux dont le crédit est épuisé comme la bourse, et qui va maintenir au travail ceux qui sont sous le coup d'une mise à pied pour demain ou après.

*size="10pt" Il est vrai que si, au lieu de prendre une dizaine de jours pour aboutir, nos braves intransigeants avaient abouti séance tenante, il n'y aurait rien eu de changé, — l'oreille qu'ils se sont finalement avisés d'ouvrir aux souffrances prolétariennes n'étant qu'une oreille d'âne.

*size="10pt" Oui, ânes — et ânes bâtés — des députés qui osent placer au premier rang des « causes de la crise » les traités de commerce, les tarifs de chemins de fer et l'assiette de l'impôt » — alors que la même crise sévit plus intense peut-être en Angleterre et aux Etats-Unis, où impôts, chemins de fer et commerce se présentent dans des conditions toutes différentes.

*size="10pt" Oui, ânes — et de leur propre aveu — les mêmes députés qui proclament que « l'heure des actes a sonné, mais qu'ils ne savent pas eux-mêmes les mesures qu'on pourrait proposer » et s'invitent mutuellement à « étudier. »

*size="10pt" Le temps est passé, bien passé, où l'on pouvait espérer tromper les estomacs ouvriers avec des fables ou des interpellations vides. Notre prolétariat a appris, lui ; il a appris, et la cause unique du mal, et le remède définitif, et le remède provisoire. Et il ne pardonnera pas plus à ceux de ses prétendus représentants qui n'ont qu'une phraséologie impuissante à mettre à son service qu'à ceux qui le laissent décimer sans ouvrir la bouche.

*size="11pt" Anes bâtés ! le qualificatif doit sembler dur aux malheureux intransigeants qui commencent à éprouver à leurs dépens la vérité du proverbe :

*size="10pt" Qui sème le vent recueille la tempête.

Les informations descriptives de l'article ont été transformées en commentaire (signalé par la séquence *{...}) et les attributs contenus dans ces métadonnées ont été édités en propriété. Les attributs des valeurs DATE, PAGE, EMBLEMMENT ont dans le même temps été combinés de façon à créer un identifiant singulier pour chaque unité documentaire, comme requis par le logiciel *Sato*.

Une fois déposés sous cette forme sur *Sato*, les fichiers de la base peuvent être exportés en format pivot XML-TEI. Conçu pour permettre de repérer et de conserver les annotations du corpus en vue de les rendre réutilisables d'un logiciel d'analyse de données textuelles à l'autre, ce format repose sur un découpage du texte en *token* ou plus exactement en « séquence discursive continue et non-vide » (Daoust & Marcoux, 2006 : 335) par opposition à un découpage en unités linguistiques (*word*). Cette segmentation préalable est matérialisée par la balise `<w>...</w>`, ayant pour attribut XML la séquence *XML : id*, à laquelle succède la référence de l'occurrence, qui sera l'identifiant singulier de cette dernière. La structure formelle du document est pour sa part traduite à l'aide des balises `<pb/>`, `<p/>` et `<lb/>`, indiquant respectivement le début d'une nouvelle page, le début d'un paragraphe et le début d'une nouvelle ligne. L'attribut *n* de la balise indique le numéro de la ligne. À titre d'illustration, voici un extrait du texte considéré plus haut en format pivot XML-TEI :

```

<pb n="1884-01-16_P/1"/>

<p><!--*{
ARTICLE :
SOURCE : Petit Comtois
DATE : 1884-01-16
NUMERO : 0168
PAGE : 2
LARGEUR : 1
EMPLACEMENT : 1M
RUBRIQUE : intérieur
TITRE :
SIGNATURE : NON
TEXTE :
}
--><lb n="1"/><milestone unit="partie" n="nil"/><milestone unit="size"
n="11pt"/><milestone unit="style" n="nil"/><w xml:id="w2">2_</w> <milestone
unit="size" n="11pt"/><w xml:id="w3">1m</w>
<lb n="2"/><milestone unit="partie" n="rubrique"/><milestone unit="size"
n="11pt"/><w xml:id="w5">intérieur</w> </p>
<p><lb n="3"/><milestone unit="partie" n="texte"/><milestone unit="size"
n="11pt"/><w xml:id="w7">On</w> <milestone unit="size" n="11pt"/><w
xml:id="w8">sait</w> <milestone unit="size" n="11pt"/><w xml:id="w9">que</w>
<milestone unit="size" n="11pt"/><w xml:id="w10">l'</w><milestone unit="size"
n="11pt"/><w xml:id="w11">extrême</w> <milestone unit="size" n="11pt"/><w
xml:id="w12">gauche</w> <milestone unit="size" n="11pt"/><w xml:id="w13">a</w>
<milestone unit="size" n="11pt"/><w xml:id="w14">délégué</w> <milestone unit="size"
n="11pt"/><w xml:id="w15">cing</w> <milestone unit="size" n="11pt"/><w
xml:id="w16">de</w> <milestone unit="size" n="11pt"/><w xml:id="w17">ses</w>
<milestone unit="size" n="11pt"/><w xml:id="w18">membres</w> <milestone unit="size"
n="11pt"/><w xml:id="w19">pour</w> <milestone unit="size" n="11pt"/><w
xml:id="w20">rechercher</w>

```

Figure II-18 : Extrait d'une unité documentaire en format *XML-TEI Sato*.

Enfin, conformément aux recommandations TEI, nos données normalisées sont introduites par un en-tête minimal, dont les blocs principaux ont été générés par *Sato* puis complétés manuellement. Cet en-tête se présente comme suit :



Figure II-19 : En-tête TEI minimal.

Comme montré ci-dessus, l'en-tête TEI des données du *Petit Comtois* est composé, par ordre de progression, d'un prologue renseignant sur le format du document, d'un élément racine renvoyant aux recommandations TEI, d'un premier en-tête englobant un espace de description bibliographique du document et d'un espace d'informations sur le texte et la source dont il dérive. Ce dernier est composé d'un bloc *refDecl* documentant la façon dont les références canoniques ont été établies, et d'un bloc *ClassDecl*, que nous avons créé pour renseigner sur toutes les codifications des emplacements et des rubriques utilisés dans le corpus. Par souci de gain d'espace, seuls les

premiers exemples des listes EMLACEMENT et RUBRIQUE ont été présentés sur la figure ci-dessus. Enfin, l'en-tête s'achève sur une balise fermante de type `</TEIHeader>`.

2.1.7. Bilan et perspectives

Outre l'acquisition du premier fonds de la base PQR3R, soit approximativement six millions de mots, les premières expérimentations sur *Le Petit Comtois* (1883-1903) ont permis de construire un protocole amélioré de numérisation mis en œuvre pour l'acquisition de grands volumes de données plein-texte structurées, portatives et documentées issues de presse ancienne.

La presse ancienne nécessite le développement d'outils de correction linguistique

La presse ancienne résiste principalement à l'acquisition de ses données sous mode plein texte en raison de la détérioration de son support d'origine. Les solutions existantes de restauration des fichiers image sont efficaces pour redresser les lignes courbées par la conservation des supports papiers en volume relié, pour atténuer les phénomènes de transparence du verso sur le recto ainsi que les empreintes, mais ne permettent pas de remédier à la perte de caractères ou d'empans de texte effacés. Le taux de reconnaissance des caractères, affecté d'une façon générale par les défauts du document, est de surcroît affaibli par les fontes typographiques utilisées et leurs espacements fluctuants, dont la lecture est souvent ambiguë pour le logiciel d'océrisation. Développé pour la lecture de données textuelles fixées sur un document de bonne qualité et dont le vocabulaire est plus restreint, le logiciel *ABBYY FineReader*, pourtant référencé comme l'un des meilleurs programmes industriels disponibles sur le marché, s'est révélé insuffisant face à la numérisation du *Petit Comtois*. Pour l'ensemble de la première campagne de numérisation (1883-1903) effectuée de 2005 à 2007, nous évaluons en effet le taux de reconnaissance du logiciel à quelque 60%. Dans sa version 9.0,

disponible depuis 2009, le moteur de recherche du même logiciel s'est révélé beaucoup plus puissant puisque son taux de reconnaissance avoisine les 70%. Malgré cette amélioration fort appréciable, la qualité du texte reconnu ne saurait suffire aux besoins des sciences des textes, en mal d'outils performants. Cette performance nous semble certes reposer sur l'accroissement de « l'intelligence artificielle » des programmes d'OCR, mobilisée pour accroître leur capacité de mémorisation de nouveaux caractères rencontrés fréquemment. L'augmentation de la qualité des données reconnues est en effet dépendante d'un enrichissement linguistique des dictionnaires convoqués par ces programmes. Mais selon nous, l'acquisition à grande échelle de données rendant compte de la matérialité du texte, implique surtout de disposer d'un outil qui articule (1) des phases de lecture automatique où chaque forme jusqu'ici non répertoriée et intégrée dans le dictionnaire serait signalée à l'utilisateur expert ; (2) à des phases de dialogue avec l'utilisateur expert qui aurait la charge du contrôle et de la correction des erreurs de lecture, nécessairement réalisées en contexte sur la base de concordances. Il s'agit donc bien pour la communauté scientifique de se doter d'une plate-forme fonctionnelle d'océrisation et de correction semi-automatique des données. Sans la mise en œuvre de procédés d'amélioration d'assistance de correction, intégrée ou non dans l'interface d'océrisation, la faisabilité d'un projet de numérisation de presse ancienne en mode plein-texte reste compromise par la nature chronophage d'une correction manuelle particulièrement laborieuse. D'après nos estimations, cette dernière représente près de 120h⁶ par million de mots, amputant d'autant d'heures le temps alloué à l'établissement fin des données. Dans le cadre du pôle *Archive, Bases, Corpus*, J.-M. Viprey, le concepteur d'*Astartex-Diatag* travaille actuellement au développement d'un environnement de correction ergonomisée.

⁶ Sur la base d'environ une heure de correction par numéro, comptant en moyenne 9000 mots.

S'affranchir des formats propriétaires dès l'acquisition des données en mode plein-texte

Une base de données textuelles implique nécessairement que les documents électroniques qu'elle est vouée à communiquer soient pérennes et exploitables par le plus grand nombre sans que leur consultation requière de disposer de la suite bureautique avec laquelle le document a été rédigé. En d'autres termes, pour être lisibles et portables, les données textuelles créées doivent obligatoirement être affranchies des formats propriétaires. Alors que la première campagne de numérisation du *Petit Comtois* opérait une conversion des données RTF en XML-TEI nécessitant maintes opérations envisagées comme autant de surcoûts, la deuxième campagne de numérisation du quotidien s'attache à enregistrer directement les données issues de l'océrisation en HTML.

Structurer les données

Pour que les données textuelles de la base puissent être totalement réutilisables, il est impérieux qu'elles soient structurées de façon normalisée. Dans ce cadre, il s'agit tout d'abord de fournir à l'utilisateur une série de renseignements contextuels sur l'article, qui pourront être convoqués comme autant de critères pour la sélection d'un corpus (source, date de parution, rubrique, signature). Dans l'objectif d'accroître la normalisation des données et dans un souci de gain de temps et d'économie de moyen, nous avons donc repensé l'organisation de la deuxième chaîne de numérisation. À ces objectifs s'ajoutait, dans la perspective de mise à disposition d'un environnement de consultation couplant image et texte, la nécessité de conserver trace de la localisation de chaque article dans une zone de pixels définie et répertoriée, afin de la mettre en relation hypertexte avec les données plein-texte correspondantes. Toutes ces visées ont convergé vers la décision de procéder à la segmentation des unités documentaires en amont de l'interface d'océrisation. En effet, dans notre premier protocole de numérisation, l'opérateur chargé de l'océrisation segmentait la zone de texte que le

programme devait lire, colonne après colonne, de façon à conserver l'ordre du discours tel qu'il se déployait dans la mise en page originale du support-papier : sans ce découpage, le logiciel procédait à une lecture horizontale du texte. L'opérateur chargé de la correction avait ensuite la charge de définir chaque unité documentaire en lui intégrant ses métadonnées. Ce protocole engendrait, d'une part, des surcoûts importants et ne tenait, d'autre part, pas suffisamment compte des niveaux de compétences induits par le balisage contextuel des données, très nettement supérieurs à ceux exigés par l'opération de correction et *a fortiori* par l'océrisation.

Aussi, le protocole de la chaîne de numérisation du *Petit Comtois* (1883-1927) se fixe de procéder à la segmentation des unités documentaires dans un logiciel de traitement d'images (à ce jour TM*Photoshop*) en amont de l'océrisation, afin d'automatiser la passe de reconnaissance optique des caractères et de conserver les coordonnées en pixels ou points PPI de l'unité documentaire dans la surface de la page du numéro, permettant ultérieurement la mise en relation hypertexte du format image et du mode plein texte. Pour éviter que le logiciel d'océrisation ne procède à une lecture horizontale d'un article se déployant sur plusieurs colonnes, l'opérateur construit alors plusieurs fichiers images, répertoriés comme composantes d'une même unité documentaire. Une fois lues par le logiciel d'océrisation comme des entités indépendantes, les données seront regroupées pour former une même unité documentaire à partir des indications fournies dans un fichier répertoire des métadonnées. Cette étape de segmentation des unités documentaires a été couplée avec l'opération de saisie des métadonnées. Cette dernière a été semi-automatisée par l'intégration d'un menu déroulant, ayant pour but de limiter les erreurs de saisie des intitulés de rubrique. Une fois l'unité documentaire segmentée, la fenêtre de dialogue reproduite ci-dessous apparaît, guidant l'opérateur dans la saisie des champs à renseigner.

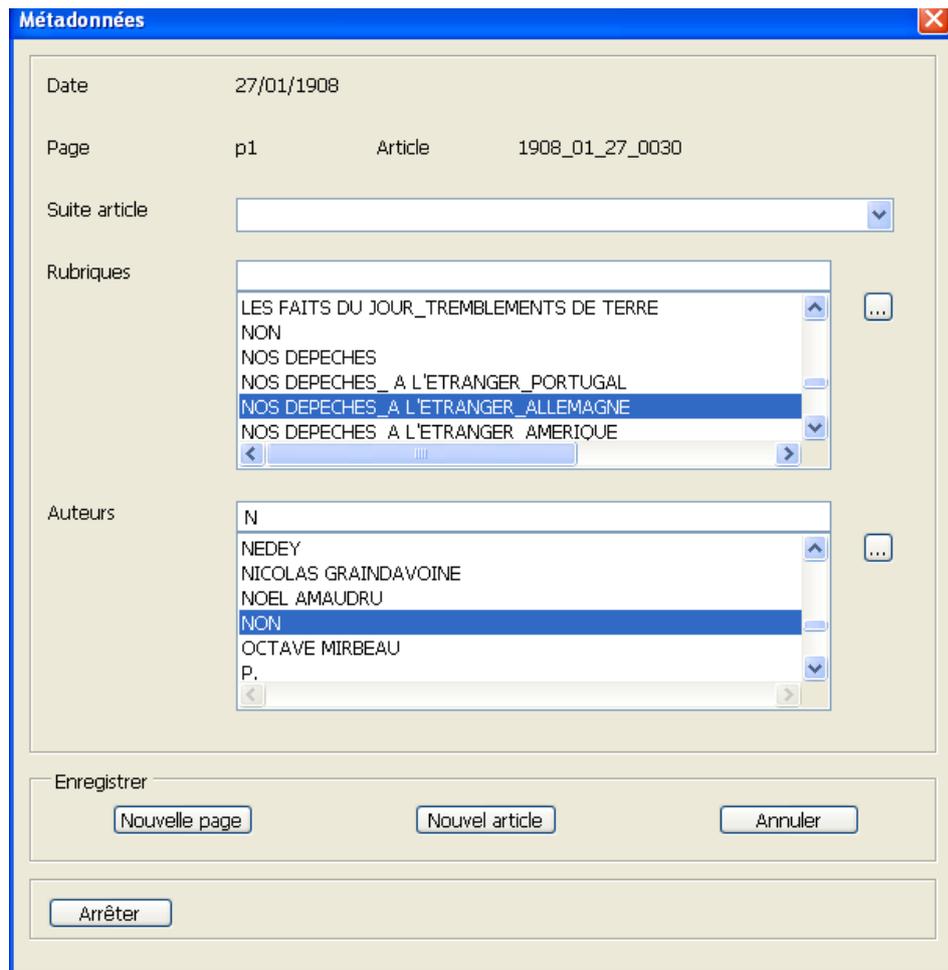


Figure II-20 : Saisie des métadonnées dans *TMPhotoshop*.

Les métadonnées ainsi saisies, une fois validées, sont enregistrées dans un fichier de format .txt dont nous avons reproduit un extrait ci-dessous. Chaque ligne de ce fichier répertorie les métadonnées d'une unité documentaire et renseigne sur la date de publication de l'article, la page du journal où il se localise, la taille du fichier image de la page, les coordonnées de l'unité documentaire sur cette page. La seconde partie de la ligne renseigne sur la rubrique d'appartenance de l'unité documentaire, éventuellement sa sous-rubrique, et son auteur.

Date	Page					
1906_01_31;	1/1;	0001/0002;	5171 px_6978 px/5171 px_6978 px;	123_1279_956_6615/951_1263_1760_2688;	OPINIONS;	G.D
1906_01_31;	1;	0003;	5171 px_6978 px;	959_3087_1756_3747;	LES FAITS DU JOUR_CONSEIL DES MINISTRES;	NON
1906_01_31;	1;	0004;	5171 px_6978 px;	959_3746_1756_4159;	LES FAITS DU JOUR_CONSEIL DES MINISTRES;	NON
1906_01_31;	1;	0005;	5171 px_6978 px;	955_4158_1752_4479;	LES FAITS DU JOUR_CONSEIL DES MINISTRES;	NON
1906_01_31;	1;	0006;	5171 px_6978 px;	955_4618_1748_4927;	LES FAITS DU JOUR_NOUVELLES PARLEMENTAIRES;	NON
1906_01_31;	1;	0007;	5171 px_6978 px;	951_4918_1756_5547;	LES FAITS DU JOUR_NOUVELLES PARLEMENTAIRES;	NON
1906_01_31;	1;	0008;	5171 px_6978 px;	955_5530_1752_5831;	LES FAITS DU JOUR_NOUVELLES PARLEMENTAIRES;	NON
1906_01_31;	1;	0009;	5171 px_6978 px;	959_5826_1752_6323;	LES FAITS DU JOUR_NOUVELLES PARLEMENTAIRES;	NON
1906_01_31;	1;	0010;	5171 px_6978 px;	967_6322_1752_6627;	LES FAITS DU JOUR_NOUVELLES PARLEMENTAIRES;	NON

Numéro de l'image	Taille de la page	Coordonnées image de l'unité documentaire	Rubrique	Sous-rubrique	Signature
-------------------	-------------------	---	----------	---------------	-----------

Figure II-21 : Exemple de métadonnées.

Garante d'une plus grande homogénéité, la semi-automatisation de la saisie des métadonnées permet aussi l'automatisation de l'océrisation ; elle représente donc un important gain de temps, non négligeable. Les métadonnées localisées dans le fichier image seront ultérieurement intégrées dans l'entête de chaque unité documentaire corrigée, conformément aux recommandations de la TEI.

Quelle richesse d'annotations programmer pour la base?

À cette étape d'avancement de nos travaux, nous n'énoncerons que les principes du balisage des données de la base de PQR3R sans en fournir un descriptif détaillé.

À cela, une première raison : un balisage fin du texte tel que l'exige l'analyse textuelle du discours, parce qu'il ne peut se faire automatiquement, représente une opération très coûteuse en temps et en main-d'œuvre, et se trouve dès lors dépendant des financements accordés au projet. Remarquons que si l'on peut aisément envisager d'opérer un balisage très fin d'un corpus de quelques milliers de mots, il en est tout autrement pour sa mise en œuvre sur un ensemble de 200 000 millions de mots, qui sera approximativement, à terme, le volume de la base PQR3R. Le volume des données à traiter est concrètement

incompatible avec un balisage manuel. Or, l'annotation des données d'une base se doit d'être systématique et cohérente, tant au niveau qualitatif que quantitatif. Il s'agit en effet d'offrir à l'utilisateur une annotation stabilisée et ce, dans l'intégralité des données de la base. Par conséquent, le balisage des fonds d'une base de quelques millions de mots ne peut reposer que sur une annotation semi-automatique. Un des chantiers les plus stimulants de la construction de bases de données enrichies nous semble ainsi relatif à la classe d'annotation linguistique. Comme nous l'avons récemment proposé (Viprey & Lethier, 2008), la lemmatisation serait ainsi un balisage porteur d'une véritable valeur ajoutée : le contrôle des passes automatiques par un chercheur garantirait la finesse des annotations et le respect du principe d'économie inhérent à la base. Un étiquetage minimal des phénomènes d'hétérogénéité signalée du discours à partir d'un repérage semi-automatique des marques formelles, serait un autre exemple de chantier à mener pour accentuer la valeur ajoutée de la base.

La deuxième raison pour laquelle le degré de finesse de balisage des données ne sera pas traité plus avant ici peut être érigée en principe : toute annotation des données doit être rationalisée en fonction des attentes des utilisateurs experts de la base. Avant d'aller plus loin dans l'établissement des données, il convient de prendre connaissance des besoins des recherches interdisciplinaires suscitées par la base, qui alimenteront en retour, nous l'espérons, la richesse d'annotations de la base.

2.2. De la base textuelle au corpus

En analyse du discours, un ensemble de données textuelles historicisées ne devient *corpus* qu'en fonction d'un projet d'analyse singulier pour lequel il a été constitué et de la problématique qui le sous-tend. La mise en série de textes résulte d'un ensemble de choix heuristiques qui déterminent en retour le sens de chacun des textes ainsi contextualisés, comme l'a pointé F. Rastier

(2001 : 92). Aussi, il convient à présent d'insister sur les choix qui ont présidé à la configuration du corpus *Petit Comtois*, avant d'en décrire les propriétés.

2.2.1. Retour sur notre projet de recherche

Initialement, notre projet de recherche s'assignait pour programme d'observer l'évolution de la production discursive du *Petit Comtois* sur la totalité de sa parution, c'est-à-dire de 1883 à 1944. Cette ambition a très vite été contrariée par le caractère chronophage de l'acquisition des données textuelles nécessaires, – ne s'accommodant que mal avec le contexte actuel de la recherche universitaire –, voire par les limites intrinsèques de toute recherche en linguistique de corpus⁷.

Notre recherche s'est ainsi définitivement tournée vers l'étude des productions discursives du *Petit Comtois* des débuts de sa parution à 1903. Ce corpus est comme un dispositif d'observation de la stratégie discursive d'un quotidien de sensibilité républicaine d'une part, et des mutations des formes d'écriture journalistique d'autre part. Si la première borne de cette tranche diachronique est naturelle, puisqu'elle correspond à la première année de parution du *Petit Comtois*, la borne de fin, 1903, nécessite d'être justifiée. Risquons-nous tout d'abord à dire l'évident : de 1883 à 1903, se jouent les 21 premières années de parution du *Petit Comtois*, organe républicain exerçant une influence des plus notables sur l'éducation politique du « petit peuple » de Franche-Comté. Ces années s'inscrivent dans un contexte socio-historique des plus intéressants: celui des débuts de la Troisième République ainsi que celui de la « naissance du journalisme moderne » pour reprendre le titre de l'ouvrage de M. Palmer (1983).

⁷ Nous ne pouvons donc qu'abonder dans le sens de B. Pincemin (1999) et approuver son constat selon lequel : « Le rapport aux données tient d'un compromis : faire avec ce à quoi on a accès, mais faire au mieux avec cela » (1999 : 415).

2.2.1.1. Bref historique des débuts de la Troisième République (1883-1903)

Pour montrer en quoi le discours d'un quotidien de presse de province de la fin du XIX^e siècle tel que le *Petit Comtois* constitue un terrain d'observation de grand intérêt, nous procéderons à un bref historique de la Troisième République. Il ne pourrait être question, dans les limites de cette recherche, d'exposer ici les origines de la Troisième République, ni même ces premiers jours depuis sa proclamation le 4 septembre 1870. Notre propos, beaucoup plus modeste, se limite à fournir quelques repères dans l'histoire de la Troisième République pendant la période couverte par notre corpus, à même d'éclairer l'évolution socio-politique du *Petit Comtois* de 1883 à 1903.

L'orientation pragmatique des débuts de la Troisième République (1880-1885)

Aux nombreuses expériences constitutionnelles initiées pendant près de trois quarts de siècle, s'ensuit au début des années 1880 une période de stabilité institutionnelle. Pour faire aboutir les principes de la Révolution, les républicains qu'on nommait « opportunistes »⁸ sont contraints de ménager le monde rural et la bourgeoisie d'affaires, en se présentant et en agissant non plus comme des révolutionnaires, mais comme des réformistes. La presse, et tout particulièrement la presse régionale, assure dans ce contexte un rôle crucial : il s'agit de « trouver le ton qui convient à ses abonnés, essentiellement ruraux » (Grévy, 1998 : 318) pour s'assurer du soutien politique du lectorat.

⁸ « [...] ceux qui, comme le nom l'indique, allaient être prudents et donner une image de la République ferme dans ses réalisations qu'elle considère comme essentielles ; mais prudente dans tout ce qui n'est pas essentiel, ou qu'elle ne considère pas comme essentielle ; modérée dans le domaine social, pour essayer de rallier le plus possible de Français ordinaires, de gens de classe moyenne, voire de notables et libéraux et conservateurs » (Agulhon, 1991 : 13).

L'avènement des républicains au pouvoir marque la mise en place d'une politique de gouvernement ayant pour objectif d' « introduire le progrès »⁹, selon la formule comtienne, sans pour autant mettre en péril l'ordre. Ne visant pas tant à refonder brutalement la société qu'à mettre en place les conditions de réalisation des principes de la Révolution, les républicains se donnent pour priorité de faire émerger des réformes susceptibles d'être acceptées par l'opinion publique. Parce que, dans une logique positiviste, elle est la pièce maîtresse d'une société *allant vers le progrès*, l'école est rendue gratuite, obligatoire et laïque de 6 à 13 ans par les lois de 1881 et 1882. Le bilan des premières années de la Troisième République ne saurait cependant être établi sur la seule base des lois votées jusqu'en 1885, qui sont d'ailleurs essentiellement libérales (lois sur les libertés de réunion publique votées le 30 juin 1881, sur la liberté de la presse du 29 juillet 1881). Pour dresser un état des lieux de l'œuvre de la politique opportuniste, il est nécessaire, comme le rappelle F. Ewald (1992), de prendre en compte le très grand nombre de projets et de propositions de lois concernant les questions sociales. Ces projets visent notamment la réglementation du travail (réduction de la journée de travail, suppression du livret ouvrier, etc.), le rapport salarial dans la grande industrie (dispositions concernant les mines et les chemins de fer) mais aussi la liberté syndicale, les questions de secours en cas de maladie, de retraite ou d'accidents du travail. Bien que ces propositions de lois n'aboutissent pas avant 1885, – à l'exception de la loi du 21 mars 1884 sur les syndicats professionnels –, elles témoignent néanmoins d'une volonté de traiter les questions sociales, fondamentalement plurielles, comme l'avait posé Gambetta dans son discours du Havre du 18 avril 1872¹⁰. Selon l'historien F. Ewald, les

⁹ Voir à ce sujet Berstein (1992 : 167-169).

¹⁰ « Croyez qu'il n'y a pas de remède social, car il n'y a pas une question sociale. Il y a une série de problèmes à résoudre, de difficultés à vaincre, variant avec les lieux, les climats, les habitudes, l'état sanitaire, les problèmes économiques qui changent à l'intérieur d'un même pays. Eh bien, les problèmes doivent être résolus un par un et non par une formule unique ».

opportunistes sont alors dans l'incapacité d'instaurer un traitement politique positif de la question sociale puisque la mise en œuvre de celui-ci impliquerait de s'inscrire en rupture avec le style des utopies révolutionnaires :

Le problème qu'ils ont à résoudre est en quelque sorte d'instituer, avec la République, un nouveau contrat social qui, tout en accomplissant ou en terminant la Révolution – ce qui reste le grand programme politique du XIX^{ème} –, se ferait contre les doctrines philosophiques et juridiques qui ont inspiré la Révolution. (Ewald, 1992 : 176).

La question coloniale va mettre en difficulté de nombreux ministères de la République opportuniste, car elle divise les républicains : si les opportunistes lui sont globalement favorables, les radicaux y sont fermement opposés, pour des raisons parfois idéologiques¹¹, mais surtout parce qu'ils jugent que les aventures colonialistes détournent l'attention des provinces perdues d'Alsace et de Lorraine. Suite à des mouvements anticolonialistes, le ministère Ferry (février 1883- mars 1885) – déjà tombé en novembre 1881 suite à la position de la France en Tunisie –, est contraint de démissionner en raison de la question du Tonkin, tout comme le ministère Brisson l'est en décembre 1885, seulement huit mois après sa mise en place.

1886-1898 : Scandales et mutations des forces politiques

Si la politique coloniale de la France donne lieu à un climat de tension, ce dernier va être aggravé par une série de scandales (1886-1898). Dans un contexte économique difficile, aggravé par le déficit financier de l'Etat, cette série d'affaires concourt à l'émergence de mouvements de contestation

¹¹ Rappelons sur ce point que J. Ferry considérait que le deuxième ordre d'idées justifiant sa politique coloniale était « le côté humanitaire et civilisateur de la colonisation » (Discours à la Chambre du 28 juillet 1885), à quoi G. Clémenceau répond le 30 juillet de la même année : « Non, il n'y a pas de droit des nations dites supérieures contre les nations inférieures. Il y a la lutte pour la vie qui est une nécessité fatale, qu'à mesure que nous nous élevons dans la civilisation nous devons contenir dans les limites de la justice et du droit. Mais n'essayons pas de revêtir la violence du nom hypocrite de civilisation. Ne parlons pas de droit, de devoir. La conquête que vous préconisez, c'est l'abus pur et simple de la force que donne la civilisation scientifique sur les civilisations rudimentaires pour s'approprier l'homme, le torturer, en extraire toute la force qui est en lui au profit du prétendu civilisateur. Ce n'est pas le droit, c'en est la négation. Parler à ce propos de civilisation, c'est joindre à la violence l'hypocrisie ».

fragilisant la Troisième République, qui marquent l'ouverture d'une nouvelle période s'accompagnant d'une mutation des forces politiques selon Jean-Marie Mayeur :

Une période nouvelle s'ouvre, marquée par une instabilité ministérielle accrue, par la montée des oppositions ennemies du régime. La République libérale et parlementaire, en ces années de difficultés économiques et sociales, déçoit les masses urbaines. Elles se tournent vers des idéologies qui répondent mieux à leurs aspirations : le nationalisme qui change de visage à la faveur de la crise boulangiste, le socialisme qui devient une force politique et paraît menacer l'ordre social (Mayeur, 1984 : 115).

La crise boulangiste marque le début de cette série d'épisodes ébranlant la stabilité de la République. Imposé comme ministre de la guerre du cabinet Freycinet (janvier 1886-décembre 1887) par Georges Clémenceau dont il était un proche, le général Georges Boulanger bénéficie très vite d'une grande popularité. Le corps militaire l'apprécie pour ses réformes, comme l'autorisation du port de la barbe ; les français le respectent pour avoir refusé d'envoyer des troupes contre les mineurs lors de la grève de Decazeville (1886) et supprimé le tirage au sort de la conscription. L'hostilité qu'il manifeste envers l'Allemagne lors de l'affaire Schnaebelé (1887) lui vaut de devenir une figure du revanchisme ; d'où son surnom de général « Revanche ». Craignant que les actes de provocations répétés de G. Boulanger envers l'Allemagne ne déclenchent une nouvelle guerre, le gouvernement Rouvier (mai 1887-décembre 1887) succédant au cabinet Goblet (décembre 1886-mai 1887) nomme un nouveau ministre de la guerre, le général Féron.

Le mouvement boulangiste prend alors son essor, la figure du Général fédéralise les représentants de la droite monarchiste, les bonapartistes, les radicaux, les nationalistes de Paul Déroulède et de la Ligue des Patriotes. Fort du soutien de ces déçus du nouveau régime, revanchards et opposants de la République, le général Boulanger remporte une série de victoires aux élections législatives partielles en 1888 avant d'être élu à Paris en janvier 1889. Ses partisans l'encouragent alors à prendre le pouvoir par la force. Accusé de complot contre la sécurité intérieure, le Général s'enfuit à Bruxelles en avril

1889. Cette crise se traduit par une prise de mesures visant à défavoriser les mouvements populistes, telles que l'abandon du scrutin de liste au profit du scrutin d'arrondissement (13 avril 1889) et l'interdiction des candidatures multiples (17 juillet 1891).

La Troisième République est également déstabilisée par une suite d'affaires politiques et financières. En octobre 1887, la presse met à jour un trafic de décorations par le député Daniel Wilson, gendre du président de la République Jules Grévy. L'enquête révèle que le député a vendu depuis un bureau de l'Élysée des milliers de décorations pour verser des subventions à des journaux de province. Relayé par les boulangistes, ainsi que par Jules Ferry et Georges Clémenceau, cette affaire, dite « des décorations », contraint J. Grévy à démissionner en décembre 1887. Le discrédit jeté par cette affaire sur le gouvernement est accentué par le scandale du Panama, mettant en cause un certain nombre de députés corrompus. Ces derniers avaient voté une loi pour l'augmentation du capital de la Compagnie universelle du Canal océanique, pourtant au bord de la faillite en raison de la mauvaise gestion de Ferdinand de Lesseps, promoteur de la société. L'inévitable faillite de la Compagnie en 1889 ruine près de 80 000 souscripteurs¹². Le scandale éclate en 1892 sous la plume du journaliste E. Drumont dans les colonnes de *La Libre Parole*, feuille nationaliste et antisémite, qui s'empresse de mettre en cause certains financiers juifs.

Dans ce climat social tendu, les années 1890-1898 sont marquées par de nombreuses mutations des forces politiques.

Alliant radicaux et républicains modérés, la concentration républicaine remporte les élections de 1889. Néanmoins, les radicaux reprochent au gouvernement de ne pas faire assez de réformes sociales et critiquent sa *politique d'apaisement*. Voulant unifier la France, le cabinet Freycinet (mars 1890-février 1892) agit en

¹² Selon C. Fohlen (1982 : 394), la population bisontine perd près de 5 millions de francs dans cette affaire.

effet en faveur d'un apaisement de la question religieuse, vers laquelle converge l'Eglise, inquiète du renouvellement du paysage politique français. Tentant de remédier à l'image préjudiciable d'une Eglise royaliste, le Pape Léon XIII, en 1890, charge le Cardinal Lavignerie de rallier lors du toast d'Alger les catholiques à la République. Cette initiative étant mal accueillie par les radicaux et les royalistes, Léon XIII publie, le 15 mai 1891, l'encyclique *Rerum Novarum* dans laquelle la politique sociale de l'Eglise est définie.

Tandis que la question sociale est au centre des débats, le socialisme s'impose progressivement comme une force politique à part entière : en 1893, les socialistes effectuent une percée en gagnant 50 sièges. Ces années sont aussi celles du développement du syndicalisme : en 1895, l'on ne compte pas moins de 450 000 syndicats en France, et ce onze ans seulement après leur autorisation. La force revendicatrice des syndicats est perceptible à travers une série intensive de grèves. À cette agitation socialiste, s'ajoute la multiplication des troubles anarchistes. C'est notamment l'attentat à la bombe perpétré par Ravachol le 11 mars 1892 : celui-ci lance une bombe dans l'enceinte de la Chambre des députés présidée par Auguste Vaillant le 9 décembre 1893. Pour réprimer le mouvement anarchiste, le gouvernement Casimir-Périer (décembre 1893-juin 1894) vote une série de trois lois qualifiées de « scélérates » par l'opposition socialiste. En décembre 1893, en l'espace de quelques jours, sont votées deux lois, celles du 12 décembre 1893 et du 18 décembre 1893 : la première autorise les juges à condamner tout acte de provocation indirecte ou d'apologie ; la seconde permet d'inculper tout membre ou sympathisant d'association de malfaiteurs. Suite à l'assassinat du président S. Carnot par un anarchiste italien, le gouvernement Dupuy (juillet 1894-janvier 1895) vote enfin la loi du 28 juillet 1894. Visant directement les anarchistes en les nommant explicitement, cette loi leur interdit tout type de propagande. Deux gouvernements se succèdent alors, dont celui de L. Bourgeois (novembre 1895-février 1896) composé uniquement de radicaux, renversé par le Sénat en raison d'un projet de réforme fiscale comportant un impôt sur le revenu. Une

relative stabilité débute avec la constitution du gouvernement Méline (avril 1896-octobre 1898), marquant la rupture entre les républicains modérés et la gauche. J. Méline instaure alors une politique d'apaisement : sans remettre en cause les lois laïques, il donne des gages à l'Eglise, et accueille en France une partie des congréganistes expulsés au début des années 1880. La France semble alors s'orienter vers une politique conservatrice. Les radicaux, relégués en province par les socialistes, attaquent sévèrement cette politique : forts du soutien électoral du monde rural, ils voient grandir leur influence au Sénat.

Vers une République radicale

Au moment où la France s'apprête à basculer dans le XX^e siècle, l'affaire Dreyfus provoque une redistribution des forces politiques, comme le met en exergue M. Rébérioux (1975).

Si l'arrestation du capitaine Dreyfus, accusé d'espionnage au profit de l'Allemagne et sa condamnation à la déportation à vie sur l'île du Diable en Guyane date de 1894, il faut attendre 1896 pour que deux journaux de grand tirage, *L'Eclair* et *Le Matin*, attirent l'attention de l'opinion sur les preuves qui ont entraîné la condamnation de Dreyfus. En 1897, *Le Figaro* publie une lettre de Mathieu Dreyfus, révélant que l'auteur du bordereau livrant les renseignements à l'armée allemande est en fait le commandant Esterhazy. De nombreux journaux sortent alors du silence, même si la classe politique ne s'empare pas de l'affaire. Un exceptionnel retentissement est donné à l'affaire lorsque *L'Aurore* publie la « Lettre au président de la République » signée E. Zola et coiffée du titre « J'accuse » par G. Clémenceau ; le plus populaire des écrivains français accuse l'armée d'avoir condamné un innocent et de ne pas vouloir le reconnaître. L'affaire divise l'opinion publique. Les dreyfusards insistent sur l'extrême légèreté des preuves, voyant dans l'affaire un complot réactionnaire, clérical et militaire, un coup du *sabre et du goupillon*, pour reprendre l'expression de G. Clémenceau, visant à détruire la République parlementaire. Les antidreyfusards, eux, mettent en avant le caractère incontestable de l'affaire jugée et sont scandalisés que l'autorité de l'armée et de la justice soit

remise en cause. Ce camp regroupe la plus grande partie des officiers et de la magistrature, les royalistes et les nationalistes, mais aussi la plupart des grands quotidiens populaires. Les antidreyfusards sont majoritairement persuadés de la culpabilité de Dreyfus, « ce qu'ils concluaient de sa race », selon la formule de M. Barrès.

L'affaire finit par générer des troubles menaçant l'ordre public. La Ligue de l'Action Française, créée en 1898, s'attache à attiser les passions pour renverser la République des « nantis et des pourris » pour reprendre encore les propos de P. Déroulède. Les actes antisémites se multiplient tandis que les dreyfusards se font molester. C'est dans ce contexte que P. Déroulède, profitant des obsèques du président Faure, tente d'entraîner l'armée dans un putsch contre l'Élysée le 23 février 1899. Le 4 juin de la même année, des antidreyfusards envahissent l'hippodrome d'Auteuil où se trouvait le président de la République E. Loubet. Ce dernier est alors frappé à coup de canne. Bien qu'à première vue anecdotique, cet événement a pourtant une charge symbolique forte et témoigne de l'état de fragilité de la République. Le 11 juin, se déroule une grande manifestation dreyfusarde en réaction aux événements du 4 juin. Elle est très vite réprimée par l'intervention de la police envoyée par C. Dupuy, ce qui révèle la partialité de ce dernier en faveur des nationalistes. Le ministère Dupuy est renversé le 12 juin.

Sur le plan politique, l'affaire Dreyfus aboutit d'une part à la constitution d'une nouvelle majorité parlementaire et à la fin de la politique progressiste. E. Loubet, devant la progression des extrémistes de droite, tente de mettre en place une nouvelle majorité énergique plus à gauche, qui permettrait de défendre efficacement la République contre ses ennemis. Le 22 juin 1899, à l'issue d'une crise ministérielle d'une quinzaine de jours, le Ministère Waldeck-Rousseau est constitué. Pour contrer les poussées nationalistes menaçant la République, le Président du Conseil forme un gouvernement de « *défense républicaine* », avec des radicaux et un socialiste du nom d'Alexandre

Millerand. J. Jaurès, à la tête du groupe parlementaire qui regroupe la plupart des députés socialistes, soutient également la majorité pour défendre le régime. L'affaire Dreyfus participe d'autre part au retour à une politique résolument laïque, à laquelle s'attellent en priorité les radicaux. Le 1^{er} juillet 1901, le parlement adopte la loi sur les associations, soumettant les congrégations religieuses à autorisation administrative. Le Bloc des gauches, regroupant la gauche laïque, remporte les élections d'avril-mai 1902 avec une majorité confortable et initie une politique radicale visant à l'aboutissement de la laïcisation du pays. Le ministère Combes (juin 1902-janvier 1905) procède le 27 juin 1902 à la fermeture de 125 écoles religieuses créées après 1901, pour fermer quelques temps plus tard toutes les écoles ouvertes sans autorisation avant 1901, malgré les protestations de l'épiscopat.

Des débuts d'une Troisième République marquée par une orientation pragmatique à la mise en place d'une République radicale à la suite de l'affaire Dreyfus, la presse républicaine de province est chargée de conquérir et de rallier à l'idée républicaine son lectorat dorénavant électoral déterminant. L'influence de cette presse régionale sur le paysage politique, supérieure à tout autre type de presse, n'est pas le seul trait la distinguant de la presse nationale : la presse de province se différencie essentiellement par son « style » spécialement adapté au lectorat provincial.

C'est précisément cette spécificité discursive que nous nous proposons d'examiner à partir des outils de l'analyse du discours, en nous attachant à observer les marques d'une formation discursive républicaine sur l'investissement de la langue, des formes lexicales et des structures rhétorico-stylistiques en diachronie.

2.2.1.2. Un contexte de mutation de la pratique journalistique

Les caractéristiques de notre corpus ont conjointement été définies dans l'intention d'observer l'état des productions discursives d'un quotidien de presse régionale tandis que *s'invente*, selon le titre de l'ouvrage de Thomas Ferenczi

(1993), la pratique journalistique à la fin du XIX^e siècle. Du côté de l'histoire de la presse et de la sociologie du journalisme, il est acquis que les dernières décennies du XIX^e siècle sont marquées, dans une version forte, par l'avènement du journalisme comme fonction sociale spécifique et *ordre de discours* différent (Chalaby, 1998) ou dans une version faible, par une « profonde mutation des journaux et du journalisme » (Bellanger *et al*, 1972 : 275).

Il convient de rappeler brièvement que le discours de presse des gazettes et des premiers quotidiens de la première moitié du XIX^e siècle est une *praxis* littéraire. En témoigne l'édition de 1851 du *Dictionnaire des professions* d'Édouard Charton, – pourtant patron de presse et fondateur du premier magazine (*Le magazine pittoresque*, 1833) –, qui, au mot « journaliste » renvoie à l'entrée « homme de lettres ». La profession de journaliste ne prend consistance sur le plan sémantique qu'autour de 1870-1880. La rupture avec le champ littéraire et la constitution du champ journalistique comme champ relativement autonome est d'abord sensible à travers l'histoire sociale de la profession de journaliste. Leur nombre ne cesse d'ailleurs de croître au cours du XIX^e siècle¹³. L'apparition d'associations de journalistes est à la fois signe et facteur de la cohésion de la profession. En province comme à Paris, des associations de journalistes¹⁴ apparaissent sur la base de sociétés de secours mutuel, réunissant rédacteurs et directeurs. Ces associations témoignent d'une prise de conscience d'une identité de groupe et

¹³ S'il se révèle impossible de chiffrer précisément le nombre de journalistes de province à la fin du XIX^e siècle, leur effectif serait multiplié par deux et demi entre 1885 et 1890, pour stagner à partir de 1890, l'élargissement des rédactions ne parvenant pas à compenser la disparition de nombreux titres (Martin, 1997 : 123).

¹⁴ La première association manifestant la solidarité de tous ceux qui participent à la réalisation intellectuelle du journal, *l'Association de la presse républicaine départementale*, est créée en 1879 à l'initiative de Léon Brière. D'autres associations nationales de la presse départementale se développent à partir de 1882, regroupant les journalistes selon leur sensibilité politique. Des associations régionales émergent ensuite, indépendamment de toute considération politique, signe d'une intensification du sentiment d'appartenance professionnelle. On citera ainsi *l'Association de la presse de l'Est*, fondée en 1901.

de la solidarité de ses intérêts, résultant notamment du nouveau statut qu'accorde la Troisième République à la presse. Devenue l'un des trois « lieux de parole » de l'espace public, avec la tribune des chambres et celle du barreau (Van Den Dungen, 2008 : 14), le journalisme tire de cette position une justification socio-politique participant très fortement à l'émergence d'une cohésion interne de la profession et par suite, à la pratique d'un journalisme de moins en moins littéraire. La presse est désormais valorisée pour sa mission de médiateur entre une République parlementaire, ses politiques et ses citoyens. Pesant sur le pouvoir et influençant l'opinion, elle est reconnue comme nécessaire au bon fonctionnement du régime, qui l'alimente en retour d'avantages.

Au-delà de ces premiers facteurs concourant à la distanciation du journalisme avec le champ littéraire, il faudrait faire état d'un ensemble de facteurs sociologiques, technologiques et politiques concourant à l'avènement d'un « journalisme moderne » (Palmer, 1983). D'emblée, remarquons que si le substrat conjoncturel des mutations du champ journalistique est bien connu des historiens, ses influences sur les contenus et les formes discursives du journal restent en comparaison méconnues, comme l'évoque C. Bellanger :

Faute d'une documentation suffisante sur l'économie de la presse en général et des entreprises en particulier, faute d'études sur l'évolution du contenu et du style des journaux ou du monde des journalistes, il est bien malaisé de dégager les lignes directrices de ces transformations (Bellanger *et al*, 1972 : 275).

Nous aurions tendance à nuancer ce constat de C. Bellanger : les grandes lignes des transformations des productions journalistiques nous semblent avoir largement été pressenties, mais ne se sont accompagnées que de très rares études sur les caractéristiques discursives formelles des formes génériques. C. Bellanger lui-même dresse une esquisse des grandes mutations du journal :

La richesse des informations transmises par les agences ; la concurrence entre les titres, qui portait moins désormais sur l'originalité des articles que sur l'exclusivité et la rapidité de transmission des informations ; la mode, généralisée après 1890, des interviews ; le goût croissant des lecteurs pour les faits divers et les sports ; l'accélération aussi de la vie politique nationale et

internationale, entraînent une profonde et constante transformation du contenu des journaux qui élargissent leur champ d'information et différencient plus nettement leurs rubriques (Bellanger *et al*, 1972 : 279).

Il nous semble primordial de souligner que les dernières décennies prennent acte de la constitution du journalisme comme champ faiblement autonome à l'égard des contraintes économiques. Tandis que le journal devient un « objet de consommation » de masse et que son lectorat s'élargit, les liens entre journalisme et logique économique se renforcent très nettement. S'inscrivant désormais dans un champ concurrentiel, le journal met en œuvre des stratégies pour s'attirer les faveurs de la masse flottante des lecteurs. Les productions discursives en sont plus ou moins directement affectées. La logique économique et concurrentielle prévalant à la fin du XIX^e siècle explique la diminution du prix de vente du quotidien, ainsi que l'adaptation du journal au goût du lecteur : le journal ouvre aussi ses colonnes à de nouveaux contenus (le sport par exemple) et à de nouvelles formes génériques, comme l'interview. Dans le même temps, le journal donne une place croissante aux contenus tels que les faits-divers et les formes génériques qui ont fait son succès.

Conjointement, l'activité journalistique a été bouleversée par les évolutions techniques majeures de traitement et de diffusion de l'information. La construction des réseaux ferroviaires français, accélérant le travail des postes, a été un premier pas vers une diffusion rapide et étendue de l'information. Le traitement de cette dernière a connu une réelle mutation avec l'installation du réseau des câbles télégraphiques dont la première ligne (Paris-Rouen) date de 1845. Dès 1878, les dépêches télégraphiques commencent à se généraliser dans les journaux, du moins pour les informations nationales ; il faut en effet attendre 1884 pour les informations étrangères (Palmer, 1983 : 40). Ces mutations techniques engendrent de profonds bouleversements du « visage du quotidien » (Guéry, 1997) : décuplant le nombre d'informations que le journal est désormais en mesure d'apporter, le télégraphe participe à une diversification des contenus du quotidien régional, et par le fait, à un élargissement de son horizon. Il est également des mutations relatives au style

même de l'écriture journalistique, désormais soumise à de plus fortes contraintes quant au nombre de signes disponibles pour traiter une information. Autre condition de production, autre genre, autre trait d'écriture : celui introduit par le télégraphe, tourné vers le factuel, sera succinct, discontinu et abrupt. En cela, l'évolution technique participe au passage d'une « presse d'opinion » à une presse dite « d'information ». Nous prendrons néanmoins nos distances avec cette terminologie réductrice en cela qu'elle tend à suggérer une évolution radicale et brutale vers une écriture journalistique dépolitisée et désobjectivée, dont nous contestons la possibilité même d'existence. L'analyse du discours et les linguistiques énonciatives montrent en effet que l'objectivité sacralisée de l'écriture journalistique est une illusion : il n'y existe pas de parole désobjectivée quel que soit le cadre pragmatique dans lequel elle s'inscrit (informer, commenter, etc.) mais uniquement un discours tentant de dissimuler ses traces par divers « simulacres énonciatifs » (Rabatel & Chauvin, 2006). À la conception d'un discours neutre, objectif, ces disciplines opposent donc celle d'un discours *objectivé* et s'attachent à mettre à jour les phénomènes langagiers et discursifs sollicités par le locuteur dans ce cadre (Rabatel, 2004a ; Marnette, 2004, Adam, 1997 ; Koren, 1996). C'est cette tension entre *objectivation* et *subjectivation* traversant nécessairement l'écriture de presse qui nous semble bien entendue dans ce remarquable extrait prescriptif de L. Jules, apportant un éclairage des plus précieux sur la « neutralité » affichée du discours d'information. Ce passage décrit sans ambages les ressorts d'une objectivation de surface de l'écriture journalistique qui dissimule son argumentation et les traces de sa subjectivité derrière un discours doxique et un régime allusif :

Le journal d'information tient à sa neutralité, qui fait sa force. Il invoquera l'honneur, l'intérêt du pays, la morale, la vertu, tous les sentiments généraux qui se prêtent aux argumentations imprécises et dispensent des raisons topiques, des motifs secrets et nettement déterminés, que l'intérêt de la propagande ne permet pas de dévoiler. Il évitera les longs articles où, bon gré mal gré, il faudrait en venir aux détails et tomber dans la politique de parti. Un court entrefilet, une phrase adroitement glissée dans un compte rendu, une citation de journal, un écho habilement présenté, une simple insinuation dans quelque

article de reportage lui suffiront amplement pour la besogne qu'il veut faire. Ces quelques mots, négligemment jetés, porteront certainement dans l'esprit du lecteur, d'abord parce que le lecteur de ces feuilles est à l'ordinaire peu cultivé, parce qu'il croit en son journal dur comme fer, ensuite parce que ces quelques mots seront les seuls que lui fera entendre le journal, touchant la politique¹⁵.

Les dernières décennies du XIX^e siècle sont certes une période charnière dans l'histoire du journalisme, notamment parce qu'elles voient l'émergence d'un « journalisme d'information ». Cependant cette opinion s'appuie sur l'analyse exclusive des grands titres nationaux. Pour notre part, il nous semble pertinent de procéder à un examen des formes et des contenus de la presse régionale sans *a priori*. Le « journalisme d'information » est-il compatible avec la fonction de porte-parole d'une sensibilité politique ?

2.2.2. Présentation du *Petit Comtois*

Le lancement du *Petit Comtois*, le 1er août 1883, renvoie tout d'abord à un contexte socio-historique d'essor de la presse nationale et régionale. À la fin du XIX^e siècle converge en effet un ensemble exceptionnel de facteurs qui participent à l'avènement de l'« âge d'or » de la presse. Sans entrer dans le détail d'un processus ayant donné lieu à une abondante littérature¹⁶, nous nous limiterons ici à évoquer brièvement ces facteurs.

2.2.2.1. Contexte général de parution

La promulgation de la loi du 29 juillet 1881 a ainsi constitué une étape déterminante dans l'évolution de la presse. Émanant d'un projet républicain visant à ériger la presse en instrument de la démocratie¹⁷, cette loi augmente

¹⁵ Jules, L. (1910), « Le Petit Parisien », in *Romans-Revues*, p. 27, cité par Palmer (1983 :18-19).

¹⁶ Parmi les ouvrages de références : Bellanger (1972), Avenel (1900), Charle (2004).

¹⁷ Rappelons le discours d'Eugène Pelletan, rapporteur de la loi au Sénat : « La presse à bon marché est une promesse tacite de la République au suffrage universel. La presse, cette parole présente à la fois partout et à la même heure, grâce à la vapeur et à l'électricité, peut seule tenir la France toute entière assemblée comme sur une place publique, et la mettre, homme par

considérablement la liberté de la presse en légiférant sur l'imprimerie, la librairie, la presse périodique, l'affichage, le colportage et la vente sur la voie publique. Les lourdes formalités administratives que connaissait toute rédaction furent levées avec la suppression du cautionnement, posant les conditions d'une liberté de publication et de diffusion. La définition précisée par ce texte de loi de ce qui était considéré comme délit de presse fut une avancée décisive pour la liberté d'expression : désormais, seuls les appels des militaires à la désobéissance, la provocation directe aux crimes et aux méfaits, la diffamation des souverains étrangers, les offenses au Président de la République étaient punissables.

L'âge d'or de la presse n'aurait pu être s'il n'avait pas bénéficié des progrès techniques qui ont marqué le XIX^e siècle. Nous avons déjà évoqué les impacts de l'évolution des transports et des moyens de communication. Il reste à mentionner les mutations des techniques de fabrication du quotidien. Sur le plan des techniques de composition et d'imprimerie, rappelons que jusqu'à la fin du XIX^e siècle, la composition des journaux était manuelle. L'opération de composition, lente et laborieuse, impliquait de reproduire immuablement les mêmes gestes que ceux utilisés depuis des décennies : le typographe levait les lettres les unes après les autres dans la casse (sorte de grand tiroir divisé en compartiments et qui servait autrefois à ranger les caractères en plomb) et les assemblait dans un composteur – une sorte de règle formée de deux lames en V– pour composer des mots et par suite une ligne. Inutile de préciser que la vitesse de production des feuilles était alors excessivement faible, et constituait un frein au traitement de l'actualité. Grâce à l'invention en 1885 de la linotype, première machine à composer en lignes-blocs, la vitesse de composition fut multipliée par dix, on atteignait les 7000 à 8000 signes à l'heure (Guéry, 1997 : 69), ce qui permettait de répondre aux besoins des quotidiens régionaux

homme, jour par jour, dans la confiance de tous les événements et au courant de toutes les questions ».

entrés dans un univers concurrentiel. Parallèlement à ces progrès en matière de composition, l'industrie de la presse connut une innovation de la plus grande importance lorsqu'en 1865, l'ingénieur français Marinoni mit au point sa presse rotative à bobines¹⁸ qui allait permettre la naissance et le développement de la presse à grand tirage.

Les progrès techniques relatifs à la production et à la diffusion des quotidiens n'ont cependant participé à l'essor de la presse que dans la mesure où il existait un public. Entendons ici l'extension du public instruit au XIX^e siècle, suite à une longue évolution remontant à la loi Guizot témoignant d'une volonté que les lois Ferry ne firent que confirmer. Cette relative démocratisation de l'instruction permit l'accession à l'information à un public plus large que celui des salons fréquentés par les notables lettrés d'une société clivée par ailleurs entre urbains et ruraux. Ce dernier clivage tend justement à s'estomper au cours du XIX^e siècle du fait d'un processus massif d'urbanisation. Parallèlement, du fait de l'évolution des méthodes de diffusion et des stratégies commerciales des quotidiens, la presse devient accessible à un public plus large. Jadis vendu à l'abonnement et très souvent partagé entre plusieurs lecteurs, le journal à 5 centimes devient un objet individuel et familial, s'offrant à la lecture des milieux modestes, mais aussi des femmes et des jeunes gens, à qui la lecture au café ou dans les lieux publics n'était pas jusqu'alors permise. Des facteurs d'ordre sociologique sont également soulignés par l'historien M. Martin (2002), qui appelle à prendre également en compte l'existence d'un phénomène de *mimesis* pour reprendre la terminologie de P. Bourdieu, à même

¹⁸ Nous en rappellerons brièvement le principe en nous appuyant sur la description qu'en donne L. Guéry (1997 :119) : « Le principe est le suivant : l'impression est obtenue par la rotation de deux cylindres, l'un portant des clichés stéréos qui représentent la page de journal en relief, l'autre un habillage souple. La rotative reçoit le papier à imprimer d'une bobine qui se déroule après son engagement entre les deux cylindres du premier élément. La bande de papier se trouve alors imprimée d'un seul côté avant de passer dans un second élément qui imprime l'autre côté ».

d'analyser en partie l'engouement pour la presse de la part de son nouveau lectorat :

Pour une somme modique, l'ouvrier citadin s'adapte aux nouvelles formes de consommation ; plus que cela, en achetant le journal, il imite le bourgeois, s'élève socialement et culturellement (Martin, 2002 : 146).

2.2.2.2. *Le Petit Comtois* : un exemple typique de journal régional sous la Troisième République ?

Après avoir évoqué le contexte général de parution du *Petit Comtois*, nous voudrions à présent proposer certains jalons dans l'histoire de ce journal, en nous appuyant sur la lecture des recherches d'historiens tels que L. Mairry (1990, 1992), de R. Marlin (1965a, 1965b) et M. Palmer (1983, 2006) ainsi que sur la consultation des archives de la famille des imprimeurs du *Petit Comtois*¹⁹. Conformément à l'évolution observée au niveau national selon laquelle les tirages de la presse régionale sont multipliés par trois en province de 1880 à 1914, la presse du Doubs connaît également un essor exceptionnel à partir de 1880. Selon R. Marlin, les quatre-cinquièmes des deux cents journaux recensés dans le Doubs²⁰ datent de cette époque au cours de laquelle chaque chef-lieu d'arrondissement se dote d'un quotidien. La durée de parution de la majorité d'entre eux se révélera cependant éphémère :

L'accroissement considérable du nombre des publications après 1880, ne doit pas faire illusion, plus des deux tiers des journaux et des périodiques qui virent le jour dans les départements vécut moins d'une année [...] (Bellanger, 1972 : 395).

¹⁹ Nous remercions chaleureusement J.-L. Millot et C. Millot pour avoir très aimablement mis à notre disposition ces précieuses archives, dont le fonds extrêmement volumineux est en cours de numérisation. On trouvera sur le site du *Petit Comtois* un certain nombre de documents (photographies, correspondances, etc.) qui seront progressivement enrichis d'autres éléments susceptibles de présenter un grand intérêt pour l'historien de la presse.

²⁰ Cf. Discours de Mairry « La Presse du Doubs » à la société d'émulation mars 1959, non publié, disponible dans les archives Millot.

À cet égard, la longévité du *Petit Comtois*, publié de 1883 à 1944 est exceptionnelle. La création d'une grande majorité des feuilles régionales répondait à un dessein politique : il s'agissait d'assurer la pénétration idéologique d'une sensibilité politique, républicaine opportuniste ou radicale, royaliste ou encore socialiste²¹, d'exposer le programme d'un parti et de lutter contre celui de ses adversaires, et par là, d'assurer la victoire électorale de tel courant politique au niveau local. La création du *Petit Comtois* ne déroge pas à cette tendance. La parution du premier numéro du *Petit Comtois* le 1^{er} août 1883 concrétise l'aboutissement d'un projet porté depuis 1881²² par Jules Gros²³, avocat bisontin et Vénérable de la Loge maçonnique, ancien sous-préfet et futur député. Comme le souligne M. Palmer (2006 : 41), J. Gros tente alors d'inciter A. Rambaud (1842-1905), normalien, chef de cabinet au ministère de l'Instruction publique de Jules Ferry, à se porter candidat dans la perspective proche des élections de l'été 1881. Pour assurer l'élection de ce dernier, J. Gros préconise la création d'un journal :

[...] en politique, on n'a guère égard qu'aux gens que l'on craint. Or, vous n'avez pas de journal, et déjà vos concurrents vous battent en brèche par insinuations.²⁴

Cette feuille devrait, selon lui, impérativement emprunter la formule du « quotidien à un sou » mise au point par M. Millaud, créateur du *Petit Journal*.

²¹ Dans le Doubs s'opposent notamment *L'Eclair Comtois* (1903-1950), organe du courant royaliste et *Le socialiste Comtois* créé en 1892.

²² Le nom du quotidien de 1881 était *Le Petit Franc-Comtois*.

²³ Jules Gros, né en 1838, descendant d'une vieille famille bisontine, était le neveu de Théodore Rousseau (peintre de Barbizon) par sa mère. Il devient avocat et fait dans le même temps ses premières armes de journaliste en bataillant contre l'Empire dans un journal d'opposition, *Le Doubs*, si violemment opposé à l'Empire et à Napoléon III que le gérant du journal récolte seize mois de prison en moins d'un an et demi de parution. Après le 16 mai, il entre dans l'administration et devient successivement sous-préfet de Montbéliard en 1878 et administrateur du Territoire de Belfort en 1879. Il donne sa démission de ce poste en 1883 au gouvernement de Waldeck-Rousseau pour des raisons politiques, et décline les fonctions de trésorier-payeur général qui lui étaient offertes en compensation, sous prétexte « qu'il ne savait pas faire une addition », et consacra son temps à la rédaction du *Petit Comtois*.

²⁴ J. Gros à A. Rambaud, 19.7. 1881. A.N. 81 AP. Archives Rambaud, cité par Palmer (2006 : 41).

Le prix qui rend le journal accessible à tous et la formule alliant information et divertissement, sont entendus comme les deux garanties qui permettront d'atteindre un large lectorat, et de le mobiliser en temps d'élections :

[...] dans les trois départements, il y a une masse de 900 000 habitants à exploiter, étant ou à peu près pouvant être atteints par le journal avant midi... Pour réussir comme journal, il faudrait se garder de faire trop de politique et se contenter d'être un journal de nouvelles. Il y a toujours le temps de se livrer aux polémiques dans les moments décisifs...²⁵.

Le projet de 1881 n'aboutit pas mais pose les bases de ce qui sera, trait pour trait, *Le Petit Comtois*, journal créé en 1883. Jules Gros et Alfred Rambaud trouvent les actionnaires et les fonds nécessaires à la création du journal d'un capital de 100 000 francs en les personnes des industriels protestants Savoye et Bersot, du cousin de Jules Gros, V. Delavelle, notaire et maire de Besançon, du peintre Fanart et du docteur Perron. « Quotidien à un sou », *Le Petit Comtois*, conformément au projet de 1881, met l'accent sur sa fonction d'information dès les premières lignes parues du journal :

²⁵ J. Gros à A. Rambaud, 19.7. 1881. A.N. 81 AP. Archives Rambaud, cité par Palmer (2006 : 46).

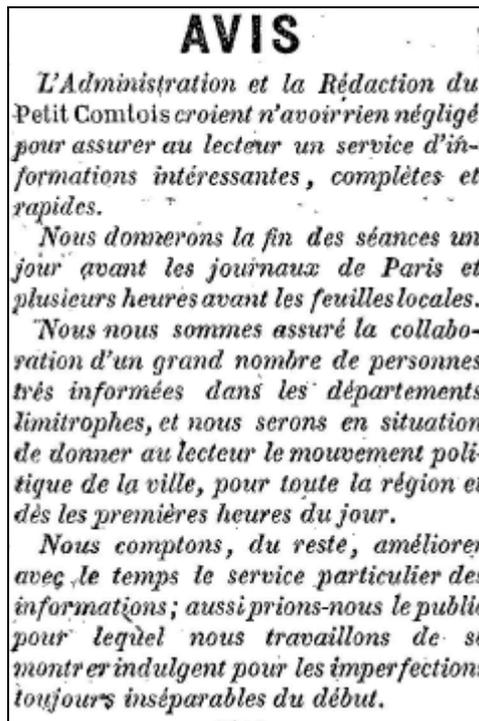


Figure II-22 : « Avis », (1883-08-01), p.1.

Cette fonction apparaît néanmoins subordonnée à l'objectif premier du quotidien, objectif qui se manifeste dès sa manchette : la feuille est avant tout un « Journal Républicain démocratique quotidien ». La rédaction met l'accent sur l'appartenance du journal à une presse d'opinion dès son premier numéro :

On nous demande qui nous sommes. Nous allons le dire : En 1869, c'est-à-dire à une époque où ceux qui nous accusent aujourd'hui de jeter la division dans le parti républicain gardaient un silence diplomatique, nous nous sommes rencontrés un certain nombre d'hommes résolus à dire leur opinion. [...]

Aujourd'hui, nous nous retrouvons les mêmes hommes, à très peu de chose près, non moins résolus à faire connaître notre sentiment, et ne réclamant en cela que le droit inaliénable qui appartient à tous les citoyens (*Le Petit Comtois*, 1883-08-01).

Dans un contexte où le camp des républicains est divisé entre une tendance radicale, ayant pour chef de file G. Clémenceau, qui appelle à une République démocratique et sociale d'une part, et d'autre part, une tendance modérée qualifiée d'opportuniste, *Le Petit Comtois* s'affiche ouvertement de la seconde tendance en 1883. Citant Gambetta, le quotidien énonce sa conception de

l'idée républicaine : marqué par le souvenir de la Commune et l'horreur de la journée du 18 mars 1870, *Le Petit Comtois* se targue de placer « la patrie bien avant [ses] préférences personnelles », plaide avec fermeté pour la réalisation de réformes impérieuses tout en reprochant aux radicaux leur utopisme, qui participe selon lui à menacer la stabilité de la République :

Le danger, il est dans l'avortement des réformes sans cesse promises et toujours éludées. Il est dans cette alliance singulière des libéraux trembleurs avec les intransigeants, qui repoussent toutes les améliorations, sous prétexte qu'elles ne sont pas suffisamment radicales, et dans la paralysie qui s'ensuit (*Le Petit Comtois*, 1883-08-01).

S'élevant « *contre l'avilissement du pouvoir exécutif* », *Le Petit Comtois* exige de la « virilité » de ses représentants, afin d'établir « *un gouvernement fort au-dedans comme au-dehors* ». La formule du quotidien rencontre dès sa parution un franc succès. Dès la fin de septembre 1883, *Le Petit Comtois* tire à 5000 exemplaires. Le quotidien fait cependant face à la concurrence de *La Petite France de l'Est*, imprimée à Dijon, dont l'audience s'étend au détriment de celle du *Petit Comtois*. Ce dernier se plaint dans ses colonnes de cette concurrence déloyale et accuse D. Wilson de soudoyer une « horde de camelots » pour s'ouvrir le marché franc-comtois :

<p>La <i>Petite France de l'Est</i>, journal de M. Wilson, gendre du président de la République, sous prétexte de se plaindre que nos vendeurs seraient autorisés à faire ce qui est interdit aux siens, insinue que le <i>Petit Comtois</i> est « payé par les fonds secrets. »</p> <p>Cette accusation ridicule nous a été adressée plusieurs fois par une « feuille de chou » avec laquelle, par respect pour nos lecteurs, nous ne voulons rien avoir à démêler.</p> <p>Mais nous n'entendons pas permettre à un journal dirigé par le gendre du chef de l'État d'insinuer ce qu'il sait parfaitement être mensonger.</p> <p>Le <i>Petit Comtois</i> est la propriété d'une société anonyme au même titre que la <i>Petite France</i>. Il est absolument indépendant du pouvoir exécutif aussi bien que du pouvoir législatif.</p> <p>La <i>Petite France</i>, qui porte la marque de M. Wilson, n'en pourrait dire autant.</p> <p>Nos crieurs ont tous reçu pour instructions de se conformer aux arrêtés de police qui réglementent la vente des journaux sur la voie publique. Si, parfois, l'un d'eux s'écarte de ses instructions, il tombe sous le coup d'une contravention et méconnaît les ordres qu'il a reçus. On ne peut nous en demander davantage.</p> <p>Pourquoi, d'ailleurs, la <i>Petite France</i> entre-t-elle sur cet terrain brûlant pour elle?</p> <p>Nous n'avons jamais cru devoir relever ce que tout le monde sait à Vesoul, à Besançon et ailleurs : à savoir que la <i>Petite France de l'Est</i> donne souvent deux journaux pour cinq centimes, et parfois pour rien ; qu'elle paie, depuis plus d'un an, des sommes folles à ses crieurs, et que, dans la Franche-Comté, elle n'en est pas moins accueillie avec une rare défiance.</p> <p>Rappellerons-nous les manifestations grotesques de l'armée de camelots que M. Wilson déchaîne de temps à autre sur nos départements? Mangin n'a jamais fait tant de tapage, même avec la collaboration de <i>Vert-de-gris</i>.</p> <p>Qui n'a vu à Besançon, notamment, ce groupe de crieurs, dont une femme portant une pancarte à l'instar de celles pro-</p>	<p>menées par les acrobates américains, et hurlant, en compagnie de ses deux auxiliaires, toutes sortes d'annonces plus ou moins fantaisistes.</p> <p>Honneur au beau sexe ! C'est l'employée de M. Wilson qui remportait la palme dans ce trio charivarique.</p> <p>« Demandez la <i>Petite France de l'Est</i>, clameur-elle, le journal donnant les nouvelles de Besançon vingt-quatre heures avant les journaux de Besançon ! »</p> <p>Du coup, Mangin était dépassé.</p> <p>Le Bisontin, né malin, riait et passait outre. Aussi la horde des camelots soudoyés par M. Wilson, et qui lui coûtait environ 1,000 pour 100, finit-elle par s'en aller chercher un public plus naïf.</p> <p>Nous n'avons rien dit de tout cela, et c'est la <i>Petite France</i> qui vient aujourd'hui nous chercher une mauvaise querelle à propos de prétendues faveurs dont nous serions l'objet à Vesoul !</p> <p>Nous n'avons ni sollicité ni obtenu de faveurs, et si la <i>Petite France</i> avait pris la peine de s'informer, elle aurait su que notre crieur de Vesoul est traité absolument sur le même pied que ses confrères.</p> <p>Quant à la ridicule accusation qu'elle nous lance à ce propos, sans doute pour se venger du succès du <i>Petit Comtois</i> dans la Haute-Saône, alors qu'elle y éprouve un lamentable échec, nous n'entendons pas qu'elle se renouvelle.</p> <p>Ceux qui prétendent que le <i>Petit Comtois</i> est payé par les fonds secrets en ont menti.</p>
---	--

Figure II-23 : Extrait du *Petit Comtois* (1884-12-18), p. 3.

La Petite France de l'Est est un des nombreux quotidiens du gendre de Jules Grévy, Daniel Wilson (1840-1919), qui utilise ses feuilles à des intérêts politiques et économiques bien compris. *Le Petit Comtois* résiste pendant un certain temps à cette concurrence en repoussant les propositions de rachat de D. Wilson. Néanmoins, lorsqu'en février 1885, J. Gros prend connaissance des rumeurs circulant selon lesquelles la société « Petite France » de Wilson compte créer un journal à un sou à Besançon, il accepte une fusion avec cette dernière. Le 22 février 1886, une société commanditée par Wilson et Stahl, dont

la direction et la gérance furent confiées aux frères Millot²⁶, chefs d'atelier du *Petit Comtois*. J. Gros accepte cette fusion pour des raisons complexes qu'on ne saurait réduire à des causes économiques. Depuis 1885, des différends relatifs à la ligne politique se font ressentir entre J. Gros et A. Rambaud, fidèle gambettiste, ainsi qu'avec les actionnaires du journal. La chute du ministère Ferry auquel succède le ministère Brisson ainsi que les difficultés du *Petit Comtois* à s'implanter dans certaines zones géographiques, que J. Gros analyse comme autant de signes de mauvais augure pour le succès de son camp politique aux élections à venir, sont également des raisons pour lesquelles la fusion semble une solution judicieuse pour J. Gros. Ce qu'il explique à A. Rambaud, condamnant pour sa part toute alliance financière et politique avec D. Wilson :

Vous semblez ne tenir aucun compte de la chute du ministère Ferry et de l'avènement du ministère Brisson. Moi, qui avais à diriger la politique du journal, je ne pouvais me permettre de pareilles licences, et j'étais obligé de compter avec les faits...Je soutiens que la seule politique gouvernementale et patriotique dans les circonstances actuelles c'est de soutenir le ministère Brisson...C'est ce terrain neutre de la politique ministérielle sur lequel je voulais placer la fusion²⁷.

Quelques mois seulement après cette fusion, le 4 novembre 1886, *Le Petit Comtois* décide de fusionner ses administrations et ses imprimeries avec *La Démocratie franc-comtoise* (radicale modérée) afin de réduire ses coûts : en avril 1887, les journaux sont transférés au 20, rue Gambetta dans un immeuble dont la construction a été commandée par D. Wilson. L'historien L. Mairry (1992 : 74) ne manque pas de relier ces éléments commerciaux aux enjeux politiques dont ils sont indissociables : « *C'est la politique de conciliation républicaine appliquée aux affaires commerciales et journalistiques !* » (Mairry, 1992 : 74).

²⁶ Le journal devint dès lors l'entreprise familiale des Millot : après la mort de Georges Millot en 1903, son frère Jean racheta toutes les parts et devint seul propriétaire de l'affaire qu'il confia à ses fils, Louis et Paul, en 1913.

²⁷ J. Gros à A. Rambaud, 27 mai 1885, cité par Palmer (2006 : 44).

Aux obsèques de l'imprimeur du *Petit Comtois* G. Millot en 1903, J. Gros déclare que « La création du *Petit Comtois* n'était pas l'œuvre de politiciens en quête de popularité pour la satisfaction de leurs ambitions personnelles ». Cette affirmation convient d'être partiellement nuancée à la lueur des remarques précédentes, pointant les facteurs politiques voire politiciens façonnant l'histoire du quotidien. L. Mairry (1992) remarque que l'évolution politique du *Petit Comtois* est indissociable des vues politiques de J. Gros. Ce dernier, bien conscient de l'évolution de la progression des radicaux dans le département dès 1889, se rallie au radicalisme en 1891, motivé par l'espoir de remporter les élections. *Le Petit Comtois*, dont la ligne politique est livrée au seul pouvoir de décision de J. Gros, devient dès lors lui aussi radical, comme le décrit L. Mairry :

En fait, depuis décembre 1890, Wilson et Stahl ne s'intéressant plus qu'aux bénéfices de la société, le journal est entre les mains des frères Jean et Georges Millot, les imprimeurs, qui ont confié la direction politique à Gros, lequel a décidé d'épouser les vues de la majorité des électeurs bisontins (Mairry, 1992 : 106).

On peut ainsi lire le 1^{er} Février 1891 à la Une du *Petit Comtois* : « Nous adhérons sans réserve au programme [radical] arrêté à Besançon par la ligue des intérêts républicains ».

Fraîchement rallié au radicalisme, le *Petit Comtois* dialogue tout d'abord de façon cordiale avec les socialistes : preuve en est l'aménagement d'une rubrique du nom de *Chronique socialiste* dans les colonnes du quotidien en 1893. Néanmoins, les rapports entre socialistes et radicaux se dégradent rapidement dans le climat de tension provoqué par les élections. À l'annonce de la nomination d'E. Charlot comme candidat du groupe socialiste bisontin aux législatives du 20 août 1893 pour Besançon-Ville, *Le Petit Comtois* décide de rompre tout échange avec cette tendance et de la combattre âprement. L'article ci-dessous, par lequel le *Petit Comtois* oppose le socialisme « pratique » des radicaux au « socialisme dogmatique et collectiviste », en témoigne :

<p>UTOPIE DU COLLECTIVISME</p> <p>Prenant texte de l'imperfection, reconnue par tous les esprits éclairés, de notre situation sociale actuelle, le citoyen Charlot prétend, comme tous les apôtres d'une idée, à la supériorité de l'école collectiviste dont il est le porte-parole.</p> <p>Qu'est-ce donc que le <i>collectivisme</i> ?</p> <p>C'est le travail produit au compte de l'Etat.</p> <p>C'est l'Etat vendeur des produits.</p> <p>C'est l'Etat distributeur des bénéfices en résultant.</p> <p>C'est l'Etat englobant toutes les fabrications, tous les commerces, toutes les industries.</p> <p>C'est l'Etat cultivateur.</p> <p>C'est l'Etat maître supérieur et direct de tous les travailleurs.</p> <p>C'est aussi l'Etat rémunérant tous ces travaux et répartiteur du produit de leur vente.</p> <p><i>Conséquence.</i> — C'est la suppression de toute initiative personnelle, c'est l'abandonnement des prédispositions intellectuelles de l'individu.</p> <p>C'est l'homme asservi à travailler de tel ou tel métier, tant d'heures par jour et payé en conséquence.</p> <p>Qui ne voit immédiatement la conséquence de ce système ?</p> <p>La collectivité, c'est donc, en application, un encouragement à la paresse ; c'est la suppression de toute supériorité individuelle, car, sûr d'obtenir son salaire en travaillant peu comme en travaillant beaucoup, l'ouvrier travaillera le moins possible. En un mot, c'est casser le grand ressort de l'émulation, qui est l'âme du progrès.</p> <p>C'est un retour à l'Etat barbare, où la collectivité, la tribu était propriétaire, produisait et distribuait les produits, et toutes ont croupi dans l'ignorance et dans la grossièreté.</p>	<p>Nous disons plus : elle avilit les caractères. L'Etat producteur et distributeur des produits, pour arriver à l'équité, devra faire surveiller le travail, l'enregistrer, contrôler les surveillants, les distributeurs, transporter, emmagasiner, conserver les produits. Toute cette armée de fonctionnaires est aussi une armée de travailleurs, et ils auront droit au partage comme les autres ; mais ils auront sur eux, de degré en degré, une autorité de plus en plus grande, puisque d'eux dépendra la répartition à intervenir en faveur des travailleurs producteurs : d'où soumission toujours de plus en plus avilissante de degré en degré, jusqu'à celui de directeur général, qui, soutenu par ses employés, n'aura pas de peine à accaparer un beau jour toute l'autorité et rétablira le <i>despotisme engendré par cet état social</i>.</p> <p>L'initiative individuelle, au contraire, élève et ennoblit le caractère, parce qu'elle laisse à chacun sa liberté personnelle, que personne ne se souciera d'aliéner pour s'asservir à l'armée de fonctionnaires que l'établissement du collectivisme entraînera d'une façon inéluctable.</p> <p>Le collectivisme, nous en avons des exemples sous les yeux, mais en petit : ce sont les couvents et les bagnes. C'est à cet état social que sera réduite la société.</p> <p>Oh ! la jolie perspective ! Nous aurions été heureux d'entendre le citoyen Charlot nous exposer <i>la mise en pratique</i> du collectivisme qu'il prône partout. Nous n'avons pas eu encore, aujourd'hui, cette satisfaction, et nous sommes bien certains qu'il ne nous la donnera pas. Entrer dans les détails, ce serait condamner son système. Esquiver les explications est bien plus commode. Nous voyons M. Alviset de Maizières en faire autant de son programme. Tous deux, ils fuient la lumière.</p> <p style="text-align: right;"><i>Un socialiste pratique.</i></p>
---	---

Figure II-24 : « Utopie du collectivisme », (1893-08-17), p.1.

Au début des années 1900, des dissidences se font ressentir dans le parti radical. Elles sont liées aux élections au Conseil Municipal dans le canton sud de Besançon. Des opposants à la décision de la Fédération des Comités qui avaient été évincés du conseil, se regroupent pour former une feuille hebdomadaire nommée *Les Nouvelles Comtoises*. Cette feuille attaque de manière virulente la municipalité et *Le Petit Comtois*. Ce dernier était calomnié dans le même temps par des hebdomadaires socialistes, en raison de la

participation de L. Millot à la propagande radicale. *Le Petit Comtois* répondait à ces attaques avec fureur jusqu'à crouler sous les procès intentés par les autres organes de presse. Au lendemain de la campagne législative de 1902, *Le Petit Comtois* avait encore plus de 200.000 francs de procès politiques à liquider. La complexité de la situation s'entend dans le passage suivant extrait d'un discours de Jules Gros, affirmant :

Ce n'était pas sous une argumentation supérieure qu'on prétendait nous écraser, mais sous le poids de l'argent²⁸.

Les positions du *Petit Comtois* ne lui valurent pas que des procès et des pertes financières. Sa position anticléricale engendra également des tensions d'un autre ordre. Après avoir été rappelé à l'ordre à plusieurs reprises, le journal fut frappé d'anathème par l'archevêque Mgr Gauthey le 25 février 1913. Les motifs de cette sanction sont énumérés dans « L'ordonnance relative à la vente, l'achat et la lecture du *Petit Comtois* » publiée dans les *Semaines religieuses* du 1^{er} mars 1913 que nous soumettons à l'appréciation de notre lecteur. Au terme de celle-ci, l'archevêque de Besançon déclare « coupable de péché grave tout fidèle qui, sans motif sérieux, [...] achètera, vendra ou lira habituellement le *Petit Comtois* ».

²⁸ Gros, J., 1903, discours aux obsèques de M. Millot, Archives Millot.

ORDONNANCE	
<p style="text-align: center;"><i>relative à la vente, à l'achat et à la lecture du journal quotidien</i> LE PETIT COMTOIS</p> <p>NOUS, FRANÇOIS-LÉON GAUTHEY, par la grâce de Dieu et du Saint-Siège apostolique, archevêque de Besançon,</p> <p>CONSIDÉRANT qu'une des principales obligations de Notre charge est de veiller à la conservation et à la défense de la foi catholique dans le diocèse au gouvernement duquel Nous avons été préposé (Léon XIII, <i>Encycl. Sapientie Christiane</i>);</p> <p>CONSIDÉRANT que, parmi les moyens à employer à cet effet, les saints Canons prescrivent la dénonciation aux fidèles du danger auquel ils exposent leur foi en lisant les écrits : livres ou journaux, qui attaquent la religion catholique dans son enseignement dogmatique ou moral et dans sa hiérarchie (Grégoire XVI, <i>Inter precipuas</i>; — Pie IX, <i>Qui pluribus</i>; — Léon XIII, <i>Officiorum et munerum</i>; — Pie X, <i>Pascendi Dominici gregis</i>);</p> <p>CONSIDÉRANT qu'au nombre de ces écrits manifestement hostiles à nos croyances et dangereux pour la foi des lecteurs catholiques, figure, à n'en pas douter, au premier rang, dans Notre diocèse, le journal quotidien de Besançon, le <i>Petit Comtois</i>;</p> <p>CONSIDÉRANT que ce journal, dans le cours des mois de novembre et décembre 1912, janvier et février 1913, a maintes fois parlé de Dieu, de la Providence, d'une façon impie et blasphématoire, numéros des 2 novembre, 4 novembre, 1^{er} décembre;</p> <p>Qu'il a mis en doute la vérité du personnage de Notre-Seigneur Jésus-Christ : « Christ hypothétique », 6 décembre; traité les récits de sa Nativité et de la crèche de « divines légendes », 25 décembre; mis, dans une scène triviale, Notre-Seigneur Jésus-Christ sur le même pied que Mahomet, 2 novembre; parlé sur le ton de la moquerie de la virginité de la sainte Vierge, 6 décembre, et de sa protection sur les catholiques, 8 janvier;</p> <p>Qu'il a attaqué la foi catholique en insinuant « que l'homme est la source de toute révélation », 17 février;</p> <p>Qu'il a parlé injustement de l'Église catholique en affirmant « l'autoritarisme intolérant et bautain de l'Église d'autrefois », 16 janvier;</p> <p>Qu'il a traité les pratiques de la religion catholique de « coutumes surannées et obscures », 11 novembre; appelé ses croyances « des croyances étroites et des dogmatismes intolérants », 4 janvier; comparé les mêmes croyances catholiques à « des fables, des hyperboles, des excentricités et des stupidités de la mythologie », 11 décembre;</p> <p>Qu'il a attaqué la morale chrétienne en affirmant « que la raison de l'homme doit faire, à la fois, sa loi, sa morale et sa science », 17 février;</p>	<p>Qu'il a raillé l'infaillibilité pontificale et parlé grossièrement du Pape, 18 décembre; prétendu qu'il « combat en fanatique toutes les conquêtes de l'esprit moderne et toute liberté », 13 février; qu'il a choisi pour archevêque de Gènes « un fanatique de la plus belle espèce », 13 février;</p> <p>Qu'il a tourné en dérision les actes de la hiérarchie et outragé des prêtres et des religieux qu'il traite « d'intolérants », « fanatiques » et « hypocrites », 4 décembre, 6 décembre, 8 février, 14 février;</p> <p>Qu'il a parlé des écoles catholiques de façon injurieuse, en leur opposant l'école laïque qui doit « affranchir les cerveaux et les consciences », 11 décembre;</p> <p>Qu'il met en danger la foi de ses lecteurs par la réclame qu'il fait aux enterrements civils et les félicitations qu'il donne à ceux qui s'affranchissent des lois de Dieu et de l'Église, 11 novembre;</p> <p>Qu'il essaie de tourner en dérision les pèlerinages, qu'il appelle « les grands bazars de la dévotion », destinés à « ruiner les humbles chapelles », 28 décembre, et l'usage du scapulaire et des médailles, 4 février;</p> <p>ATTENDU que les citations et les expressions visées ci-dessus sont très caractéristiques de l'esprit et de la tendance générale du <i>Petit Comtois</i> et y sont assez fréquentes pour qu'on puisse juger sagement que ce journal est une publication antichrétienne et dangereuse pour les lecteurs catholiques,</p> <p>A CES CAUSES ET POUR CES MOTIFS, Nous avons décrété et décrétons ce qui suit :</p> <p>1^o Devra être considéré comme coupable de péché grave tout fidèle qui, sans motif sérieux, dont Nous tenons à rester seul juge, achètera, vendra ou lira habituellement le <i>Petit Comtois</i>. (Nous admettons toutefois qu'on puisse s'adresser, pour ce sujet, à Nos vicaires généraux comme à Nous.)</p> <p>2^o Ce péché, en ce qui concerne l'admission aux sacrements, sera traité, par les confesseurs, comme les autres péchés d'habitude, selon les règles de la théologie morale.</p> <p>Et sera Notre présente Ordonnance, avec la lettre qui la précède, lue et publiée dans toutes les églises de Notre diocèse, le dimanche 2 mars, quatrième dimanche de carême.</p> <p>Donné à Besançon, en Notre résidence de la Porte Noire, sous Notre sceau et Notre seing et le contre-seing du Chancelier de Notre archevêché, le 25 février 1913, au septième anniversaire de Notre sacre.</p> <p style="text-align: right;">† FRANÇOIS-LÉON, Archevêque de Besançon.</p>

Figure II-25 : Ordonnance relative à la vente, à l'achat et à la lecture du journal (1893-03-01), (Archives Millot).

La mobilisation de 1914 décima la rédaction et l'administration du *Petit Comtois*. L. Millot, M. Cordelier et M. Colin, MM. Perrin et Prost furent très vite mobilisés. Ne restaient avec P. Millot, que J. Gros et G. Graff pour assurer la marche du journal. Le quotidien traversa néanmoins la première guerre mondiale sans arrêt de publication, mais se heurta à la censure, comme en témoigne G. Graff :

Le soir, vers dix heures, chaque fois que l'on entendait dans la rue le roulement de la voiture qui amenait « ces jeunes gens », — comme nous appelions entre nous messieurs les censeurs, — Jules Gros ramassait à la

hâte ses papiers dans sa serviette et partait pour ne point rencontrer nos visiteurs quotidiens.

Le matin, à onze heures, quand notre doyen arrivait au journal, il déployait le *Petit Comtois* du jour. Lorsque, sur l'ordre de la Censure, une phrase avait été enlevée dans son article, Jules Gros, frottait ses deux genoux avec ses deux mains et semblait relire la phrase échoppée puis il lançait un claironnant : « Quels Jean F..... ! »²⁹

La première guerre mondiale n'avait fait que renforcer le statut d'institution du *Petit Comtois*, dont le succès de vente se confirmait. À la veille de la deuxième guerre mondiale, le tirage du *Petit Comtois* dépassait les 35 000 exemplaires quotidiens, soit le double du tirage des autres journaux bisontins réunis. L'avènement de la deuxième guerre mondiale occasionne des difficultés sans précédents pour le quotidien. Après avoir subi une réduction de 90% de son contingent de papier, un ordre de la Kommandantur suspend officiellement la publication du *Petit Comtois* pour « mauvaise volonté » et « résistance » le 22 mai 1944. Le transfert du matériel de l'imprimerie prévu vers l'Allemagne est évité de justesse par l'avancée des alliés. À la libération, certains anciens ouvriers du *Petit Comtois* s'emparent de l'imprimerie et du journal, et publient une feuille socialo-communiste intitulé *Le Comtois* dès le 9 octobre 1944. La société Millot intente et gagne en décembre 1946 une série de procès en diffamation, mais ne peut récupérer son matériel d'impression. Ce n'est qu'en 1961 que M. Louis Millot peut réintégrer son bureau d'où il avait été chassé par les Allemands en 1944. Tout espoir de publier un nouveau numéro du *Petit Comtois* est cependant anéanti en raison de la situation économique difficile du secteur de la presse écrite.

2.2.3. Description du corpus

Regroupant des textes intégraux et non des énoncés, le corpus *Petit Comtois* correspond à ce que l'on nomme en lexicométrie une « série textuelle

²⁹ Témoignage de Graff, archives de la famille Millot.

chronologique » (Salem, 1988) : il s'agit d'une série périodique homogène de textes produits dans des situations d'énonciation similaires.

Un corpus de presse homogène

Homogène, notre corpus ne regroupe que des textes produits dans des conditions institutionnelles identiques par le même *sujet communicant collectif* (Charaudeau, 2002 : 227) : *Le Petit Comtois*, que nous envisagerons comme l'acteur social responsable des productions discursives. L'unicité du *Petit Comtois* comme « être de discours » doit bien évidemment être contrebalancée par la nature profondément dialogique et hétérogène du discours de presse. Le dialogisme du discours de presse renvoie tout d'abord à l'hétérogénéité constitutive de tout discours, afférant aux formes de relations que tout énoncé entretient avec les énoncés produits antérieurement ainsi qu'avec les énoncés à venir, que l'on peut appréhender à travers les notions de *dialogisme* (Bakhtine, 1977) et d'*interdiscours* (Pêcheux, 1975). De façon concomitante, un *dialogisme représenté* (Moirand, 2004) donne sa pleine dimension dans le discours de presse, au travers d'une monstration des discours hétérogènes qui ponctuent le fil de sa matérialité. Sous le vernis unifiant du sujet communicant collectif, se donne à lire une mosaïque de voix hétérogènes, émises par des *sujets communicants individuels* (journalistes, correspondants, politiciens, lecteurs, etc.). Cette hétérogénéité se traduit dans l'organisation matérielle des textes, et s'explicite notamment *via* les signatures des articles du corpus, même s'il ne s'agit que d'une des manifestations les plus visibles de l'hétérogénéité du discours de presse. Remarquons tout d'abord que la plupart des articles ne sont pas signés, ce qui témoigne d'une prise en charge énonciative maximale de la part de la rédaction. Nous distinguerons parmi les articles signés les signatures internes, c'est-à-dire émanant directement de la rédaction du *Petit Comtois*, des signatures externes. Les signatures renvoyant à l'équipe de rédaction du *Petit Comtois* sont d'ailleurs très rares et uniquement utilisées pour

les avis³⁰ et pour les éditoriaux au début du XX^e siècle. Les articles signés en interne correspondent soit à la signature explicite du journaliste en charge d'une rubrique, à l'instar de Ch. Charolet-Chonion, responsable de la rubrique « Chronique parisienne », soit à une signature sous pseudonyme pour des articles généralement à valeur polémique : ainsi se cache derrière « Nicolas Grindavoine » la plume de J. Gros, derrière celle de « Jean Turquis », celle de Louis Millot, et enfin, derrière « le Moine de la Vallée » et « Jean Valjean », Léon Jouffroy. La majorité des articles du *Petit Comtois* sont cependant signés par des personnalités extérieures. Il s'agit tout d'abord des brèves signées d'un correspondant spécial ou de l'agence Havas mais également des articles de politique intérieure émanant de quotidiens nationaux que *Le Petit Comtois* reproduit intégralement sans commentaire, et qui sont empruntés au journal *Le Soleil*³¹, ou *La Justice*. Notons que la reproduction intégrale d'un article publié dans un organe national n'est pas sans évoquer les pratiques des premiers quotidiens régionaux de presse du début du XIX^e siècle dont les colonnes étaient remplies par une reproduction d'articles de la presse nationale. *Le Petit Comtois* héberge également bon nombre de tribunes³² ouvertes aux personnalités politiques locales ou nationales, dans lesquelles est offerte une réaction à un fait d'actualité et qui alimentent bien souvent une polémique locale. La plupart des signatures extérieures se concentrent néanmoins dans les pages consacrées aux chroniques régionales. Il s'agit tout d'abord d'institutions qui demandent au quotidien de diffuser une information officielle, telle que la Mairie de Besançon³³, ou la Compagnie P.L.M, désireuse de renseigner les usagers de retards de trains³⁴. Ces articles correspondent ainsi

³⁰ Cf. « Au lecteur », (1893-01-01).

³¹ Cf. article « Flibustiers de fin de siècle », (1896-01-04).

³² Cf. article « Le Black-rot », (1896-01-05), signé « Paul Decker-David, Ingénieur agronome, député du Gers ».

³³ Cf. article « Caisse d'Épargne et de prévoyance » du 01-01-1893, signé « Le maire, président du conseil de direction, VUILLECARD ».

³⁴ Cf. article « Compagnie P.L.M » du 28-01-1893, signé « Le chef de gare ».

aux fameuses « Prière d'insérer » qui font de la presse régionale du XIX^e siècle l'écho des institutions. Si les chroniques régionales sont principalement constituées d'articles rédigés par des correspondants spéciaux permanents du quotidien, les chroniques régionales concentrent également un très grand nombre d'articles signés de lecteurs, qui, s'érigeant en lecteurs-correspondants, adressent un courrier à la rédaction en vue d'être publiés. L'identité de ces locuteurs reste très majoritairement cachée derrière un pseudonyme (*Jean Meurs, Spectator*), des initiales (P... ; R...), ou un syntagme cocasse comme « un groupe de curieux », « un assistant content de la journée », « Un zoulou », etc.

À cette hétérogénéité de surface, nous opposons fermement, à la suite de M. Ebel et P. Fiala (1983) les phénomènes de régulation institutionnelle de la circulation des discours par la rédaction. Cette dernière reste en effet maîtresse de la circulation des discours dans ses colonnes, qu'il s'agisse de la reproduction des communications de lecteurs-correspondants, d'articles intégraux de quotidiens nationaux ou de professions de foi. Le contrôle de la publication des discours dans *Le Petit Comtois* est d'ailleurs aisément repérable, dès lors que ne sont reproduits que des articles ou des professions de foi ne relevant presque exclusivement que de la formation discursive républicaine. Quant aux courriers des lecteurs-correspondants, de nombreux avis du journal explicitent les réglages institutionnels opérés, signalant notamment que les articles jugés trop virulents par la rédaction ne seront pas publiés. La sélection des textes publiés renvoie à la politique rédactionnelle du quotidien, étroitement liée à sa position politique.

Un corpus « transgénérique »³⁵

Notre corpus se caractérise par une forte hétérogénéité textuelle, liée à notre décision d'observer l'ensemble des unités rédactionnelles du quotidien.

³⁵ Le terme « transgénérique » est une proposition empruntée à A. Krieg (2000 : 91).

Cependant, le terme d'unité rédactionnelle est tout d'abord à entendre par opposition aux annonces publicitaires : nous avons ainsi exclu les avis, les publicités et les petites annonces. Nous avons procédé à un deuxième filtrage des unités dites rédactionnelles, relatif, cette fois, à leur consistance textuelle. Toute unité rédactionnelle se présentant comme une énumération de noms propres ou encore de chiffres a été exclue du corpus, à l'instar de la *Chronique du bien*, correspondant à une énumération des actions méritantes effectuées par des concitoyens, des bulletins financiers et météorologiques, se résumant respectivement à des listes de chiffres et de températures, des rubriques *État civil*, *Caisse d'Épargne*, *Petite bourse du soir*, etc. Chacune de ces rubriques pourrait bien évidemment donner lieu à des études instructives permettant d'enrichir nos connaissances sur la presse régionale du XIX^e siècle. Néanmoins, ces rubriques ne présentent qu'un intérêt limité dans le cadre de notre exploration du discours du *Petit Comtois*. Dans la même perspective, parce qu'ils excédaient les limites de capacité de traitement d'une analyse textuelle du discours assistée par informatique, listes et tableaux n'ont pas été inclus dans notre corpus. Les œuvres publiées en feuilleton au rez-de-chaussée de chaque numéro du quotidien, ont quant à elles été exclues de notre étude parce qu'elles constituaient un objet de recherche excédant le discours journalistique d'un point de vue méthodologique ; et d'un point de vue logistique, parce que leur numérisation était superficielle dans le sens où des versions imprimées de qualité typographique supérieure de ces œuvres étaient disponibles.

D'où une inévitable *hétérogénéité générique* de notre corpus. Loin de remettre en cause la propriété d'exploitabilité de notre corpus (Pincemin, 1999), la diversité des formes génériques se justifie par la visée de notre projet de recherche : analyser l'évolution des diverses et multiples formes journalistiques d'un quotidien régional de presse. Si le genre n'est pas ici posé comme un facteur contribuant à l'homogénéité du corpus, celle-ci est assurée par d'autres modes de cohésion tels que l'homogénéité énonciative et la diachronie.

Un corpus diachronique

Parce qu'il rassemble des textes produits dans les mêmes conditions d'énonciation en diachronie, ce corpus pourra donc être qualifié de série chronologique (Salem, 1998). Cette série chronologique constitue un dispositif d'observation des évolutions, des ruptures et des continuités des productions discursives du quotidien régional. Puisque notre corpus sera étudié selon sa propre norme, la diachronie constituera l'un des ressorts essentiels de la contrastivité du dispositif d'analyse. De par l'angle de notre projet de recherche, la diachronie est un candidat privilégié comme critère de partitionnement du corpus dans ses phases de traitement. L'ampleur diachronique de notre corpus, couvrant la période 1883-1903 (soit 21 années de parution), a impliqué de mettre en œuvre une pratique d'échantillonnage rigoureuse. Rappelons encore une fois que chaque numéro du *Petit Comtois* sous mode texte a exigé un travail minutieux de copiste, ne s'accommodant que très mal avec les conditions de réalisation des nouvelles thèses. Parce que l'opération de définition d'un échantillon est garante de la représentativité du corpus³⁶, il faut entrer dans le détail de la pratique d'échantillonnage ici suivie. Un mois de numéros quotidiens du *Petit Comtois* a été sélectionné pour représenter une année de parution. Cet ensemble partiel de données peut légitimement prétendre à représenter la population des numéros du *Petit Comtois* publiée de façon ininterrompue de 1883 à 1903. Le mois de parution, variant entre 28 et 31 numéros, constitue une unité quantitative raisonnablement substantielle et qualitativement raisonnée. La représentativité qualitative du corpus est notamment autorisée par le caractère redondant et cyclique du discours de presse : de jour en jour, le discours prend forme dans diverses formes génériques à la fois stabilisées et évolutives tandis que la temporalité du jour et

³⁶ Notamment Habert (2000).

celle de la semaine structurent le journal et déterminent la récurrence des unités rédactionnelles.

Pour des raisons logistiques liées aux conditions d'acquisition des données, le mois sélectionné pour représenter l'année de parution n'a pas été fixe : pour l'année 1883, nous avons sélectionné les 31 numéros du mois d'août, premier mois de parution du quotidien. Les années de publication continues de 1884 à 1893 ainsi que l'année 1896 sont échantillonnées par le mois de janvier. De même que les années continues 1897-1900, 1894 et 1895 sont représentées par le mois de février. Le mois de mars a enfin été choisi pour représenter la dernière série chronologique du corpus (1900-1903). Que les mois de référence utilisés soient différents pour chaque année pourrait donner lieu à des discussions méthodologiques : on pourrait notamment nous reprocher que cette configuration soit de nature à perturber le traitement thématique de notre corpus. Néanmoins, les manipulations pré-exploratoires sur ce corpus ont révélé que ces choix ne semblaient pas biaiser les analyses menées, ce qui reste bien entendu à confirmer par la réinscription de ce corpus dans une série élargie composée du *continuum* des numéros du quotidien. Basée sur une analyse des items les plus fréquents du corpus à partir desquels sont évaluées les proximités et divergences entre les partitions diachroniques de notre corpus, notre méthodologie d'exploration tend à neutraliser les rares éléments de calendrier. Pour plus de précautions, nous avons de surcroît exclu les noms de mois dans le prétraitement du vocabulaire du corpus.

Année de parution	Mois sélectionné	Nombre de numéros parus dans le mois
1883	Août (premier mois de parution)	31
1884	Janvier	31
1885	Janvier	30 (pas de parution le 02-01-1885)
1886	Janvier	30 (pas de parution le 02-01-1885)
1887	Janvier	31
1888	Janvier	31
1889	Janvier	31
1890	Janvier	31
1891	Janvier	30 (pas de parution le 31-01-1891)
1892	Janvier	30 (pas de parution le 02-01-1892)
1893	Janvier	30 (parution d'un numéro unique pour le 02-01-1893 et le 03-01-1893)
1894	Février	28
1895	Février	28
1896	Janvier	30 (parution d'un numéro unique pour les 02-01-1896 et 03-01-1896)
1897	Février	28
1898	Février	28
1899	Février	28
1900	Mars	31
1901	Mars	31
1902	Mars	31
1903	Mars	31

Tableau II-2 : L'échantillonnage du corpus *Petit Comtois*.

La sérialité du corpus comme aide à l'interprétation

Au-delà des critères de constitution garantissant l'exploitabilité du corpus, envisagé comme sa capacité à recevoir des traitements statistiques significatifs, nous voudrions nous arrêter sur le régime matériel du discours de presse et les difficultés qui en découlent quant à sa lecture.

En raison de la sérialité et de la temporalité de leur production qui les destinent à être lus puis repris, ou non, dans les numéros suivants, les textes de notre corpus, observé(s) hors de leur co(n)texte sont fréquemment opaques. Pour le lecteur du XXI^e siècle, la lecture de certains textes du *Petit Comtois* se révèle un exercice parfois périlleux, et ce, quelle que soit la nature générique des articles envisagés et le statut des unités textuelles observées (titre ou article). À cet état de fait, un constat indépassable : hors de son co(n)texte, l'article de journal perd de sa transparence (Moirand, 2004b : 135).

C'est le propre de la matérialité scripturale que d'inscrire tout texte comme renvoi à un co-texte et à un intertexte : cette dimension qu'on pourrait qualifier à la suite J.-M. Viprey (2005b : 59) d'*hypertextuelle*, nous semble néanmoins particulièrement exacerbée dans l'écriture journalistique du XIX^e siècle. Parce que la lecture du journal est supposée linéaire dans l'espace du plan et quotidienne, les textes du *Petit Comtois* mobilisent intensément une mémoire (inter)discursive supposée partagée entre le journaliste et son lecteur. L'exemple suivant permet de prendre la mesure de cette forme de renvoi à une situation extralinguistique et par suite, à d'autres textes :

Ch. II. (1) « Conseil des ministres », Dépêches de nuit Informations, (1899-01-25)

Paris, 24 février.

Le conseil des ministres, réuni ce matin au ministère des affaires étrangères, sous la présidence de M. Loubet, s'est longuement occupé des incidents de la journée d'hier.

Le président de la République a fait savoir qu'il a envoyé 1,000 fr. à la caisse de l'Association de prévoyance de la préfecture de police, et une somme égale à la caisse des services payés des gardiens de la paix.

Le caractère déictique du syntagme nominal souligné par nos soins compromet la compréhension du texte dans son intégralité. Dans le même temps que l'opacité de ce syntagme ne précise pas le sujet de la discussion tenue au conseil des ministres, l'enchaînement des deux phrases de la dépêche est rendu ambigu : doit-on voir ici une simple succession d'informations sans rapport ? Ou la succession des deux phrases pointe-t-elle bien un rapport entre les deux informations ? Si oui, lequel ? S'agit-il d'un rapport de cause à conséquence ?

L'aspect sériel du corpus *Petit Comtois* et l'hétérogénéité des unités rédactionnelles sélectionnées tend à répondre partiellement à ce type de difficultés de compréhension et d'interprétation. En raison de sa nature sérielle, le corpus *Petit Comtois* contient en lui-même le co(n)texte nécessaire à l'interprétation des renvois à un traitement précédent de l'information. La prise de connaissance du référent des « incidents de la journée d'hier » s'accomplit ainsi aisément par la navigation dans le corpus. En date du 24 février 1899, une série de dépêches traite ainsi des obsèques de Félix Faure, obsèques durant lesquelles les nationalistes de la Ligue des patriotes avec à leur tête P. Déroulède, tentent un coup d'état en ralliant à eux l'armée, comme l'indique la brève suivante :

Ch. II. (2) « Les exploits de Déroulède », (1899-02-24)

Boulevard Voltaire, M. Déroulède et les membres de la Ligue des patriotes ont acclamé les Saint-Cyriens et les fantassins, aux cris nourris de : Vive l'armée ! Ils conduisirent les soldats jusqu'à la caserne de la rue de Reuilly, où ils tentèrent de pénétrer dans la caserne. M. Déroulède, seul, y entra ; mais on ne l'en laissa pas sortir avant que tout danger d'une manifestation soit écarté.

C'est ainsi que la prise de connaissance de l'intertexte permet d'explicitier le rapport de cause à effet entre les deux phrases du texte précédent : la prise de mesures financières au bénéfice de la police et des gardiens de la paix vise à assurer le soutien de ces derniers au gouvernement, dans un contexte politique des plus instables.

Restituant les textes dans leur régime intertextuel, notre corpus inclut une partie des ressources interprétatives du corpus nécessaires à l'interprétation du

discours de presse régionale du XIX^e siècle. Dans le même temps que la configuration de notre corpus offre un réseau « hypertextuel » au texte, il en livre certains éléments contextuels au sens d'encyclopédiques et culturels. Ce point nous semble revêtir d'autant plus d'intérêt que le contexte de notre analyse est nécessairement disjoint de celui de la production du discours du *Petit Comtois*.

La sérialité de notre corpus se révèle donc précieuse face à un régime discursif profondément allusif. Un second exemple nous permettra d'insister sur cette particularité du discours du *Petit Comtois*, recourant à l'allusion pour activer ou (ré)activer des polémiques avec l'ennemi clérical, et entretenir dans le même temps une connivence avec son lectorat. Comportant une allusion des plus curieuses à l'« automobile » d'un « richissime curé » qui renonce à faire campagne, le texte suivant ne manquera pas d'éveiller la curiosité du lecteur :

Ch. II. (3) Chronique régionale, Jura, (1898-02-28)

On lit dans le *Réveil des campagnes* :

« Il paraît que notre richissime curé, M. Sallot, a renoncé à poser sa candidature contre M. Genoux, député radical.

Il s'est aperçu qu'il allait courir à un échec certain ; aussi a-t-il dissous son comité socialiste chrétien. Son journal ne sera pas fondé, et **de son automobile on n'en a plus entendu parler**. C'est vraiment dommage. »

D'abord perplexe devant un tel énoncé, nous avons procédé à une requête du substantif *automobile* dans le corpus, qui s'est révélée très congruente. Un certain nombre d'occurrences de cet item s'inscrit dans des textes à valeur polémique traitant des modalités de campagne d'un illustre clérical, M. Outhenin-Challandre, Sénateur de Haute-Saône. Celui-ci avait parcouru son arrondissement en automobile pour faire campagne. Les républicains du *Petit Comtois* ont alors érigé ce moyen de transport³⁷ en symbole de la richesse des cléricaux. Par suite, il faut lire dans le texte ci-dessus une métonymie chère aux

³⁷ Ce qui avait d'ailleurs valu à M. Outhenin Challandre le surnom de « Teuf-Teuf » par *Le Petit Comtois*.

républicains locaux, où le terme d'« automobile » réfère à la campagne électorale cléricale.

Ce sur quoi nous voudrions insister n'est donc pas tant l'intérêt d'une lecture assistée par informatique de la presse ancienne, que l'intérêt d'observer le discours de presse au travers d'un dispositif sériel de grand volume.

Un corpus de très grande taille

En termes d'unités documentaires, notre corpus équivaut à 26 967 articles. La taille des articles, évaluée en nombre de mots, est extrêmement hétérogène. Comme le représente la figure 21, 6% de la totalité des articles du corpus ont une taille comprise entre 6 et 25 mots, 23,3% des articles ont une taille comprise entre 25 et 50 mots. La tranche de taille la plus significative est comprise entre 50 et 100 mots puisqu'elle caractérise plus de 28,8% des articles du corpus. Presque un quart des articles du corpus ont une taille comprise entre 100 et 250 mots. Les articles de taille supérieure sont relativement rares, puisque seulement 5,2% des articles du corpus ont une taille comprise entre 500 et 1000 mots, tandis que le pourcentage d'articles supérieurs à 1000 mots s'élève à moins de 2,3%.

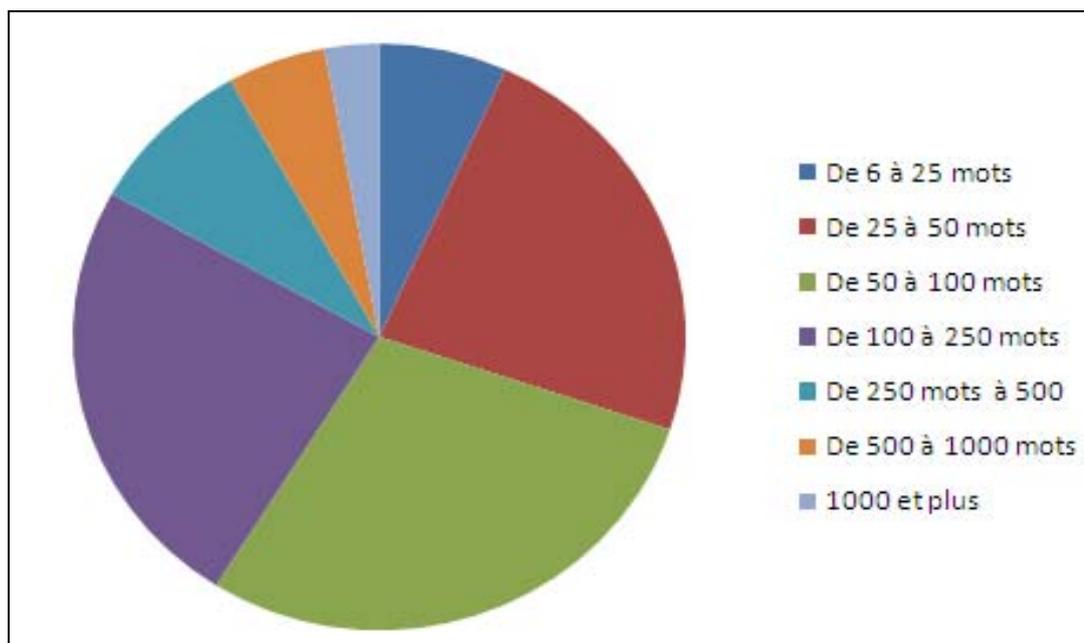


Figure II-26 : Taille des articles du corpus *Petit Comtois*.

Notre corpus représente approximativement 5,5 millions d'occurrences. À titre de comparaison, remarquons que le corpus pionnier du *Père Duchesne* réunit en 1986 comptait environ 141 180 mots et que les vastes corpus s'évaluent aujourd'hui en millions de mots : songeons notamment au très grand corpus constitué par D. Mayaffre (2004) réunissant près de 1800 discours présidentiels sur la diachronie 1958-2003, avoisinant les 4,4 millions d'occurrences. Notre corpus est constitué de 21 sous-corpus annuels quantitativement équilibrés, qui comptent chacun plus de 150 000 occurrences, comme on peut l'observer sur les figures II-27 et II-28.

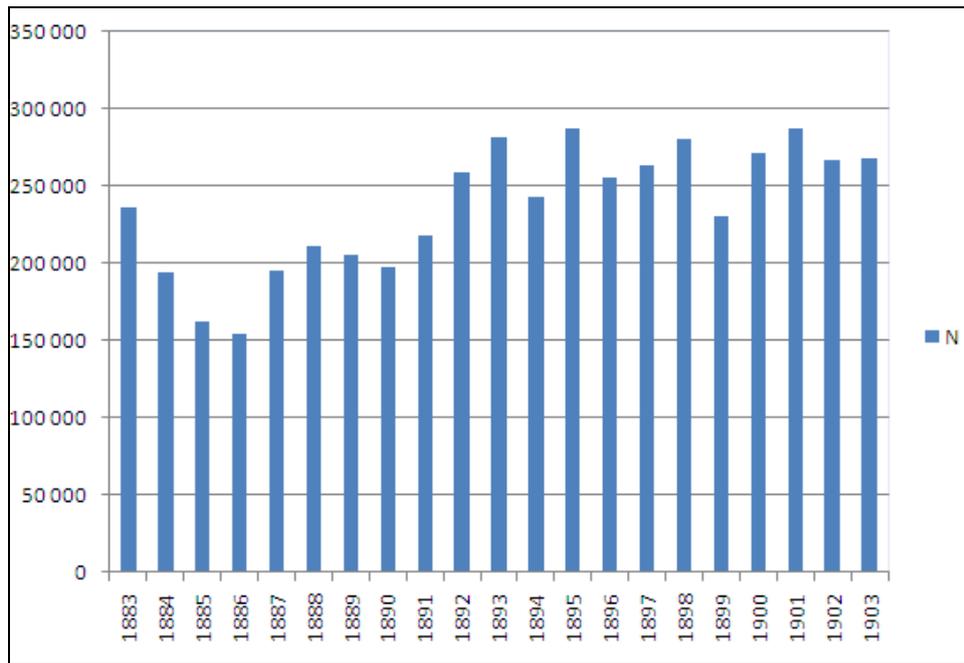


Figure II-27 : Nombre d'occurrences (N) dans les sous-corpus.

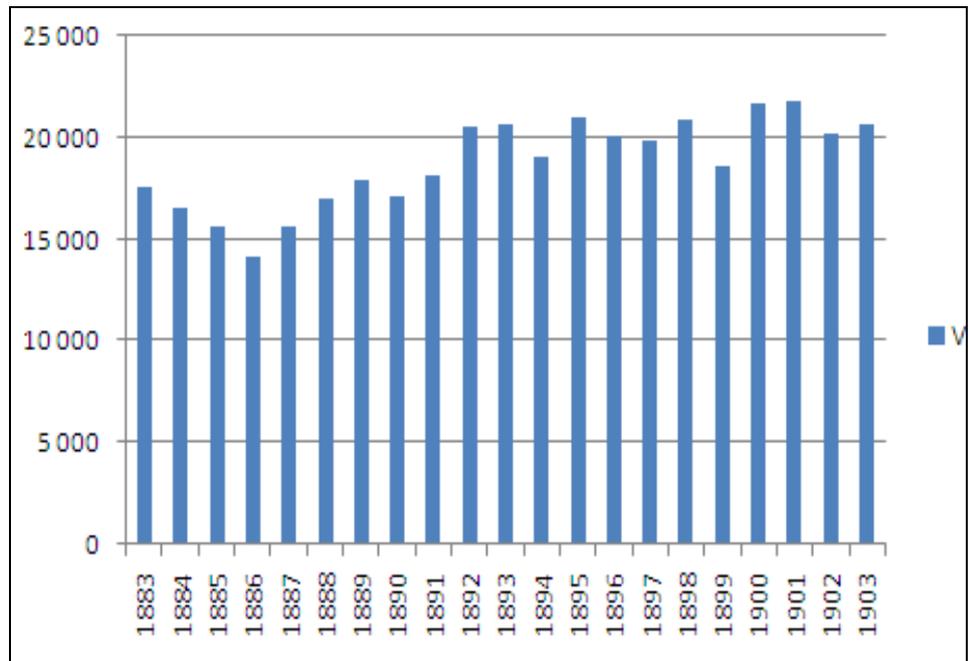


Figure II-28 : Nombres d'items (V) dans les différents sous-corpus.

Notre corpus présente donc les caractéristiques quantitatives garantissant son exploitabilité³⁸ : il est suffisamment long pour être traité par la statistique et ses sous-parties sont suffisamment consistantes et équilibrées pour que soit autorisée une analyse du corpus à partir de sa norme endogène, de laquelle on peut attendre des résultats pertinents et significatifs.

Conclusions partielles

L'acquisition des données textuelles du *Petit Comtois* a représenté, en temps consacré, la majeure partie de notre travail de recherche.

Ce temps de nos recherches a tout d'abord livré les premiers fonds d'une base de données plein-texte établies, normalisées et portables, en vue de leur très prochaine diffusion à la communauté scientifique, de manière à favoriser l'étude d'une source historique de grand intérêt et la comparaison des outils et des traitements lexicométriques. Sous l'angle des résultats, cette étape de nos recherches ne saurait néanmoins être résumée à la simple acquisition d'un vertigineux volume de données textuelles. À notre sens, le processus de numérisation allant de l'étape d'identification des obstacles techniques pesant sur la numérisation de la presse ancienne à la proposition de solutions méthodologiques aptes à faciliter la constitution de données textuelles répondant aux normes scientifiques contemporaines, constitue en soi un résultat. Fruit d'un travail collectif de laboratoire, le protocole expérimental développé dans le cadre de notre travail de recherche propose des améliorations non négligeables pour la constitution d'une base plein-texte de presse répondant aux normes et aux besoins des sciences du texte.

Extrait du premier état de la base plein-texte de PQR3R résultant de ce processus de numérisation, le corpus *Petit Comtois* a été construit par une

³⁸ Cette propriété de notre corpus repose sur « une exigence d'homogénéité » suivant la terminologie de Benzécri (1973).

procédure d'échantillonnage de sorte à constituer un dispositif d'observation des mutations du discours du quotidien sur la période 1883-1903. Un tel corpus dont le volume avoisine les 5 500 000 mots, exige inévitablement de recourir à une analyse assistée par informatique et aux outils de la statistique textuelle (Lebart & Salem, 1994). Dans notre perspective, l'apport de ces outils ne se réduit pas à faciliter grandement notre (re)lecture des quelque 27 000 articles du *Petit Comtois* observés. Nous les convoquons surtout pour leur potentiel heuristique, c'est-à-dire leur capacité à faire émerger du corpus lui-même, analysé à partir de sa propre norme, ses régularités et ses ruptures, ainsi que pour disposer d'un accès privilégié à la textualité. Ce faisant, le corpus *Petit Comtois* se pose lui-même comme un terrain d'expérimentations qui interroge les processus méthodologiques et conceptuels par lesquels peuvent être appréhendés des faits discursifs singuliers et pertinents extraits de la disparité et la masse textuelle d'un très grand corpus.

Chapitre III.

La mise en forme de l'information dans *Le Petit Comtois*

Envisageant l'*aire scripturale* (Peytard, 1975) du journal comme un support signifiant, nous nous proposons dans ce chapitre de décrire la mise en forme de l'information dans le quotidien *Le Petit Comtois*. Postulant que le discours ne saurait être dissocié du cadre où il s'actualise, nous nous attacherons à examiner les dispositifs par lesquels les énoncés du *Petit Comtois* prennent à la fois forme et sens. Nous décrirons ainsi la mise en page de ce quotidien, en tant que celle-ci, résultant de choix éditoriaux propres, confère au quotidien son identité propre.

Après avoir décrit les dispositifs sémiologiques adoptés et leurs incidences sur le parcours de lecture du journal, notre examen de la mise en page du *Petit Comtois* se prolongera par une analyse des formes et des fonctions du péri-texte. En partant de la distinction mise au jour par J.-M. Adam et G. Lugin (2000a), nous analyserons tour à tour le péri-texte du journal, entendu comme l'ensemble des éléments relativement stables du journal, et le péri-texte de l'article, dont la variabilité est à relier à la singularité de l'article qu'il surplombe.

Nous orienterons plus particulièrement notre attention sur le rubriquage, en tant qu'élément péri-textuel du journal, que nous définirons comme un système de cases pré-allouées dans lesquelles sont distribués des contenus nécessairement divers et fluctuants. En raison du découpage du monde qu'il

donne à lire, de la hiérarchisation des informations qu'il assure, le rubriquage sera envisagé comme un mode d'accès privilégié à l'identité du *Petit Comtois*.

Enfin, nous observerons brièvement le péri-texte de l'article tel qu'il se déploie dans l'ensemble du quotidien, pour nous concentrer plus longuement sur ses formes et fonctions dans une rubrique particulière, les *Dépêches de nuit*, à partir d'un sous-corpus de titres substantiel, dépouillé de manière systématique.

3.1. Description matérielle de l'objet imprimé

Le lecteur du XXI^e siècle fréquentant la collection imprimée du *Petit Comtois* ne manquera pas de remarquer la grande sobriété de présentation du quotidien. D'un format de 30 cm par 43 cm en 1883, un numéro de ce journal constitue une aire scripto-visuelle dont la consistance textuelle est extrêmement dense. Un rapide sondage permet d'évaluer le nombre moyen de mots à 1800 par page, sans compter l'espace du rez-de-chaussée. Si certains numéros exceptionnels atteignent les six pages, le quotidien est généralement constitué de quatre pages de 1883 à 1903. Les trois premières pages sont exclusivement rédactionnelles ; la quatrième dédiée à la publicité et aux annonces.

La « mise en page » du *Petit Comtois* est rudimentaire et repose sur l'exploitation de la seule dimension verticale du plan. De 1883 à 1893, texte et péri-texte ne sont justifiés que sur une seule colonne, à l'exception du titre-bandeau du journal. Aucune unité rédactionnelle n'échappe à cette organisation rigoureuse : plus qu'elles ne cohabitent dans l'aire scripturale, elles se succèdent, rivées à la colonne. Pour cette raison, le terme de « mise en page » se révèle donc abusif. La rigidité de l'exploitation du plan a pour corollaire l'unicité du parcours de lecture suggéré : l'unique logique affleurant à la surface des pages du *Petit Comtois* est celle de la progression linéaire du texte, se déployant de ligne en ligne, de colonne en colonne. Loin d'être incité à une lecture (s)élective et flottante, le lecteur, comme s'il était en face d'un orateur de la tribune, est ainsi contraint de suivre la parole de la rédaction, qui s'inscrit

dans une logique temporelle inéluctable. En 1892, *Le Petit Comtois* rompt avec la structuration en quatre colonnes qui le caractérisait depuis le début de sa parution. Il s'organise alors autour de sept colonnes pour revenir à cinq colonnes en 1896. Ces évolutions ne s'accompagnent néanmoins d'aucune mutation de l'exploitation du plan. De 1883 à 1903, la colonne prime, et par elle, la voix du journaliste. C'est avant tout l'état des techniques de composition qui déterminent l'exploitation du plan précédemment décrite et le mode de lecture qu'elle induit. Il est en effet plus commode pour le compositeur typographe de construire des lignes au nombre de caractères restreint afin de les déposer sur la galée. Travailler sur un nombre limité de caractères par ligne facilite la tâche de correction, qui est plus confortable que celles de lignes longues. De fait, l'introduction des monotypes et des linotypes n'a eu que très peu d'incidences sur l'évolution de cette pratique déjà très normalisée.

Dans l'espace des colonnes, comme si le texte se suffisait à lui-même, l'austérité est de mise quant à l'emploi des constituants typographiques. Les visées d'économie d'espace prédominent très nettement sur la recherche de l'esthétisme : les polices utilisées sont très peu diversifiées, signe d'une faible attention apportée aux propriétés esthétiques de l'aire scripturale ; les cas d'utilisation de la graisse des pleins et des déliés sont très raisonnés. Seuls les titres de rubriques et les titres des articles viennent rompre la densité textuelle de la page, tout en restant de taille discrète (entre sept et neuf de corps). Constituant une déchirure dans l'espace de la colonne, ces éléments péri-textuels sont le lieu d'une certaine diversité typographique, qui se justifie par leur fonction dissociative. Les titres de rubriques sont en effet composés en lettres, minuscules ou majuscules, avec des graisses plus ou moins épaisses. La variation de police et de taille de police est un des ressorts d'une hiérarchisation limitée de l'information, complétée par l'utilisation de « filets » signalant une rupture dans l'espace vertical des colonnes. Ces traits rectilignes servent à délimiter les articles. La diversité de ces filets est restreinte : on

distingue des filets droits, des filets anglais (renflés en leur centre) comportant parfois en leur centre des anneaux décoratifs, plus ou moins épais.

lité vient d'ouvrir une baraque où il expose une femme aux formes splendides, aux lèvres de grenade mûre, aux cheveux d'or.
La Belle Panama
annonce l'enseigne. Et au-dessous :
On peut toucher

FAITS DIVERS

En faveur de la crémation

Un cas de léthargie s'est produit à Saint-Gatien-des-Bois, près de Honfleur.

M. Félix Daubenesque, jeune homme de vingt ans, était atteint, il y a vingt jours, d'une fièvre typhoïde.

Le dimanche 18 décembre, il mourait, du moins en apparence, et l'inhumation était fixée au mardi.

A l'heure dite, la cérémonie funèbre eut lieu ; mais, comme le caveau de famille n'était pas prêt, le cercueil fut déposé dans une fosse provisoire.

Le mercredi, on fit l'exhumation.

Toutefois, les ouvriers n'ayant pas encore terminé le caveau, le cercueil fut de nouveau déposé dans l'église, et une personne fut chargée de veiller près de lui.

Voilà que, soudain, un bruit sourd se fait entendre, et ce bruit sort... du cercueil !

Grand effroi, comme on le pense, chez la personne qui veillait !

Le cercueil est ensuite transporté au presbytère.

On retire le couvercle et on aperçoit M. Félix Daubenesque parfaitement en vie, ne comprenant pas pourquoi on l'avait ainsi enveloppé : il sortait de léthargie.

Comme on le voit, ce jeune homme l'a échappé belle ; il doit de la reconnaissance aux ouvriers qui n'avaient pas terminé le caveau en temps utile.

VARIÉTÉS

LES INHUMATIONS PRÉCIPITÉES

La résurrection terrifiante de l'enterré vivant de Saint-Gatien-des-Bois, que nous racontons d'autre part, remet une fois de

Figure III-1 : Exemple de filets droits (1893-01-04), p. 3.

La réponse de M. le président Grévy au pape est un fait accompli; toutefois, on ne connaît pas encore la teneur de ce document. Attendons la publication de cette lettre pour essayer de deviner la pensée vraie qui a donné lieu à cet échange de correspondance.

DÉPÊCHES DE NUIT

Service de notre correspondant spécial.

INFORMATIONS GÉNÉRALES

Paris, le 4 août, 8 h. 45 du soir.

Conseil des ministres.

Dans le conseil des ministres qui a eu lieu ce matin, M. le président de la République a signé un décret nommant M. Foucher de Careil ambassadeur de France en Autriche.

Figure III-2 : Exemple d'un filet anglais (1883-05-08), p.1.

Jou, d'Arrens, ne le voyant pas rentrer le soir, partit à sa recherche; on finit par le retrouver mort sur la neige; il avait succombé à une congestion cérébrale déterminée par le froid.

HAUT-RHIN

Belfort. — Le *Journal de Belfort* publie l'information suivante :

« La *Lanterne* d'hier publie une correspondance de Belfort disant qu'un rapport sur la gestion de M. Jolivet, receveur de l'hospice et du bureau de bienfaisance, a été envoyé par la municipalité au préfet, qui doit le transmettre au parquet.

« Nous n'avons pas à prendre la défense de M. Jolivet; mais nous pouvons signaler cette étrange façon de procéder.

« On fait connaître aux journaux de Paris des renseignements qui devraient rester secrets, au moins tant qu'une poursuite judiciaire n'est pas décidée. »

Note du journal. — L'information qu'on vient de lire n'a pas paru seulement dans la

Figure III-3 : Exemple de filet droit décoré (1893-01-09), p.3.

En 1893, une nouveauté apparaît dans *Le Petit Comtois* : l'illustration. Majoritairement employées en page intérieure¹, ces illustrations se limitent tout d'abord à reproduire le portrait d'hommes politiques ou de personnalités au moment où ceux-ci font l'actualité. Ainsi l'illustration ci-dessous donne à voir l'aspect physique de C. de Lesseps, tandis que se déroule le procès des administrateurs de Panama devant la première chambre de la cour d'appel de Paris.



Figure III-4 : Portrait (1893-01-18), p.2.

Selon J.-F. Tétu (2008), l'apport de cet énoncé iconique consiste à renforcer l'effet de réel visé par le quotidien :

¹ Une seule illustration, reproduisant le portrait du nouveau « M. Le général de France » (1893-03-31), s'ancre en Une en 1893.

Quel est donc l'apport de la gravure ? Simplement un renforcement de cet effet de réel, spécifique de toute image figurative, c'est-à-dire analogique du monde naturel. Que la gravure reproduise les traits d'un individu bien connu ou qu'elle soit une construction élaborée à partir d'un texte descriptif ne change rien de fondamental à la question : elle renforce de toute sa puissance figurative le discours qui la sous-tend (Tétu, 2008 : 64).

L'utilisation d'illustration reproduisant des portraits reste néanmoins très modérée puisque sa fréquence d'emploi n'excède pas, en mars 1893 une fréquence mensuelle de 23%, alors qu'on n'observe aucune « illustration-portrait » de ce type en février 1894 et 1895.

En 1896, on observe une utilisation intensive des gravures pour illustrer certains articles et tout spécialement les faits divers. Cette tendance que nous qualifierions volontiers de « sensationnaliste » tend à s'affirmer en diachronie, ce qu'il conviendrait de confirmer par l'observation d'un échantillon élargi de numéros. En 1901, la totalité des illustrations servent ainsi les faits divers. Il s'agit de présenter le visage de l'accusé (1901-03-11, Chronique régionale, Jura, « Affaire Duparchy », p. 2) ou encore de montrer le visage de « la femme découpée en morceaux » (1901-03-15, p.2). Ainsi, l'article « Meurtre du brigadier Jourdy » (1896-01-14), présentant les verdicts de la cour d'assises du Doubs au sujet d'une affaire locale ayant fait grand bruit, s'agrémente d'un portrait de chacun des onze accusés, que le quotidien introduit en ces termes :

Cette malheureuse affaire ayant passionné et passionnant encore notre paisible population bisontine, chez laquelle les relations de bonne amitié avec les militaires sont de tradition, nous avons cru bien faire en mettant sous les yeux de nos lecteurs les portraits des personnes qui étaient accusées, devant la cour d'assises du Doubs, d'avoir coopéré au meurtre du malheureux brigadier Jourdy (1896-01-14, p. 2).

Toujours en 1896, l'illustration s'expose en Une et prend la forme d'une série de dessins humoristiques à fréquence hebdomadaire, nommée « Heureuse Epoque ».

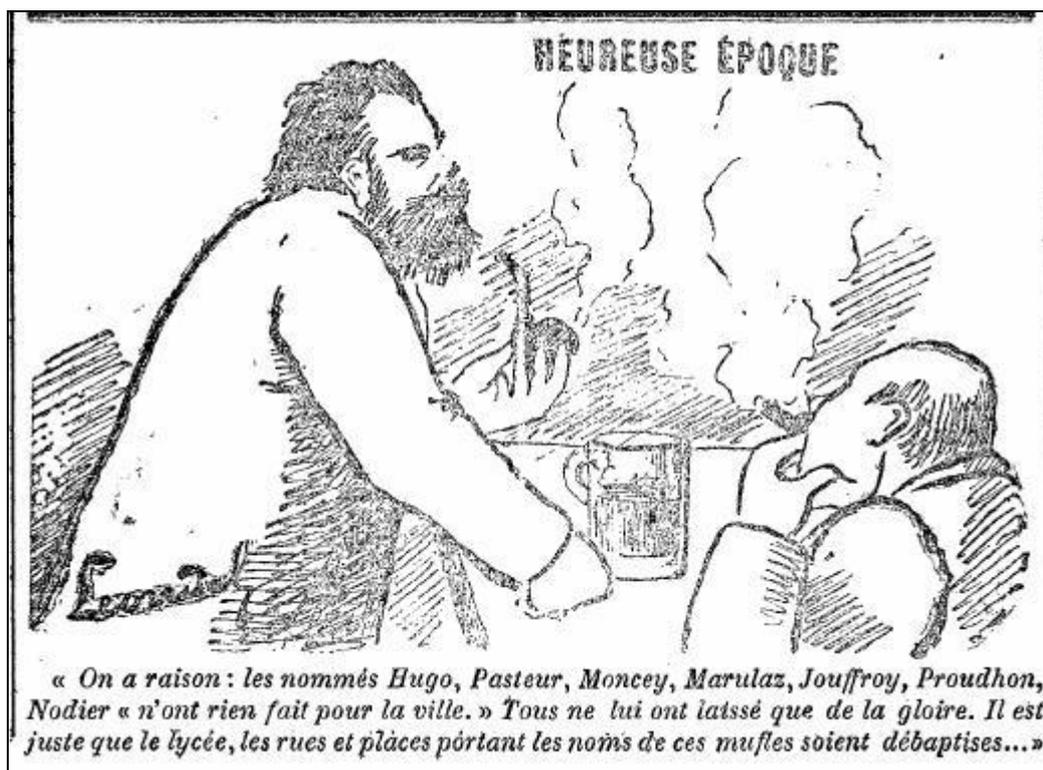


Figure III-5 : « Heureuse Epoque » (1896-01-18), p. 1.

L'illustration ci-dessus participe de manière ironique à une polémique locale ayant fait grand bruit en 1896. Loin de « créer un effet de réel », cette illustration ancre le décor de l'énoncé verbal : une discussion de comptoir entre deux familiers. L'énoncé doit être ici entendu comme une référence à une polémique locale, née des déclarations d'un conseiller municipal Bisontin, M. Ahr. Ce dernier, en pleine séance du conseil municipal, s'était opposé avec virulence à la proposition de construire une statue en l'honneur de Victor Hugo : il jugeait inacceptable que Besançon « eût été obligée de payer les 5,000. fr. [...] pour ce projet, Victor Hugo n'ayant jamais rien fait pour la ville. » (1896-01-22, « Un cas risible »).

En 1897, 50% des gravures² illustrent les prescriptions du *Courrier de la Mode* et complètent les séquences textuelles descriptives. Présentées comme des

² Soit deux des quatre illustrations répertoriées en 1897.

« croquis », ces illustrations visent à montrer les tendances vestimentaires aux lectrices et à expliciter la description.

De 1898 à 1900, ainsi qu'en 1902 et 1903, aucune illustration ne pare la Une ou les pages intérieures, reflétant la parcimonie avec laquelle le quotidien utilise les énoncés iconiques et privilégie la sobriété. Si l'austérité est de mise dans les trois premières pages du *Petit Comtois*, la dernière page présente des caractéristiques tout autres. Réservée aux unités non-rédactionnelles, c'est-à-dire aux avis, à la publicité et aux annonces, la dernière page se caractérise par une mise en page plus souple, où l'horizontalité est exploitée. Certaines annonces se déploient ainsi sur une largeur de deux colonnes, et se caractérisent par des tailles et des polices plus diversifiées. Ainsi, sur la figure III-6, reproduisant une page d'annonces et de publicités, on observera des encarts justifiés sur deux et trois colonnes, voire sur la totalité de la largeur de la page, comme on peut le constater dans l'extrême partie basse de la figure. Contrairement aux pages intérieures, la figure III-6 se caractérise par un jeu sur les variables typographiques (majuscules, gras, caractères graphiques décoratifs, accolades) et un fort encrage, l'ensemble des éléments convergeant vers la captation de l'attention du lecteur.

UTILISATION RATIONNELLE DES COURS D'EAU



TURBINES

AMÉRICAINES

Rendement garanti au frein de 80 0/0

TURBINES RATIONNELLES
pour toutes chutes et tous débits

ROUES hydrauliques — ROUES Pelton

Montage d'arbres, Transmissions, Paliers, Cloches, Poulies
APPAREILS DE LEVAGE

DOUGE FRÈRES

Ingénieurs-Constructeurs à Besançon

SPÉCIFIQUE LERCK *Reçu par les Médecins sur demande*

PLAIES PURULENTES et eczémas NATURELS ECZÉMAS, DARTRES

HÉMORROÏDES
Quatre par jour
Suppositoires
Lerck

MAUX DE JAMBES-VARICES-ULCÈRES
Lerck

Exposition Universelle Paris 1900
GRAND PRIX
Sole produit importé en France par Docteur Lerck

Dépôt pharmacies : Gray, Russ; Vesoul, Moynier, Bideux, Veyssel, Belfort, Kroll, Delsart, Petit-Jean; Danjoutin, Blanc; Montbéliard, Bernard; Mook; Baume-les-Dames, Falvre; Besançon, Pichon, Jacques, Robessey, Gauthier, Birtsch; Pontarlier, Denis. 88352



Royal Windsor

LE CÉLÈBRE
RÉGÉNÉRATEUR DES CHEVEUX

AVEZ-VOUS DES CHEVEUX GRIS?
AVEZ-VOUS DES PELLICULES?
VOS CHEVEUX SONT-ILS FAIBLES OU TOMBENT-ILS?

SI OUI,
Employez le ROYAL WINDSOR, qui rend aux Cheveux gris la couleur et la beauté naturelles de la jeunesse. Il arrête la chute des Cheveux et fait disparaître les Pellicules. Il est le SEUL Régénérateur des Cheveux médaillé. Résultats Inespérés. Vente toujours croissante.
Exiger sur les flacons les mots Royal Windsor. Se trouve chez les Coiffeurs-Parfumeurs en Grosse et Petite Rosace.
ENTREPOT : 28, rue d'Enghien, PARIS
Envoi franco sur demande du Prospectus contenant détails et attestations

Figure III-7 : Publicités avec illustration (1903-03-29), p. 4.

3.2. Le rubriquage du *Petit Comtois* (1883-1903)

3.2.1. Approche générale du rubriquage

3.2.1.1. Retour sur l'étymologie du mot « rubrique »

Un bref retour sur l'étymologie et le parcours de la notion de rubrique se justifie. Un examen de ses occurrences au XIX^e siècle *via* le Trésor de la Langue Française Informatisé (TLFI) montre tout d'abord qu'un flou sémantique

caractérise ce terme. La rubrique est tout d'abord définie comme un « titre d'un livre, d'un chapitre, écrit autrefois en rouge », suggérant une acception de la rubrique comme un élément péri-textuel établissant une rupture typographique avec l'espace blanc de la page, renvoyant à l'étymologie même du mot que rappelle le TLFi : « de *rubrica*, terre rouge ». Dans la première moitié du XIX^e siècle, la pratique journalistique emprunte le terme aux champs liturgique et littéraire, pour l'assimiler à un titre indiquant la provenance et la date des nouvelles qu'il surplombe. L'énoncé suivant, extrait d'*Albert Savarus*, d'Honoré de Balzac, atteste de ce sens prévalant jusqu'à la Monarchie de Juillet :

Elle lui tendit un numéro de gazette qu'elle avait dans la poche de son tablier, lui indiquant l'article suivant sous la rubrique de Florence, au 25 mai³.

Ainsi entendue, la rubrique a principalement pour fonction de construire une rupture dans l'aire scripturale du quotidien et d'indiquer l'horizon spatio-temporel de la nouvelle. Avec l'accroissement et la diversification du flux de nouvelles liés au développement d'un journalisme d'information, s'impose pour les grands quotidiens nationaux la nécessité d'organiser la matière informationnelle dans l'espace du journal. Le témoignage d'Hippolyte de Villemessant est à cet égard tout à fait révélateur :

[...] mon premier besoin, quand j'eus fondé mon journal, fut de le diviser par cadres. Je rêvais pour cela quelque chose d'aussi clair, d'aussi simple comme organisation que ce que je voyais dans le commerce, dans les grands magasins. [...] De même pour un journal : il est indispensable que le lecteur sache où il trouve la Causerie, les Echos de Paris, ceux de la Chambre, les articles Variétés, les Tribunaux, le rayon des faits divers, celui des théâtres, etc., et ne soit pas obligé d'errer à la pêche de tel article, comme cela arrive dans maints journaux, qui jettent pêle-mêle toutes les nouvelles importantes dans leur première page et n'ont plus rien à mettre dans les trois autres (Villemessant, de, H., 1873 : 40).

Le témoignage de ce praticien, directeur du *Figaro*, fait état de deux besoins essentiels pour la rédaction, qui préfigurent les fonctions de la rubrique dans les années 1860. Tout d'abord, une fonction d'organisation et de gestion du flux d'informations des nouvelles, qui renvoie à l'organisation et à la division mêmes

³ Balzac, H. (de), 1842, *Albert Savarus*, p. 124, cité dans le TLFi, en ligne, consulté le 10.05.2008.

du travail d'écriture journalistique. Le rubriquage peut en cela être défini comme une configuration du journal en cases allouées à des unités rédactionnelles stables. Par la structuration du quotidien en rubriques récurrentes au fil des numéros et à l'emplacement stabilisé, on vise également une fonction de ritualisation de la lecture du journal. Selon l'historien Louis Guéry (1997), une telle organisation, appelée « cuisine » dans le jargon journalistique de la fin du XIX^e siècle, est pour la première fois mise en œuvre dans le quotidien *La liberté* du 1^{er} avril 1866.

3.2.1.2. Le point de vue des théoriciens sur la rubrique

Il est stimulant de confronter la description de la rubrique par les praticiens du journalisme contemporains de sa création au point de vue des théoriciens⁴. Nous introduirons notre propos en relevant l'écho de l'analyse du fonctionnement du rubriquage donnée par T. Herman et G. Lugin (1999a), au témoignage d'H. de Villemessant :

Pour être explicite, le journal présente un système de classement composite, pluri-sémiotique, en certains points comparable à celui d'un supermarché. Ayant la même double exigence de faire vendre et de composer avec des objets très divers, le supermarché nécessite non seulement un classement précis et efficace, mais également attractif. [...] Dans l'optique d'un classement efficace, le supermarché fait appel à un arsenal de systèmes sémiotiques allant de la pancarte à l'étiquette du produit et destinés à permettre au consommateur potentiel de se retrouver au mieux dans l'espace du magasin. Enfin dans la perspective d'une certaine attractivité, on sait bien que le marketing a depuis longtemps développé des stratégies visant à augmenter les ventes, allant de l'étude de l'éclairage au travail sur les passages obligés du consommateur en passant par le rangement des produits, à la hauteur des yeux par exemple [...] (Herman, Lugin, 1999a : 17).

Structurer la masse initialement diffuse et hétérogène de la matière informationnelle

Organiser la masse initialement diffuse et hétérogène de la matière informationnelle à partir d'espaces pré-ordonnés dans lesquels sont

⁴ Exposer le point de vue des théoriciens de la mise en forme du quotidien implique nécessairement de faire état de travaux datant du XX^e et du XXI^e siècle, puisqu'aucune théorisation de l'objet journal n'est existante avant le début du XX^e siècle.

réparties les nouvelles, telle est bien la fonction première mise en avant par les auteurs de *Formes et fonctions des rubriques dans les quotidiens romands*, de même que le spécifiait H. de Villemessant. Tout comme l'entendait ce dernier, T. Herman et G. Lugin établissent que cette première fonction a pour corollaire de sous-tendre, sur le plan de la réception, une ritualisation de la lecture. Le rubriquage, en tant que configuration stable du quotidien survivant à la périodicité quotidienne du journal, permet de fait une lecture sélective de l'information. Tout comme dans une boutique que le consommateur parcourt en fonction des rayons qui l'intéressent, un lecteur construit son propre parcours de lecture en fonction des rubriques sollicitant son intérêt.

Rubriquer ou comment hiérarchiser l'information

Loin de réduire le rubriquage à un système de classification, T. Herman et G. Lugin introduisent au gré de leur comparaison du rubriquage avec les rayonnages de supermarché les notions de stratégie et de hiérarchisation de l'information. Tout comme un supermarché est régi par des principes visant à mettre en valeur des produits par des moyens divers (signalétique, choix de leur emplacement, etc.), le quotidien est construit en fonction d'une stratégie de hiérarchisation de l'information. Ce sont notamment l'emplacement topographique dans l'espace du journal et le jeu sur les variables typographiques qui concourent à une telle mise en valeur de l'information, comme le rappelle P. Charaudeau :

Dans la presse écrite, la nouvelle est présentée selon un certain nombre de critères de construction de l'espace rédactionnel et iconique, censés correspondre au degré d'importance qu'on veut lui attribuer : l'emplacement (à la Une ou en page intérieure, en haut ou en bas de la page, en sur-titre, titre ou sous-titre) ; la typographie (taille et corps des caractères d'imprimerie dans la titraille) ; la quantité de surface rédactionnelle (et iconique) comparée à celle des autres nouvelles, en pourcentage (Charaudeau, 1997 : 161).

Au même titre que la typographie, l'emplacement et la surface occupés par la nouvelle, le rubriquage est un ressort de la hiérarchisation de l'information. Par l'ordre et l'intitulé des rubriques, la rédaction du journal attribue une valeur à une information, en lui accordant une plus ou moins grande exposition par

rapport aux autres nouvelles. Parce qu'il donne à lire les choix de la rédaction, le rubriquage constitue ainsi un terrain d'observation privilégié de l'identité d'un journal. Le rubriquage témoigne notamment de la façon dont le journal traite l'information, selon que celle-ci devient dans ses colonnes *événement* ou nouvelle de seconde importance. Plus : indépendamment de son rapport au texte qu'il encadre, le rubriquage, par le seul système d'intitulé de ses titres de rubrique, est un lieu d'accès privilégié à l'identité du quotidien. Car, au travers de ses catégories, le rubriquage propose et véhicule une lecture du monde intimement liée au positionnement socio-discursif du journal. C'est là une des hypothèses fortes de M. Mouillaud et J.-F. Tétu :

La « mise en rubrique » constitue les « références » selon lesquelles les contenus peuvent se distribuer ; ces références n'existent que dans la mesure où le journal lui-même les institue, les constitue comme références. [...] La mise en rubrique ne saurait être la classification du monde « réel » (Mouillaud, Tétu, 1989 : 64).

Ce sont précisément ces attitudes du *Petit Comtois*, que nous souhaitons maintenant examiner, en entreprenant la construction d'un répertoire des rubriques du quotidien durant ses vingt premières années de parution, en observant leur stabilité en terme de fréquence, de surface, et d'emplacement dans l'espace du journal.

3.2.2. Approche quantitative des rubriques du *Petit Comtois*

Pour procéder à une étude quantitative des rubriques déployées dans le quotidien, nous nous sommes librement inspirée de la méthodologie élaborée par l'historien J. Kayser (1963) en vue de l'analyse morphologique du journal. À cet effet, nous avons observé les 630 numéros du *Petit Comtois* composant notre corpus, selon la méthodologie que nous décrivons ici.

3.2.2.1. Méthodologie d'analyse quantitative des rubriques

La première phase de cette classe d'analyse a consisté à identifier les rubriques du *Petit Comtois* par un examen de la mise en page du quotidien au

fil de sa parution. Notre attention s'est plus particulièrement orientée vers la mise au jour d'un éventuel système de variables visuelles et linguistiques signalant les titres de rubriques. Pendant la première année de parution, ces variables se combinent de façon fluctuante et ne revêtent aucune valeur systématique. Des filets de différentes tailles surplombent le titre des rubriques, portant d'un numéro sur l'autre sur la moitié (cf. figure III-8) ou la totalité de la largeur de la colonne (cf. figure III-9). Certains titres de sous-rubriques (cf. figure III-10) sont surmontés du même filet que celui qui définit rubriques. La majuscule est alors le seul indice explicite permettant de distinguer le titre rubrique du titre de sous-rubrique :

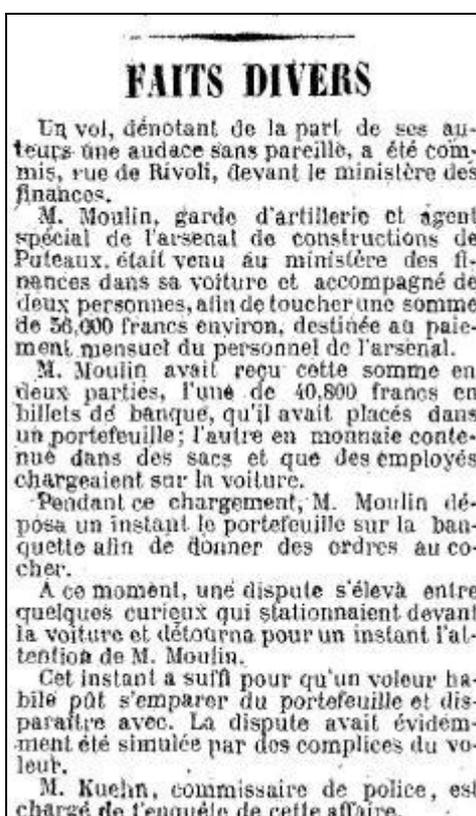


Figure III-8 : Titre de rubrique en gras et casse majuscule (1883-08-05), p. 4.



Figure III-9 : Titre de rubrique en casse majuscule et gras (1883-08-01), p.1.

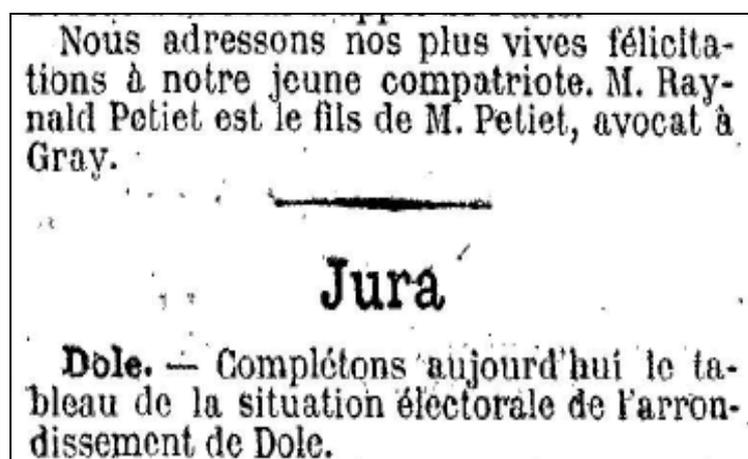


Figure III-10 : Titre de sous-rubrique en gras et casse minuscule (1883-08-05), p. 3.

Les combinaisons de variables visuelles marquant la différenciation des espaces se stabilisent à partir de la deuxième année de parution du *Petit Comtois*. Le titre de rubrique est alors précédé par un épais liseré horizontal encadré de deux traits droits de la largeur d'une colonne (cf. figure III-11). Concernant la signalétique typographique proprement dite, un jeu de variations sur la taille et la valeur des caractères est utilisé pour signaler un titre de rubrique : chaque rubrique se voit attribuée une police particulière, en petites capitales et tailles de corps de deux à trois fois supérieures à celles du texte.

Les figures ci-dessous illustrent de telles combinaisons de signalétiques typographiques :



Figure III-11 : Jeu de variations typographiques (1883-08-06), p. 4.



Figure III-12 : Double filet signalant une rubrique (1886-01-03), p.2.

Progressivement, les titres de rubrique se différencient plus nettement des titres de sous-rubrique. Ainsi, en 1896, chaque titre de sous-rubrique est signalé pour sa part par un liseré fin enflé, comme ci-dessous :



Figure III-13 : Filet enflé surmontant un titre de sous-rubrique (1896-01-10), p.3.

D'un point de vue linguistique, l'intitulé de rubrique est généralement formé d'un syntagme nominal sans déterminant ni prédicat, comme on peut le constater sur les figures ci-dessus. Nos observations rejoignent en cela la

définition des caractéristiques formelles du titre rubrique proposée par M. Mouillaud et J.-F. Tétu selon lesquels « Les titres-rubriques sont des noms sans déterminant et dépourvus de prédicat » (1989 : 118). Cela étant dit, il convient de remarquer qu'à partir de 1892, *Le Petit Comtois* utilise quelques intitulés de rubriques pourvus de déterminant, comme *L'esprit du jour*. Ce type de construction reste cependant très minoritaire.

Au-delà de l'observation de ces variables visuelles, notre repérage des rubriques s'est bien évidemment appuyé sur l'examen de leur récurrence. De fait, nous avons été amenée à distinguer des rubriques permanentes et des rubriques périodiques, récurrentes pendant un nombre limité de numéros, liées à une affaire ou à un événement donné. Une fois ce système de variables mis au jour, nous avons constitué un répertoire des rubriques identifiées. Trois objectifs ont déterminé les modalités de construction de ce répertoire. Nous avons tout d'abord voulu disposer d'informations sur la fréquence de chaque rubrique en pourcentage, afin de pouvoir examiner à la fois la structure synchronique du quotidien pour chaque année et de pouvoir comparer ces fréquences en diachronie. À cet effet, nous avons dépouillé de façon systématique tous les numéros d'un mois de parution du *Petit Comtois*, - celui-là même retenu comme échantillon pour notre corpus-, et nous avons répertorié toutes les rubriques publiées dans chaque numéro. Nous avons ensuite comptabilisé le nombre d'occurrences de chaque rubrique dans le mois et avons transformé ces valeurs absolues en pourcentage, de sorte à obtenir une fréquence mensuelle de parution de chaque rubrique. Dans le même temps, nous avons axé notre approche quantitative vers l'observation de la superficie de chaque rubrique, partant du postulat selon lequel :

L'importance de la surface vient de ce qu'elle détermine l'importance des textes qui la composent. À elle seule, une surface ne dit rien [...] seulement un rapport entre cette surface-ci et celles qui l'entourent. Mais dans la mesure où l'unité référentielle est perçue par cette surface, elle se trouve ipso facto ou plus exactement, elle est rendue comparable à d'autres, sous le seul rapport de l'importance [...]. Sa taille n'est pas une plus-value mais la condition même de son existence (Mouillaud & Tétu, 1989 : 68).

L'organisation du quotidien du XIX^e siècle présente une organisation qui ne tient pas compte d'autres propriétés du plan que celle de la verticalité. Cela nous a amenée à prendre pour unité de quantification la colonne. Le choix de cette unité d'évaluation nous semble se justifier par l'objectif de notre démarche, visant à obtenir des ordres de grandeur afin d'établir des comparaisons acceptables, et ne nécessitant donc pas de recourir à une unité plus fine telle que le nombre de mots ou de lignes, ou encore le millimètre. Grâce à une évaluation réalisée à partir des échelons suivants, les résultats obtenus se révèlent particulièrement fiables :

1. moins d'un quart de colonne
2. 1 quart de colonne
3. Entre $\frac{1}{4}$ et $\frac{1}{2}$ colonne
4. Entre $\frac{1}{2}$ colonne et $\frac{3}{4}$ de colonne
5. Entre $\frac{3}{4}$ colonne et 1 colonne
6. Entre une colonne et une colonne $\frac{1}{4}$, etc.

Parallèlement à ces deux premiers axes d'observation, l'emplacement de chaque rubrique, année par année, a été répertorié afin d'analyser la valeur accordée à une rubrique par sa topographie dans le quotidien et de vérifier la stabilité⁵ de la maquette. Ce niveau d'observation a été déterminé par l'option théorique selon laquelle l'emplacement topographique d'une rubrique lui confère une valeur plus ou moins grande. Ce postulat doit néanmoins être précisé : l'emplacement topographique n'est qu'un des nombreux ressorts de la mise en valeur de l'information, parmi lesquels la typographie, l'intitulé du titre de rubrique, etc. La nécessaire combinaison de ces éléments est d'ailleurs à l'origine de la « formule de mise en valeur » d'une unité rédactionnelle que J. Kayser (1963) théorise comme résultant de la somme de l'emplacement, du titrage et de la présentation⁶. Il ne s'agira donc pas d'évaluer la valeur directe

⁵ Notre évaluation se fonde sur l'observation d'une tendance de régularité qui n'implique pas l'absence d'altération ponctuelle.

⁶ « Ainsi, la mise en valeur d'un texte résulte de la somme de trois éléments qu'on peut objectivement isoler et mesurer_ l'emplacement (E), le titrage (T) et la présentation (P) : $Mv) E + T + P$ » (Kayser, 1963 : 152).

d'une rubrique en fonction de son emplacement topographique dans le quotidien, mais de la mettre en perspective avec un ensemble d'observations. Nous avons enfin observé l'évolution en diachronie des articles non-rubriqués, mais simplement titrés, qui représentent une surface relativement importante de la deuxième page du quotidien. On trouvera en annexe, pour chaque année, les tableaux récapitulatifs de nos observations sur la fréquence, l'emplacement et la surface occupée des rubriques.

3.2.2.2. La « cuisine » du *Petit Comtois*

De l'examen des tableaux disponibles en annexe, il ressort en premier lieu que la maquette du quotidien se révèle très stable sur la période 1883-1903. La stabilité des emplacements topographiques alloués à chaque case renvoie à la fois aux conditions matérielles et techniques rudimentaires de production du journal, mais aussi à sa ligne éditoriale.

À la stabilité de la maquette du quotidien, correspond un perpétuel remaniement de la « cuisine » du *Petit Comtois* : par le jeu de l'émergence et de la disparition de nombreuses rubriques, par des variations d'intitulés, le rubricage, de 1883 à 1903, ne cesse en effet de se modifier. Aussi décrivons-nous brièvement l'évolution du visage du *Petit Comtois*.

De 1883 à 1886

À la parution de son premier numéro, seize rubriques⁷ structurent *Le Petit Comtois* et définissent son identité. Confirmant son statut de titre de journal régional, la rubrique *Chronique régionale* occupe une place centrale dans le quotidien, tant par le nombre de colonnes qui lui est attribué (cf. figure III-14) que par son emplacement dans les pages intérieures du quotidien, ainsi que par sa fréquence (figure III-15).

⁷ Sont exclues les rubriques non-rédactionnelles : publicité, bulletin météorologique, bulletin financier, bulletin commercial, annonces, Spectacles et concerts, Halles et Marchés.

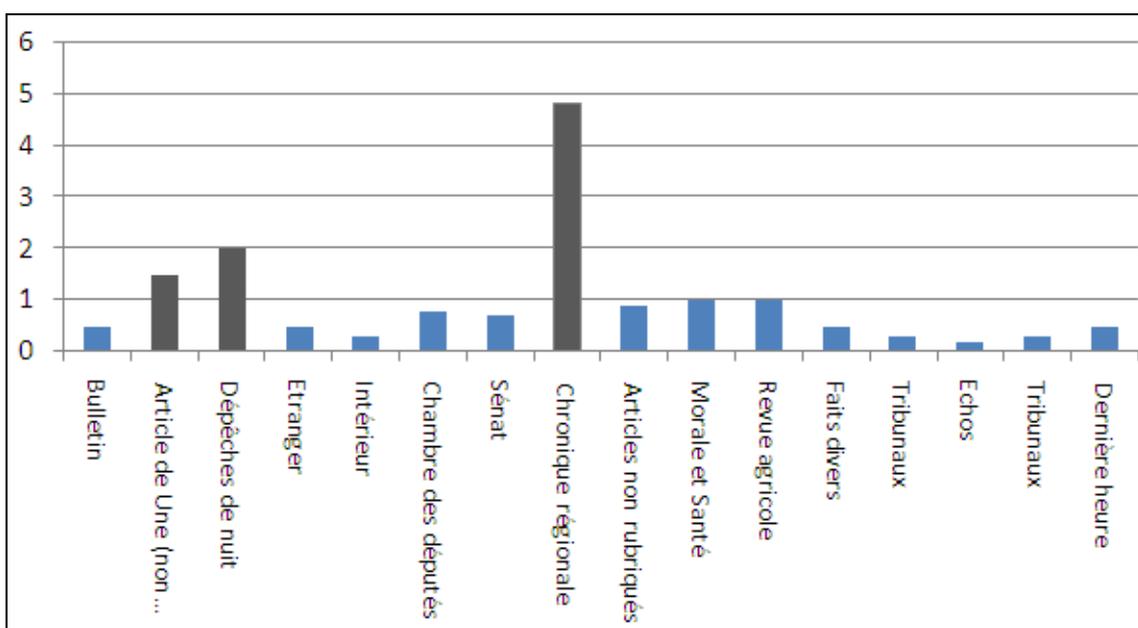


Figure III-14 : Répartition de la surface du quotidien en 1883 (en nombre de colonnes).

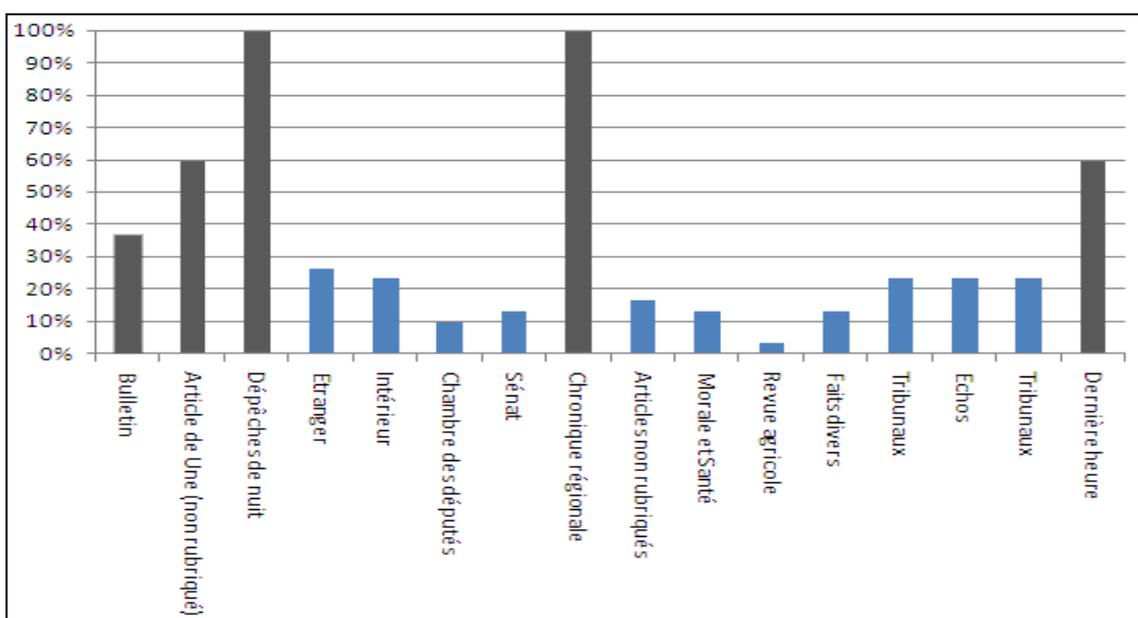


Figure III-15 : Fréquence de parution des rubriques en 1883 (en pourcentage).

Les *Dépêches de nuit* constituent l'autre bloc essentiel du quotidien, par le nombre de colonnes qu'il occupe et par sa fréquence quotidienne, mais aussi par l'emplacement de choix qui lui est réservé, c'est-à-dire la Une,

« devanture » du journal pour reprendre l'expression de J. Kayser (1963 : 85). Grâce à cet emplacement stratégique dans l'architecture du journal, la rubrique est d'emblée exposée au regard du lecteur. Il nous semble très intéressant d'observer que les dépêches et leur style télégraphique, leur insistance sur le factuel, sont précédées d'articles à haute prise en charge énonciative, où se donne à lire la voix de la rédaction. Ces articles occupent l'emplacement le plus important de la Une, appelé dans le jargon de l'imprimerie l'*ouverture*. Situés dans la partie gauche de la tribune, ils « ouvrent » le discours du quotidien. C'est ainsi que de 1883 à 1885, la rubrique du *Bulletin* donne, selon les termes de Jules Gros, le « mouvement essentiel du jour »⁸. Comme on peut le constater à la lecture des exemples suivants, la rédaction attire l'attention du lecteur sur une sélection d'événements présentés en dépêches et les commente, tout en exerçant une fonction d'appel. Imposer une lecture hiérarchisée par la rédaction de l'événement et inciter le lecteur à partager cette hiérarchie-là : telle est bien la fonction du *Bulletin*.

Ch.III (01), Bulletin, (1883-08-03)

Besançon, le 2 août, 11 h, du soir.

Le fait important de la journée est la clôture de la session parlementaire. MM. les députés et sénateurs vont se répandre dans les départements, où l'agitation électorale va prendre vraisemblablement une intensité plus considérable. Il sera curieux d'observer, dans cette lutte, si le mouvement d'opinion qui s'affirme en ce moment contre le cumul des mandats électifs ira en augmentant.

Le cumul des mandats dont nous parlons n'a rien de commun avec le cumul que la Chambre aurait voulu interdire entre les fonctions judiciaires et les mandats électifs. [...] Nous appelons à cet égard l'attention de nos lecteurs sur la partie de nos informations télégraphiques, où il est traité de l'impression produite sur le public par la séance de mercredi dernier.

[...] Puis, M. le général Thibaudin, formellement interpellé par M. Madier de Montjeau, est monté à la tribune, et en quelques mots a coupé court aux tentatives des tacticiens de l'extrême gauche qui ne lui pardonneront pas aisément, sans doute, leur désappointement.

On pourra consulter à cet égard, avec fruit, une courte correspondance de Paris que nous croyons utile de signaler au lecteur, et que nous insérons dans nos informations générales.

⁸ Cf. Lettre du 22 janvier 1884 de Jules Gros à un actionnaire du *Petit Comtois*.

Ch.III (02), Bulletin, (188-08-21)

L'étranger ne nous apporte aujourd'hui aucun contingent de nouvelles bien saillantes. Disons cependant qu'il résulte de nos dépêches que les rapports de la France avec l'Angleterre seraient moins tendus en ce moment qu'il y a quelques jours, par suite de ce qui se passe à Madagascar. [...]

A l'intérieur, les scrutins de ballottage ont achevé l'écrasement des adversaires de la République, qui n'ont plus la majorité que dans neuf assemblées départementales, sur 90. **C'est le seul point de notre politique intérieure sur lequel il nous paraisse utile d'insister aujourd'hui.**

Le *Bulletin* est publié en alternance avec un article non rubriqué qui occupe le même emplacement topographique. Généralement non signé, cet article prend position sur un fait d'actualité en engageant la responsabilité du journal. Véritable *vitrine idéologique* (Herman, Juper, 2000 :142) du *Petit Comtois*, cet article porte les traits micro-linguistiques d'un article de commentaire où les marques énonciatives, plutôt que d'être gommées pour feindre l'objectivité, sont pleinement assumées. À titre d'illustration, il est intéressant d'observer l'article suivant. La teneur idéologique de son propos est clairement explicitée : il s'agit pour la rédaction de réagir à une « doctrine discutable » tenue par le sénateur A. Oudet et par là même, d'exposer la conception du *Petit Comtois* de l'idée républicaine au travers d'une description orientée des clivages politiques entre radicaux et opportunistes.

Ch.III (03), « Deux politiques », (1883-08-08)

Nous nous sommes promis de ne pas descendre dans l'arène électorale, et nous tiendrons notre promesse. Mais nous n'avons pas pris l'engagement de passer sous silence des doctrines qui nous paraîtraient discutables [...]:

Dans la réunion de dimanche dernier, au théâtre de notre ville, l'honorable M. Oudet, sénateur, s'est exprimé ainsi : « Il existe une doctrine qui veut un gouvernement fort, et qui tend à affaiblir au profit du pouvoir central les ressorts sur lesquels repose l'autonomie de l'individu. »

Nous reconnaissons que la phrase est sonore, mais nous cherchons vainement ce qu'elle peut signifier. Qu'est-ce, en effet, que l'autonomie de l'individu? Est-ce que les citoyens ne jouissent pas aujourd'hui de toutes les libertés politiques qu'ils ont le droit de revendiquer ?

[...] Nous croyons donc que l'honorable M. Oudet aurait pu, sans inconvénient, retrancher de son discours ce qu'il a appelé « l'autonomie de l'individu, » s'il faut en croire, du moins, le compte-rendu officiel, en quelque sorte, de la séance. [...]

Depuis quelques années, en effet, deux écoles de républicains sont en présence : ceux qui parlent de gouvernement fort, et qu'on traite aussitôt d'autoritaires, parce qu'ils ne veulent pas permettre que, sous prétexte de république, on fasse à

la France une situation humiliée en Europe, et ceux qui, par conviction où par calcul électoral, crient au régime autoritaire et même à la tyrannie, dès qu'un citoyen se permet de penser et de dire qu'à force de diminuer le pouvoir exécutif, on pourrait bien compromettre la République et la liberté.

Nous appartenons à la première de ces deux écoles ; M. Oudet appartient ou semble appartenir à la seconde. [...]

Qui a tort, qui a raison, de nous deux? L'avenir le dira. En attendant, nous professons, nous, au risque d'effaroucher les intérêts électoraux, que la République ne doit pas être le gouvernement des ganaches, et nous nourrissons le secret espoir que le peuple pense comme nous, parce qu'il prise avant tout la virilité.

En 1884, des *Dépêches de la journée* apparaissent en Une au côté des *Dépêches de nuit*. Si la rédaction regroupe dans les *Dépêches de nuit* les informations parvenues à la rédaction la veille au soir de la parution, la case *Dépêches de la Journée*, introduit pour sa part les informations réceptionnées la veille au matin, comme on peut le constater sur la figure ci-dessous :



Figure III-16 : « *Dépêches de la journée* » (1885-01-06), p.1.

La Une est donc l'emplacement privilégié d'une actualité politique nationale et internationale fraîchement rapportée, dont le traitement est prolongé en haut de la deuxième page par la rubrique *Au jour le jour*. Cette rubrique, composée

d'une succession de courts énoncés traitant de nouvelles internationales, est assimilable du point de vue générique à une revue de presse⁹.

Ch.III (04), « Au jour le jour », (1884-01-28)

La *Gazette de Moscou* dit que les renseignements publiés par les journaux étrangers, touchant les réformes gouvernementales que l'on préparerait en Russie, sont dénués de fondement. Le journal russe ajoute qu'il faut que l'on décide si la Russie veut conserver son Eglise et ses lois, ou si l'on doit assister à l'avènement de quelque chose de nouveau et d'étranger. [...]

Le *Berliner Tagblatt* s'est occupé de la brochure de M. Maas, qui propose la neutralisation de l'Alsace-Lorraine, et il émet l'avis qu'une nouvelle guerre victorieuse contre la France, où les Alsaciens-Lorrains combattraient côte à côte avec les Allemands, ferait plus pour la germanisation des provinces annexées que ne feraient cinquante années d'administration pacifique. On n'est pas plus aimable que ces bons Allemands ! Gageons qu'ils n'en continueront pas moins à dire que la France leur cherche des querelles.

Les pages intérieures : du régional au divertissement

Outre quelques articles non-rubriqués en deuxième page, les pages intérieures du quotidien hébergent principalement, comme nous l'avons signalé précédemment, la *Chronique régionale*. Cette rubrique présente la particularité d'être subdivisée sur le plan géographique au fil des numéros : ainsi, les informations sont classées suivant le département dans lequel les événements se déroulent. Le même ordre se répète inmanquablement, symbolisant un ensemble de cercles concentriques dont le cœur serait la rédaction du quotidien installée à Besançon : Doubs, Jura, Haute-Saône, Haut-Rhin, Suisse. Suivant le même principe de proximité, la sous-rubrique « Doubs », qui ouvre la rubrique, débute par une section « Besançon ».

⁹ À noter que la « Revue de presse » en tant que rubrique n'apparaît dans le *Petit Comtois* qu'en 1932 pour devenir quotidienne en 1935.

CHRONIQUE REGIONALE

Doubs

BESANÇON.

On nous annonce que la Société d'émulation, dans sa dernière séance, a émis le vœu que le nom de Chifflet fut rendu à la rue Neuve-Saint-Pierre, et que la nouvelle rue construite sur les terrains Fachard portât le nom de Dunod.

Il paraît même que ce vœu a été transmis au conseil municipal, et que celui-ci a déjà délibéré à ce sujet.

Le bal de la COMTOISE.

Comme nous le pressentions hier, le bal de la société la Comtoise a conservé son animation et sa gaieté jusqu'à une heure très avancée. Ce n'est qu'à six heures du matin que les plus intrépides danseuses ont quitté la place. Encore ont-elles usé de leur aimable influence pour décider la commission à organiser très prochainement un nouveau bal, dont le succès égalerait certainement celui de samedi.

Il se forme actuellement à Besançon un cercle nommé la Bazoche bisentine, et dans lequel seront admis les avocats, notaires et avoués, les clercs de notaires et d'avoués et les jeunes gens faisant partie de l'enregistrement, des contributions directes et du bureau des hypothèques.

Ce cercle compte actuellement soixante-dix membres, et, bien que l'organisation n'en soit pas encore complète et que l'autorisation n'ait pas encore été délivrée par M. le préfet du Doubs, les demandes affluent tous les jours.

La salle choisie par les membres du cercle est celle du café Foissotte, qui servait autrefois de lieu de réunion aux membres de la loge maçonnique.

Le brouillard avait rendu hier le pavé très glissant, ce qui a causé plusieurs accidents, heureusement sans importance.

A quatre heures, un cheval de fiacre s'est abattu sur la place Saint-Pierre, en face le magasin Thiriet, mais sans se faire aucun mal.

A quatre heures et demie, un autre

classe de mathématiques du collège pour le récompenser de son courage et de sa présence d'esprit en sauvant, au péril de sa vie, les jeunes Pourcelot, Canonge et Vuez, élèves des écoles primaires, qui, le 27 décembre dernier, avaient disparu sous la glace à un endroit où l'eau avait plusieurs mètres de profondeur.

Nous joignons aussi nos compliments les plus chaleureux à ceux de M. le ministre. La jeunesse ne manque ni d'élan ni de générosité, mais il est encore plus satisfaisant de la voir agir avec sang-froid et adresse dans les circonstances périlleuses. Notre jeune compatriote a été à la fois brave et prudent : il a bien mérité l'attention délicate de M. le ministre.

La semaine dernière a été très belle et la température a été relativement très élevée.

La neige continue à couvrir les montagnes.

Ferrière-sous-Jougne. — M. Charles Robbe, de Verrière-sous-Jougne, qui était parti avec la première portion de la classe de 1879, vient d'être nommé sous-lieutenant d'artillerie de marine et envoyé à Cherbourg.

Morteau. — Un vol avec effraction a été commis, le 14 janvier, à 9 heures du soir, dans la ferme de Mi-Bois.

Les voleurs ont enfoncé la porte du logement de M. Jacquier, horloger, et ont déménagé tous les meubles que renfermait la chambre à coucher.

Plainte a été portée à la gendarmerie.

Haute-Saône

Lure. — Tribunal correctionnel.

Audience du 14 janvier 1884.

Plusieurs Italiens travaillent dans une entreprise, à Pomoy, petite commune de l'arrondissement de Lure.

Le 1^{er} janvier, à neuf heures et demie du soir, un certain nombre de ceux-ci sortaient d'une auberge; sans doute ayant trop bu, ils voulaient boire encore, c'est pourquoi ils voulurent entrer dans un autre établissement, mais le patron refusa sa porte.

Figure III-17 : Extrait (1884-02-21), p.2.

Au traitement de la vie locale, succède un ensemble de rubriques réservées aux faits divers ainsi qu'au divertissement. On rencontrera en troisième ou

quatrième page les histoires drôles des *Échos* et les allusions railleuses de la *Petite gazette* :

Ch.III (05), « Etrennes utiles », *Échos*, (1884-03-01)

« Mon petit garçon, dit à Loulou un ami de papa, je vous apporte mon cadeau de jour de l'an.

- Ce n'est pas un livre, j'espère bien !
- C'est un bel atlas géographique universel.
- Universel, ça veut dire que tous les pays y sont?
- Précisément.
- Eh bien, donnez tout de même. Je n'aurai plus à étudier la géographie.
- Mais, au contraire !
- Puisque j'aurai tout là dedans quand je voudrai savoir, pas besoin de l'apprendre !

Ch.III (06), *Petite Gazette*, (1884-01-01)

C'est un usage universellement observé que celui de formuler force souhaits à l'occasion du jour de l'an. Ainsi bien comptons-nous ne pas y manquer pour notre part. [...] Nous souhaitons en outre :

A M. le président du conseil, un kilogramme de patience ;

A M. Clovis Hugues, le Code de savoir-vivre, de M. de Bassanville ; [...]

A M. Bernard (du Doubs), une boussole;

A M. Viette, moins d'amis; [...]

Au *Démocrate*, une syntaxe et le catéchisme poissard ; [...]

Nous aurions bien souhaité à la *Démocratie*, une direction fine; mais comme la tour de Babel semble être son idéal, nous ne pouvons que lui donner ce que nous offrons à tous : nos meilleurs souhaits de prospérité, de santé, de tolérance et de l'affermissement de la République par des réformes pratiques et sérieusement poursuivies.

Variété, qui apparaît en 1884, traite pour sa part de l'actualité culturelle régionale : il s'agit par exemple de présenter aux lecteurs les œuvres du peintre « Jules Grenier, de Baume» (1884-01-05). Le contenu de cette rubrique est donc sensiblement proche de la case *Beaux-arts*.

La fréquence de publication de *Dernière heure* en fait une des rubriques permanentes du quotidien, qui tend d'ailleurs à s'intensifier de 1883 à 1886 : en 1883, sa fréquence est de 60%, pour avoisiner les 80% en 1884 et en 1885, pour enfin devenir quotidienne en 1886. Pouvant être située en troisième ou en quatrième page, cette case abrite les dépêches parvenues à la rédaction après

la mise en page des feuilles¹⁰ : d'où son intitulé à caractère temporel, par ailleurs sensationnaliste, que ne manque pas de railler le chroniqueur Jacques du *Petit Parisien*. Agacé devant l'éclosion de cette rubrique dans bon nombre de quotidiens depuis le Second Empire et par sa prétention à présenter une actualité qui n'est que prospection, il qualifie cette case en ces termes :

L'une des plus vastes plaisanteries du journalisme à sensation, c'est incontestablement la rubrique intitulée : Dernière heure.

Pour ma part, en toute sincérité, je confesse n'y avoir jamais rien trouvé. Il paraît que, dans la « dernière heure », il ne se passe aucun événement.

Et cependant il n'est de feuille bien informée qui ne cultive avec acharnement cet alinéa bizarre, dont l'habitude fâcheuse est de rien contenir.

À la première colonne, vous lisez régulièrement :

Voir la dernière heure à la quatrième page.

Ou :

Voir la dernière heure au Bulletin de la Bourse.

Empressé, vous y courez, à cette quatrième page ou à ce Bulletin de la Bourse et vous y découvrez...Quoi ? Des télégrammes dans le genre de ceux-ci :

« On nous annonce_ sous toutes réserves_ que le roi des Patagons a failli avoir la jaunisse.»¹¹

La quatrième page : publicités et informations pratiques

La quatrième page, consacrée aux annonces dont la surface allouée ne va cesser de croître au fur et à mesure de la publication, héberge également des rubriques de renseignements économiques comme *Bulletin commercial*, *Halles et Marchés*, *Bourse de Paris*.

¹⁰ Plus exactement, les archives de la famille Millot font état de « nouvelles parvenues à la rédaction après 21h ».

¹¹ Jacques, « Dernière Heure », *Le Petit Parisien*, numéro du 3 novembre 1876, cité par Thérenty (2007 : 97).

HALLES & MARCHÉS	
MARCHÉ A LA CRIÉE.	
Louis THUNOT, facteur.	
Vente du 27 juillet au 4 août 1888.	
400 kil. mûres, de 1 fr. à 1 fr. 50 le kil.	
40 kil. saumon, de 3 fr. 50 à 4 fr. le kil.	
45 kil. sole, de 2 fr. à 2 fr. 50 c. le kil.	
132 paniers écrevisses, de 3 fr. à 4 fr. 50 le pan.	
292 poulets, de 1 fr. 75 à 2 fr. pièce.	
90 kil. lapin, de 1 fr. 50 à 1 fr. 60 le kil.	
164 pigeons, de 1 fr. 80 à 2 fr. 20 la paire.	
18 canards, 2 fr. pièce.	
317 douzaines œufs, à 75 c. la douzaine.	
60 kil. saucisses, à 2 fr. le kil.	
620 kil. abricotés, de 70 à 75 c. le kil.	
90 kil. champignons frais, à 2 fr. le kil.	
L'arrivage de la semaine a été de 200 colis, du poids brut de 2,256 kil.	

Figure III-18 : «Halles et Marchés » (1883-08-05), p.4.

BOURSE DE PARIS				
Du 4 août.				
Comptant.	Dernier cours.	Cours du jour.	Hausse.	Baisse.
Rente 3 %	79 93	79 95	» »	» »
Rente 3 % amortissable ancien.	81 70	81 95	» »	» »
Rente 3 % nouv.	» »	» »	» »	» »
Rente 4 1/2 %	111 80	112 »	» 20	» »
Rente 5 %	105 »	109 »	» »	» »
Terme.				
Rente 3 %	80 12	80 15	» 3	» »
Id. amortiss.	» »	» »	» »	» »
Rente 5 %	109 12	109 10	» »	» 02
Italien 5 %	80 80	90 70	» »	» 10
Autrichien 1876 (4 % or)	» »	» »	» »	» »
Egyptien unifié.	365 »	365 75	» »	1 25
Société générale.	525 »	522 50	» »	2 50
Credit Foncier de France	1308 75	1310 »	1 25	» »
Banque ottom.	735 »	735 »	» »	» »
Credit mobilier espagnol	205 »	208 75	3 75	» »
Actions Suez	2471 25	2468 75	» »	2 50
Actions Chemins autrichiens	675 »	675 »	» »	» »
Act. Nord Espag.	510 »	510 »	» »	» »

Figure III-19 : « Bourse de Paris » (1883-08-05), p.4.

L'organisation que nous venons de décrire demeure relativement stable jusqu'en 1886, où des remaniements notables sont effectués.

De 1886 à 1903

Les changements de « cuisine » opérés en 1886 affectent essentiellement la Une du *Petit Comtois*. Autonomes jusqu'en 1886, les rubriques *Chambre des députés / Sénat / Etranger / Intérieur* sont intégrées à la

rubrique *Dépêches de nuit*. Parallèlement à ce réaménagement de la Une, la rubrique *Au jour le jour* disparaît.

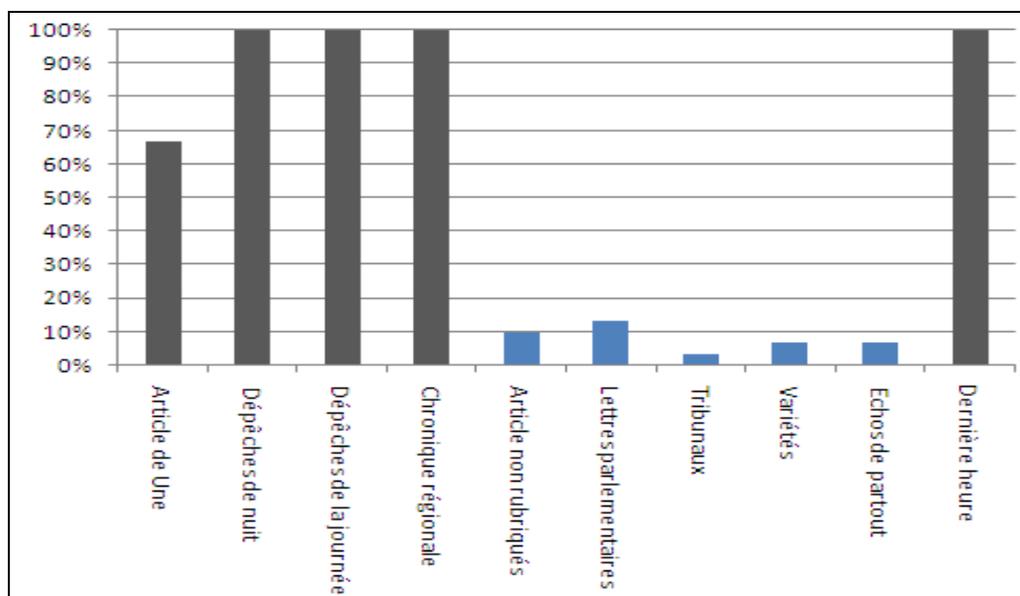


Figure III-20 : Fréquence de parution des rubriques en 1886 (en pourcentage).

En 1887, les *Dépêches de la journée* sont abandonnées. La même année, la rubrique *Bulletin* disparaît au profit de l'article de Une non rubriqué dont la fréquence de publication s'intensifie : publié à hauteur de 57% en 1887, l'article de Une voit sa fréquence croître jusqu'à devenir quasi-journalière en 1890 où elle apparaît dans 93% des numéros considérés. De 1890 à 1895, la fréquence soutenue de parution de cet article est stable pour tendre à décliner à partir de 1896. À partir de 1900, l'article de Une non-rubriqué ne paraît en moyenne que moins d'une fois tous les deux jours.

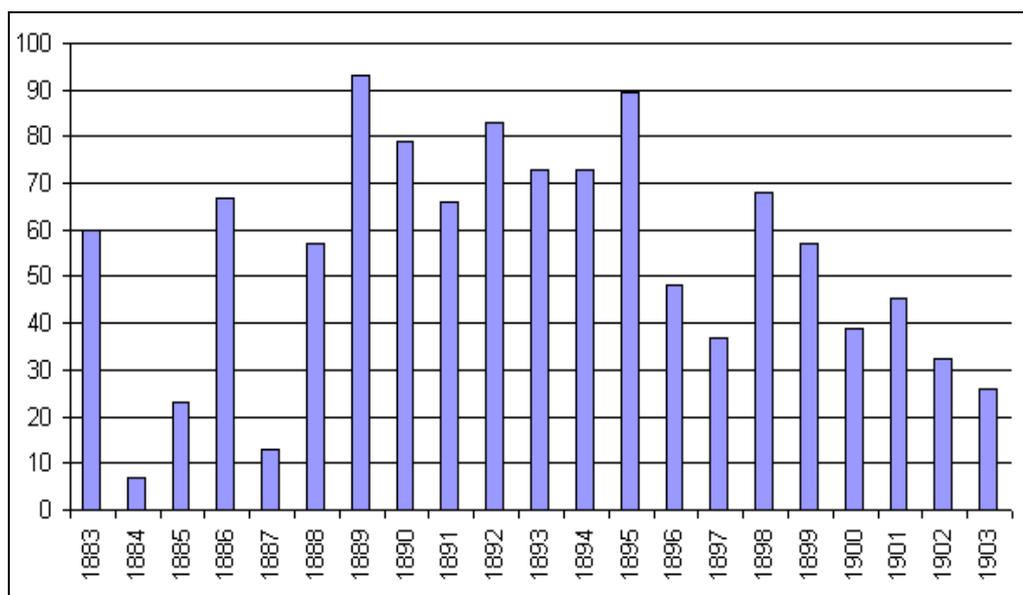


Figure III-21: Fréquence de parution de l'article de Une non-rubriqué en diachronie (en pourcentage).

Des rubriques périodiques s'installent également en Une, à l'instar des *Lettres de Paris* en 1888, par lesquels J. Gros, retenu à Paris, commente l'actualité politique.

Ultérieurement, les pages intérieures font à leur tour l'objet de modifications. Le quotidien s'enrichit de nouvelles cases, parmi lesquelles la *Petite Chronique*, tenue par un journaliste du nom d'Alexandre Tanchard. Ces chroniques, se proposant « de faire quelques excursions dans le passé »¹² afin de « déridier la rate »¹³ du lecteur, sont également le lieu d'articles d'opinion revenant sur l'actualité :

Ch.III (07), Petite chronique, (1899-02-23)

Je ne puis laisser passer sous silence les événements qui viennent de surgir ces jours derniers. M. Félix Faure, président de la République, est mort frappé d'une attaque d'apoplexie. Certes, ce lugubre événement n'aurait présenté, à mon avis, qu'une importance secondaire, s'il ne s'était présenté au moment même où la République traversait une crise aiguë.

En effet, ralliés à la conscience élastique, nationalistes braillards, centre-gauchers ventrus, royalistes découragés, impérialistes cherchant pâture, tout s'était uni,

¹²Cf. *Petite chronique* du 1899-02-18, signée par A. Tanchard.

¹³Op. cit.

coalisé, pour élever à la présidence de la République le Boulanger en redingote qu'ils dirigerait à leur gré, comme un polichinelle bien dressé.

Cet homme, est-il besoin de le nommer, c'était M. Méline, cet échappé de la commune recueilli aujourd'hui par la réaction et encensé par le clergé. Or, sénateurs et députés, réunis en congrès, ont mis fin à ce beau rêve en nommant un républicain de la veille, l'honorable président du Sénat, M. Loubet. [...]

Cette nouvelle organisation s'accompagne d'une augmentation des rubriques ponctuelles événementielles qui, pendant quelques numéros, le temps d'une « affaire » ou d'une campagne électorale, sont consacrées au traitement de l'actualité. Pour leur part, les articles non-rubriqués tendent à occuper une surface de plus en plus importante : s'ils occupaient moins d'une colonne de 1883 à 1886, ils occupent de 1886 à 1898 *a minima* 1,2 colonnes. En 1894 et 1898, la surface allouée à ces articles avoisine les deux colonnes. À partir de 1899, leur superficie décroît cependant assez nettement.

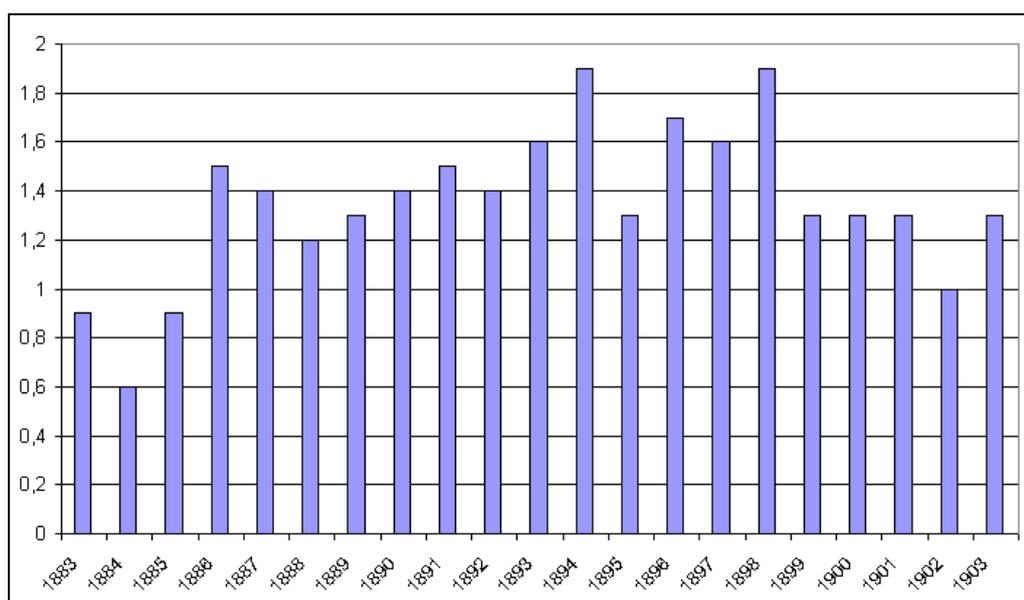


Figure III-22 : Surface occupée par les articles non-rubriqués en diachronie (en nombre de colonne).

De même, la surface occupée par la *Chronique régionale* tend à croître dès 1891 pour se stabiliser autour d'une surface égale ou supérieure à six colonnes à partir de 1895.

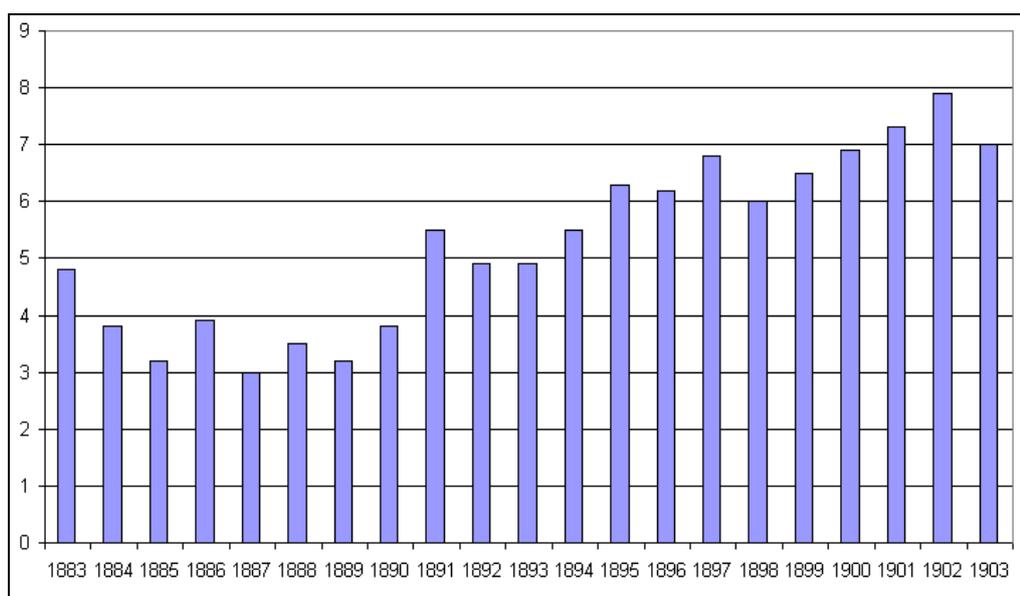


Figure III-23 : Surface occupée par « *Chronique régionale* » en diachronie (en nombre de colonnes).

Après 1890, les évolutions du rubriquage se concentrent essentiellement en troisième page. De nouvelles rubriques créées en 1893 comme *Causerie* entretiennent le lecteur de publications ou d'éléments du folklore régional :

Ch.III (08), Causerie, (1893-01-16)

Je retrouve dans ma petite bibliothèque un volume déjà vieux - il a 11 ans - mais toujours actuel. Peut-être n'aurait-il pas attiré mon attention, si son auteur n'était pas mort récemment, et il serait resté là, oublié au milieu d'autres moins sérieux cependant, puisqu'il concerne notre ville.

Je veux parler d'un livre qui a pour titre : *Besançon et ses environs*, par Auguste Castan, correspondant de l'Institut, et qui date de 1881. Ce n'est point une œuvre de fantaisie, tant s'en faut ! et l'auteur ne nous donne que des renseignements d'une rigoureuse exactitude ; mais, grâce à leur variété et la science de l'écrivain, elle est intéressante, surtout pour un Bisontin qui croit bien connaître la ville qu'il habite et s'aperçoit qu'il y a vécu comme un simple passant. [...]

Si elle apparaît également en 1893, la *Chronique agricole* connaît une durée de vie éphémère. Une nouvelle rubrique *Agriculture* voit le jour l'année suivante. Malgré une thématique similaire, on notera que la variation d'intitulé révèle une sensible différence de traitement, comme le suggèrent les extraits présentés ci-dessous. Alors que les articles de la *Chronique agricole* tendent à s'assimiler à

des articles de pratique-conseil, ceux de la rubrique *Agriculture* témoignent d'une prise en charge énonciative plus affirmée et traitent plutôt des questions relatives à l'agriculture sous un angle politique :

Ch.III (09), « Le jardin. Culture de l'Asperge », Chronique agricole, (1893-01-07).

Les lecteurs du *Petit Comtois* nous permettront bien de consacrer notre article d'aujourd'hui au jardin. Tout le monde connaît ce petit coin de terre que possède la plus humble chaumière campagnarde. C'est là que le cultivateur récolte les légumes dont il a besoin. L'importance du jardin est donc capitale dans un ménage; cependant il est bien souvent négligé. [...]

Ch.III (10), « Syndicats agricoles. Suite », Agriculture, (1894-02-12)

Pour ne pas sortir de mon programme, je ne veux m'occuper, en ce moment, que des syndicats agricoles pour drainages, curages, assainissements et irrigations. Voilà des améliorations qui sont négligées faute d'entente, d'accord entre les intéressés, et aussi, il faut bien le dire, faute des ressources nécessaires. C'est pour la formation de ces syndicats qu'il faut quelqu'un pour prendre l'initiative et arriver à la solution qui doit transformer notre agriculture, qui traverse en ce moment, et depuis déjà bien des années, une crise dont on ne saurait se dissimuler la gravité et dont elle ne sortira victorieuse qu'à la condition de savoir largement et hardiment modifier ses vieilles habitudes de pratique traditionnelle. Je vais examiner cette situation ; puis il sera facile d'indiquer le remède. Mais, hélas! Voudra-t-on l'appliquer, malgré ou peut-être à cause de sa simplicité ?

À l'opposé, alors qu'une rubrique comme *Échos* disparaît en 1890, *L'esprit des autres* apparaît en 1896 : en dépit d'une variation importante de l'intitulé, les deux rubriques présentent un contenu similaire. On comparera ainsi la similitude du calibrage et des fonctions de l'exemple (5), reproduisant un article de la rubrique *Échos* et l'exemple (11), extrait de la rubrique *L'esprit des autres* :

Ch.III (11), L'esprit des autres, (1896-01-06)

Entendu à la porte de l'église d'un village de la côte normande, le jour de la fête du pays :

« Papa, regarde donc le plumet du suisse, comme il est grand !

-En effet, mon enfant, mais ce n'est rien à côté de celui qu'il aura ce soir ! »

3.2.3. Des critères sous-jacents au système de rubriquage qui se révèlent composites

Quelques remarques doivent être formulées concernant les critères sur lesquels repose le système de rubriquage que nous venons de décrire. Dès les premiers numéros du *Petit Comtois*, les informations sont mises en forme selon un mode de classement indéniablement composite.

Un premier critère de rubriquage d'ordre temporel prévaut dans les intitulés des rubriques *Dépêches de nuit / Dépêches de journée / Au jour le jour / Dernière heure / Chronique de la semaine*. Si l'on opère une réduction sémantique, on remarquera en effet que ces intitulés partagent la caractéristique de se rapporter à la temporalité. *Dépêches de nuit, Dépêches de journée, Au jour le jour, Dernière heure* renvoient pour leur part à une temporalité quotidienne. Ces intitulés reflètent la volonté de faire coïncider la temporalité du monde avec celle du numéro du journal. Ce faisant, le quotidien met en avant sa « modernité ». Le rapport à l'actualité est en effet un des points essentiels de rupture entre la pratique journalistique du début du XIX^e siècle et celle de la fin du XIX^e siècle, comme le rappelle M.-E. Thérénty :

L'actualité journalistique du premier XIX^e siècle est donc une notion toute relative, d'abord en raison des délais postaux qui régissent la distribution des quotidiens et des revues. Les enquêtes menées par les historiens sont très instructives : en 1835, jusqu'à 220 ou 250 kilomètres de Paris, les abonnés recevaient leur quotidien le lendemain de sa parution. Passé cette distance, et jusqu'à 400 kilomètres, ils n'en disposaient que le surlendemain. Au-delà de 400 km, le Sud-Ouest était desservi plus vite que le Sud-Est : à Bordeaux, on recevait son journal le surlendemain de sa parution parisienne alors qu'à Marseille il fallait attendre quatre jours (Thérénty, 2007 : 91).

La coïncidence de la temporalité du monde avec celle du quotidien est une des conséquences directes des progrès techniques des moyens de communication survenus dans la seconde moitié du XIX^e siècle. Si le réseau télégraphique français est opérationnel dès 1855, le recours aux dépêches télégraphiques dans le champ journalistique se généralise seulement vers 1878 pour les informations françaises et 1884 pour les informations étrangères (Palmer,

1983 : 40). Dans le champ concurrentiel journalistique de la fin du XIX^e siècle, le traitement rapide de l'information proposé par un quotidien est donc un argument de vente, un élément de stratégie éditoriale décisif. Rappelons que dès son premier numéro, *Le Petit Comtois* met en exergue son traitement rapide des informations et se targue d'offrir « *le service d'information le plus complet et le plus rapide* » (Edition du 1^{er} août 1883). Les intitulés des rubriques basées sur la temporalité quotidienne correspondent ainsi à la mise en valeur d'une technologie récente, au travers de laquelle *Le Petit Comtois* se met en scène comme un quotidien, certes régional, mais d'information. Dans ce contexte, l'intitulé donné à chaque rubrique devient très important. À une heure où l'importance d'une information est donc fonction de la rapidité de son traitement, les rubriques qui mettent en avant l'actualité de leur contenu assurent en effet une indéniable fonction de hiérarchisation de l'information.

Le système de rubriquage du *Petit Comtois* repose également sur un mode de classification géographique. La rubrique *Chronique régionale*, relative à l'aire d'influence du quotidien, en témoigne, de même que les intitulés *Intérieur* et *Étranger*. Ce mode de classement géographique reflète tout autant l'identité régionale du *Petit Comtois* qu'il traduit sa prétention à être un quotidien d'information, revendiquée dès le premier numéro de sa parution. Présenter l'information sous l'axe de la proximité (*Chronique régionale*) et du lointain (*Étranger*) concourt en effet à créer l'illusion d'un discours d'information complet, qui donnerait à lire les événements survenus dans le monde entier. On doit à P. Charaudeau (2005b : 109) d'avoir souligné l'intérêt particulier suscité chez le lecteur par les nouvelles provenant d'un lieu proche de celui de l'instance de réception, ce que confirme l'historien R. Marlin, s'interrogeant sur le succès du *Petit Comtois* :

[...] les nouveaux lecteurs voulaient un horizon plus large que leur cadre quotidien, tout en étant bien aises que l'on parlât d'eux. Dans les journaux bisontins, la « chronique régionale » ne cessa d'augmenter, grâce aux nouvelles expédiées par des correspondants locaux mais arrangées de façon présentable par un secrétaire de rédaction, dont le rôle était de « faire la

cuisine » avec la littérature souvent médiocre qu'on lui envoyait (Marlin, 1965b: 184).

Que *Le Petit Comtois* institue en rubrique les nouvelles de l'étranger est à relier au même contexte concurrentiel que celui qui explique les intitulés de rubriques basés sur la temporalité quotidienne :

Dès le départ, un des principaux arguments concurrentiels de la presse d'information a été de transcrire aussi *instantanément* que possible des faits survenant aussi *loin* que possible¹⁴.

Un troisième principe de rubriquage thématique se dégage des rubriques *Agriculture*, *Chronique agricole*, et *Courrier de la mode*. Leur émergence est plus tardive que les rubriques précédemment citées reposant sur une organisation géographique ou relative à la temporalité. Elles amorcent les débuts d'un rubriquage thématique, qui ne sera effectif qu'à partir de 1910 dans *Le Petit Comtois*. L'instauration de ces thématiques (l'agriculture, la mode) en rubriques est un indice de la cible visée par le quotidien. En découpant le monde suivant ces catégories, ce quotidien d'une région rurale offre des sujets d'intérêts au « petit peuple » agricole. La rubrique *Courrier de la mode* témoigne d'une adaptation à l'élargissement du lectorat aux femmes dont prend acte la fin du XIX^e siècle.

Enfin, un quatrième mode de classement se donne à lire dans les intitulés des rubriques *Faits divers* / *Echos* / *Variétés* / *Petite gazette* / *Bulletin* / *Chronique*, renvoyant cette fois à la généricité des articles. Tous ces éléments se réfèrent à des formes d'écriture issues de la matrice littéraire. Ce faisant, en recourant à des catégories génériques, le quotidien montre son incapacité à procéder à un découpage du monde, comme le suggère M. - E. Thérenty :

D'autres rubriques fondées sur la forme même de la nouvelle plutôt que sur les nouvelles elles-mêmes constituent de fait des non-rubriques, des sortes de trous noirs de l'information et des endroits où le journal avoue lui-même son impossibilité à ordonner le monde : elles se dénoncent quelquefois dès leur intitulé même, comme la chronique, les faits divers ou les faits-Paris (Thérenty, 2007 : 83).

¹⁴ Will, N. (1976), *Essai sur la presse et le capital*. Paris : Union générale d'Éditions, p. 102, cité par Palmer (1983 : 13).

Selon cette perspective, impliquant qu'une rubrique procède nécessairement à un classement thématique du monde, il serait donc inapproprié de considérer les *Faits-Divers* ou autres *Variétés* comme des rubriques. Cette conception s'est d'ailleurs imposée au XX^e siècle, puisque les professionnels du journalisme les qualifient aujourd'hui¹⁵ de « fausses rubriques » (Agnès, 2002 : 59). Le recours à ces intitulés nous semble pointer une spécificité du journal du XIX^e siècle, encore très lié au champ littéraire, dont il emprunte les formes pour construire son propre discours.

3.2.4. Remarques conclusives sur le rubriquage du *Petit Comtois*

De 1883 à 1903, le rubriquage du *Petit Comtois* est ainsi en perpétuel réaménagement, par un mouvement de création et de suppression de rubriques, d'ouverture à de nouvelles thématiques¹⁶, mais aussi par un jeu de variation des intitulés de rubriques. Ces évolutions prennent néanmoins forme sur un fond d'invariant. De 1883 à 1903, la Une, principal dispositif de contact avec le lecteur, apparaît comme une devanture où se côtoient éléments factuels et articles d'opinion. Les pages intérieures du quotidien restent majoritairement, en termes de surface, des pages locales, auxquelles succèdent une dernière page, fondamentalement lieu des divertissements et de la culture, mais aussi de la publicité, dont la surface ne cesse de croître au fil de la parution.

Que reflète l'évolution diachronique du rubriquage du *Petit Comtois* ? Envisageons tout d'abord ce système sous l'angle de ses invariants. La stabilité des emplacements des cases rédactionnelles et la stabilité de la plupart des titres de rubrique, ainsi que la surface qu'elles occupent, mérite d'être éclairée à

¹⁵ On notera que les professionnels du journalisme recommandent très fortement aujourd'hui d'éviter ce type d'intitulés de rubriques, qualifiées de « fausses rubriques » : « Attention aux fausses rubriques ! Une rubrique doit indiquer un domaine de l'actualité, adapté à la spécificité de la publication, pas un genre journalistique » (Agnès, 2002 : 59).

¹⁶ Ou du moins à la mise en valeur de thématiques traditionnellement traitées dans le quotidien que celui-ci instaure en rubrique, à l'image de la « Chronique agricole » (1893).

la lumière de l'état des technologies de composition et d'organisation du travail rédactionnel caractéristiques de la presse du XIX^e siècle. Ce sur quoi nous voudrions insister, c'est le besoin impératif pour la rédaction de disposer d'un système stable de cases pré-allouées formatées permettant de distribuer le plus rapidement les contenus de chaque numéro afin de les soumettre à la composition et par suite, à l'impression. C'est là l'essence même du rubriquage : constituer un système permanent à même de s'adapter au flux de contenus divers et fluctuants. La permanence diachronique d'une rubrique telle que *Dernière heure* et sa stabilité topographique s'expliquent ainsi par le besoin pour la rédaction de disposer d'une case souple où seront distribuées les dernières nouvelles parvenues à la rédaction sans que soit chamboulée la maquette du numéro déjà constituée. C'est également parce que la totalité de la surface disponible de chaque numéro du *Petit Comtois* est conçue en termes de feuillets, autrement dit d'unités aux formats préétablis, que la plupart des rubriques occupent une surface relativement stable en diachronie. Comme l'attestent les archives de la famille Millot, la composition du *Petit Comtois* s'effectue en effet « feuillet par feuillet » selon une routine très précise. Ainsi, le rédacteur en chef de ce journal spécifie que le format du *Bulletin* doit être impérativement compris « entre 30 et 50 lignes ».

La permanence diachronique et topographique du système de rubriquage du quotidien autour de rubriques centrales, telles que *Dépêches de Nuit*, *Chronique régionale*, *Faits Divers* pour ne citer que celles-ci, doit être également rapportée à la stratégie éditoriale du quotidien, à l'ethos discursif (Amossy, 1999) du *Petit Comtois*. Ces rubriques construisent l'identité du *Petit Comtois*. Celui-ci est avant tout un quotidien régional rapportant les événements survenus dans un espace géographique proche de l'instance de réception. Mais si *Le Petit Comtois* est un quotidien régional, son horizon excède cependant très largement le local. Il fait ainsi la part belle à la vie politique nationale, en rapportant sur la totalité de la période observée les

débats de la chambre des députés et du Sénat¹⁷. La vie politique nationale est également très largement commentée dans les rubriques *Bulletin*, *Chronique de la Semaine*, et autres rubriques ponctuelles, à travers lesquelles le journal explicite son appartenance à une presse d'opinion. La fonction d'information du quotidien, loin d'être antinomique avec l'expression forte d'une sensibilité politique, est fortement actualisée. En outre, sa capacité à donner à lire l'actualité française et étrangère est l'argument de vente principal du quotidien, qu'il expose stratégiquement en Une. Une mise en perspective du rubriquage du *Petit Comtois* avec les quotidiens nationaux qui lui sont contemporains est à cet égard tout à fait révélatrice. Cette comparaison pointe à la fois le recours du *Petit Comtois* à des rubriques communément répandues et la singularité de sa stratégie éditoriale. Marc Angenot, fréquentant les quotidiens nationaux de l'année 1889, constate la similitude de leur système de rubriquage et la succession invariante de l'ordre des rubriques. Selon cet auteur, il est possible d'établir un type moyen idéal de structuration du journal de la fin du XIX^e siècle, dont nous reproduisons la description :

En première page, on trouve le « Bulletin du jour » ou le « Premier-Paris » (ce que nous [M. Angenot et ses contemporains] appelons l'éditorial) signé d'un polémiste attitré, mis en valeur à la première ou aux deux premières colonnes. Suivent les « échos politiques » et mondains et les « tartines rédactionnelles » de commentaire d'actualité, de chronique parisienne, eux aussi signés ou initialisés. Au rez-de-chaussée, se trouve le premier feuilleton, un roman ou bien, dans la presse « sérieuse », de la critique littéraire et dramatique. En seconde et troisième pages, les articles, dépêches, notes et entrefilets, se succèdent en désordre apparent. Les affaires internationales et les variétés parlementaires et judiciaires en page 2 ; les faits divers, échos du monde des spectacles, revue commerciale, bulletin financier en page 3. En page 4, on trouve d'ordinaire du non-rédactionnel : météorologie, liste des théâtres et spectacles, nécrologie, cours de la bourse, petites annonces, offres d'emploi et publicités (Angenot, 1989 : 536).

Cette description de la structuration d'un journal-type implique que le système d'intitulés de rubriquage du *Petit Comtois* est somme toute assez similaire à celui de la presse nationale du XIX^e siècle. Ce parangon articule des rubriques

¹⁷ L'intégration de ces débats dans la rubrique « Dépêches de nuits » par la création de sous-rubrique « Sénat » et « Chambre des députés » ne contredit en rien la permanence de leur traitement de 1883 à 1903.

d'opinion, de divertissement et d'information, dont les intitulés renvoient à un système de critères sous-jacents composites. Néanmoins, il convient de remarquer la place des dépêches dans le quotidien typique du XIX^e siècle décrit par M. Angenot. Dans celui-ci, contrairement à ce qu'on peut observer dans le *Petit Comtois*, les dépêches, reléguées en deuxième page, ne semblent pas faire l'objet d'une mise en valeur particulière, qu'elle soit d'ordre typographique et encore moins topographique. Il nous semble bien voir là la confirmation de la stratégie éditoriale du *Petit Comtois*, qui expose une technologie de communication en Une pour marquer sa modernité et sa différence par rapport aux autres quotidiens.

L'évolution du système de rubriquage perceptible de 1883 à 1903 peut tout autant être envisagée comme les effets d'une stratégie éditoriale. Développer des sujets qui interpellent son lectorat tout en diversifiant son contenu serait donc un des facteurs justifiant l'émergence de nouvelles rubriques, comme le *Courrier de la mode et Agriculture*. Pour défendre sa place dans le champ journalistique, *Le Petit Comtois* s'adapte ainsi aux goûts de son public, le « petit peuple » rural de Franche-Comté. La création d'une rubrique *Dépêche de la journée* vise pour sa part à renforcer l'image de quotidien d'information du *Petit Comtois*, exposant en Une un traitement de plus en plus rapide de l'actualité dans un contexte concurrentiel. Le jeu d'émergence et de disparition des rubriques est en effet directement lié à la situation financière du quotidien. Dans un contexte de tensions entre le conseil d'actionnaires du journal et Jules Gros, relatives à la situation financière du quotidien et des dissensions politiques en raison de bulletins politiques de la plume du journaliste jugés trop virulents par le conseil, Jules Gros met à profit son séjour à Paris pour s'improviser correspondant spécial du *Petit Comtois* afin d'économiser les frais de correspondant spécial engendrés par l'acquisition des dépêches. Pour pallier le manque de diversité des dépêches, Jules Gros se propose alors de tenir une nouvelle rubrique, les futures *Lettres de Paris*. En témoigne cette lettre du 22 janvier 1884 dans laquelle il défend ce projet :

M. Obein qui a l'esprit moins rigide que le mien suivra plus aisément la ligne de conduite que vous voudrez bien lui donner. Il n'y aurait donc qu'à lui faire prendre ma place, pendant que je me rendrais à Paris, où je serais le correspondant du journal. Le *Petit Comtois* y gagnerait donc déjà les 280 fr. mensuels que l'on donne à M. Sylvin. [...]

Pour racheter le défaut d'intérêt qui résulterait de la suppression d'une partie des Dépêches, je pourrais m'occuper à une chronique parisienne, mettant principalement en relief l'œuvre franc-comtoise, et rédigé sur le mode gai qui ferait de notre journal une sorte de *Petit Figaro* régional.

L'intérêt ne manquerait donc pas au *Petit Comtois*, et il enterrerait ainsi les polémiques qui paraissent avoir effrayé les actionnaires, bien qu'à mon avis, elles aient été dans la rigueur de la situation¹⁸.

3.3. Formes et fonctions des titres

Titrer un article est une pratique qui s'est développée plus tardivement que celle du rubriquage. Comme le rappellent M. Mouillaud et J.-F. Tétu (1989 : 65), il est ainsi rare de rencontrer des titres dans la presse jusqu'à la fin du XIX^e siècle. Dans la continuité de notre examen de la mise en forme de l'information du *Petit Comtois*, nous examinerons l'évolution des formes et fonctions de l'appareil titulaire dans ce quotidien.

3.3.1. Approche générale du titre

Le titre est le seul objet péritextuel à avoir donné lieu à un ensemble conséquent de recherches, même si cette littérature se révèle cependant restreinte en comparaison avec le nombre d'études prenant pour objet la presse écrite. On citera parmi les travaux les ouvrages de J.-P. Sueur (1968), à l'origine des travaux menés sur les titres de journaux par M. Mouillaud (1982) et J.-F. Tétu (1989), de J. de Bonville (2008), de F. Frandsen (1990), de P. Fiala et de B. Habert (1989). Plus récemment, à la suite des publications de L.H. Hoek (1981), Sullet-Nylander (1998) procède à une description à la fois syntaxique, sémantique et rhétorique du titre.

¹⁸ Lettre du 22 janvier 1884 de Jules Gros à un actionnaire inconnu, extraite des archives de la famille Millot.

3.3.1.1. Retour sur les fonctions du titre

D'un point de vue sémiotique, les titres, de par leurs caractères formels (gras, italique, majuscule) instaurent une rupture avec la densité textuelle de l'espace de la page. Pour la rédaction, le titre, tout comme le rubriquage, est un dispositif de structuration de l'aire scripturale (Peytard, 1975) et de mise en forme de l'information. *Entaille* typographique, le titre assure une fonction de dissociation. Cette dernière fonction est le ressort d'une captation de l'attention du lecteur : les titres attisent la curiosité et invitent à la lecture. Ce sont ces caractéristiques que les praticiens du journalisme contemporains mettent en exergue. C. Furet (1995) répertorie ainsi cinq fonctions essentielles du titre : accrocher le regard, donner au lecteur envie de lire l'article, permettre le choix de lecture, structurer la page et contribuer à l'image du journal. Ce dernier trait fonctionnel du titre pointé par C. Furet nous semble essentiel et justifie la pertinence d'une étude du discours d'un quotidien ayant pour préalable une analyse de ses titres. Le titre est un lieu privilégié d'accès aux attitudes d'un quotidien. À travers un examen systématique et/ou comparatif de l'appareil titulaire se donnent à lire une stratégie éditoriale et un positionnement socio-discursif sous-jacents.

Du côté des théoriciens, qu'ils soient linguistes, analystes du discours ou spécialistes de l'information et de la communication, la réflexion sur le titre est orientée principalement sur le fonctionnement sémantique et pragmatique de ce dernier. Adhérant à une opinion communément répandue dans la vulgate et dans les manuels de journalismes contemporains¹⁹, T.A Van Dijk postule que la fonction première du titre est de résumer l'information délivrée dans l'article qu'il précède :

To wit, the category of Headline in a news discourse, has a fixed form and position in news items in the press. At the same time, this Headline has a very specific thematic function : it usually expresses the most important topic of the news item (Van Dijk, 1985 : 69).

¹⁹ Notamment Martin-Lagardette (2000).

Cette position est largement contestée, puisque linguistes, sémioticiens et analystes du discours tendent à considérer que le fonctionnement du titre excède la simple fonction de résumé. C'est notamment la position tenue par F. Frandsen (1990), qui remarque d'ailleurs que certains titres n'assurent aucunement une fonction de résumé de l'information.

À ces considérations sur les fonctions assurées par le titre correspond une réflexion sur l'autonomie du titre : capacité du titre à être lu indépendamment de l'article qu'il surplombe. Les manuels de journalisme contemporains précisent ainsi que le titre doit être compris « sans qu'il soit nécessaire de lire le texte » (Furet, 1995 : 104). Nombreux sont les théoriciens qui abondent dans ce sens, à l'instar de L.-H. Hoek (1981 : 2) précisant que « le titre est autonome par rapport au texte », et de P. Charaudeau :

Les titres, dans l'information, sont d'une importance capitale ; car, non seulement ils annoncent la nouvelle (fonction " épiphanique"), non seulement ils conduisent à l'article (fonction guide), mais encore ils résument, ils condensent, voire ils figent la nouvelle au point de devenir l'essentiel de l'information. Le titre acquiert donc un statut autonome ; il devient un texte à soi seul, un texte qui est livré au regard des lecteurs et à l'écoute des auditeurs comme tenant le rôle principal sur la scène de l'information (Charaudeau, 1983 : 102).

Il nous semble percevoir une corrélation forte entre la position adoptée par le théoricien concernant les fonctions du titre et sa capacité d'autonomie, notamment perceptible dans la définition suivante du titre proposée par F. Sullet-Nylander :

[...] un texte reprenant le contenu essentiel (ou considéré comme tel par le rédacteur) de l'article, sans que cela n'engendre de dépendance « textuelle » quelconque vis-à-vis de celui-ci (Sullet-Nylander, 1998 : 307).

3.3.1.2. Lectures du titre

Les systèmes de classification de l'appareil titulaire varient en fonction des fonctions attribuées au titre. Nous souhaitons ici présenter trois propositions terminologiques majeures dans la littérature scientifique sur le titre.

Les deux premières propositions que nous exposerons brièvement reposent sur des critères principalement sémantiques.

Pour sa part, au terme d'une analyse sur l'appareil titulaire du *Monde*, M. Mouillaud (1982) propose une classification des titres de presse selon une opposition entre *titre informationnel* et *titre référentiel*. Ayant pour fonction principale de résumer l'article qu'il précède, le titre informationnel est un énoncé par lequel est réalisée une assertion sur le réel. Le titre référentiel, au contraire de l'énoncé informationnel, ne produit pas d'information sur le monde, mais fait référence à un savoir sur un état du monde. En cela, le titre référentiel est une « charnière » entre le monde et le journal, mais surtout entre le journal et la collection :

[...] le titre référentiel est, d'ailleurs, lui-même dédoublé dans son objet et dans son statut; informatif, il vise le journal; visant le monde, il ne le fait que sous la forme d'un rappel, d'une mobilisation qui est un changement de statut; il est ainsi orienté dans deux directions; une flèche est dirigée vers l'amont, elle agraphe le journal au monde (à une connaissance antérieure du monde) et à sa propre collection; une autre flèche est dirigée vers le journal et vers l'aval de la lecture; elle dit « au sujet de quoi » l'information va être produite (Mouillaud, 1982 : 83).

Le titre rubrique, qui relève de la catégorie du titre référentiel, renvoie ainsi à un état du monde thématique ou géopolitique : c'est notamment le cas des titres de rubriques tels que *Société*, *Culture*, *Étranger*, etc. Le titre anaphorique, renvoie quant à lui à un processus historique, comme dans l'exemple cité par M. Mouillaud et J.-F. Tétu (1989 : 119) : *La menace de démission de M. Begin*.

Parce qu'elle est élaborée pour classer les titres du XIX^e siècle, nous signalerons également la proposition de classification établie par J. de Bonville. S'interrogeant sur les particularités des titres de la presse québécoise de la fin du XIX^e siècle au début du XX^e siècle, Jean de Bonville a procédé à l'analyse de 1269 titres de faits-divers, corpus résultant d'une sélection biannuelle de 50 textes entre 1885 et 1935. L'auteur distingue tout d'abord les titres *descriptifs* prévalant jusqu'aux années 1890 et qui tombent en désuétude à partir de 1915. Des titres comme *La variole*; *La tempête*, *La journée d'hier* sont

qualifiés de descriptif parce qu'ils « évoquent une situation ou un état de fait » (Bonville (de), 2008 : 35). La seconde catégorie de titres distinguée par J. de Bonville sur le même corpus est celle des titres *narratifs*, qui relatent brièvement une action. L'auteur range dans cette catégorie tant les titres tels que *Des bandits font feu sur un marchand*, que les syntagmes nominaux plus courts, mais qui, selon lui, évoquent un schéma d'événement comme dans *Vol à la poste* ; *Grand incendie* ou encore *Assaut grave*. Cette classification nous semble fragile à plusieurs titres. Intrinsèquement, les frontières entre titres narratifs et titres descriptifs sont contestables. Un syntagme nominal comme *La tempête*, exemple cité par l'auteur pour illustrer la deuxième catégorie, implique cependant un schéma d'événement dans la même mesure qu'un syntagme tel que *Grand incendie*.

De surcroît, ces classifications uniquement basées sur des critères sémantiques présentent la faiblesse de négliger la diversité pragmatico-énonciative des titres de presse, ce que nous mettrons en lumière au cours d'un examen liminaire des titres du *Petit Comtois*.

La troisième classification que nous signalerons ici diffère des précédentes en cela qu'elle délaisse les critères purement sémantiques pour se baser sur la polyphonie. Dans le même temps, cette proposition pointe la diversité des fonctions pragmatico-énonciatives du titre. On doit cette classification à J. Peytard (1975) qui établit un *distinguo* entre les titres verbaux, exprimant les pensées ou les dire d'un personnage et les titres non-verbaux, définis comme ne relevant pas de la responsabilité énonciative d'un personnage. On remarquera que la définition suivante de cette dernière classe de titres reconnaît à ces derniers la diversité des visées du titre, à savoir qu'il peut, selon J. Peytard, commenter, décrire et signaler :

Nous utilisons ici les concepts de « verbal » et « non verbal » pour classer les intitulés : appartient au « verbal » tout intitulé qui retranscrit les paroles et les pensées d'un personnage; au « non-verbal », tout intitulé qui commente, signale, décrit, sans être mis en rapport avec un personnage (Peytard, 1975 : 45).

3.3.2. Les titres dans *Le Petit Comtois*

En premier lieu, il ressort d'un parcours flottant des numéros de notre corpus que si les pages du *Petit Comtois* comptent un grand nombre de rubriques, elles n'hébergent que très peu de titres. Un examen attentif de la répartition des titres dans les pages du *Petit Comtois* confirme cette première observation et met au jour une indéniable corrélation entre l'appareil titulaire et le rubriquage. Cette corrélation se réalise en diachronie à travers plusieurs cas de figures que nous présenterons ici.

Titre et rubrique tendent à entrer dans une logique d'antagonisme, et ce, tout spécialement dans les débuts de parution du quotidien. On observe qu'un nombre important de rubriques ne tolèrent aucun titrage de leurs articles sur la totalité de la période 1883-1903. Les articles des rubriques *Bulletin*, *Au jour le Jour*, *Chronique*, *Chronique parisienne*, *Petite Chronique*, *Échos* ne sont ainsi jamais titrés. L'incompatibilité entre titre et rubrique est également perceptible au regard de la répartition des titres dans le quotidien et ses unités rédactionnelles : la majeure partie des titres du corpus occurrent en effet dans des espaces non-rubriqués. C'est par exemple le cas des articles titrés mais non rubriqués de première page comme (12) :

Ch.III (12), « Touchantes pudeurs », (1884-01-04)

La *Démocratie* se voile aujourd'hui la face, parce que nous avons écrit hier, que « tous les gouvernements sont bons, s'ils sont acceptés par la nation. »

Nous comprenons mal cet accès de pudeur chez une feuille qui professe encore plus résolument que nous-même le dogme de la souveraineté populaire. Il est vrai que notre confrère est obligé par destination de présenter quotidiennement à l'admiration électorale de ses concitoyens des gens qui s'entendent si peu sur la constitution politique à donner au pays, qu'il peut bien oublier le lendemain ce qu'il a écrit la veille sur l'omnipotence du suffrage universel.

Notre confrère pense peut-être aujourd'hui que la République doit être préexistante à la volonté nationale. S'il veut être logique, c'est en effet l'opinion qu'il doit professer lorsqu'il a pour directeur de semaine celui de ses hommes d'Etat qui prétend qu'on doit corriger les erreurs du suffrage universel par les effets persuasifs de la Trique, ce qui est le comble du libéralisme, comme on sait! [...] Ah ! messieurs, vous nous excommuniez au nom de la doctrine ! Eh ! bien en voilà des principes ! Nous voulons, nous, organiser la république, et non la désorganiser ! Et ce qui vous gênera peut-être davantage, c'est que ces principes,

nous n'entendons pas les formuler seulement pour la galerie électorale, mais que nous comptons bien en provoquer l'application. [...]

Au passage, on remarquera que ce titre ne peut que difficilement être analysé à la lumière des classifications précédemment citées. « Touchantes pudeurs » est un syntagme nominal qui ne peut nullement être envisagé comme un titre informationnel et encore moins comme un titre référentiel. Il n'informe pas sur le monde, pas plus qu'il n'instaure un horizon d'attente thématique. Il ne correspond de surcroît ni à la catégorie de titres descriptifs ni à celle des titres narratifs, en cela qu'il n'évoque aucune situation ni n'implique un schéma d'événement. La particularité du titre (12) exige de prendre en compte des phénomènes d'ordre pragmatico-énonciatifs et rhétoriques : cet énoncé se caractérise en effet par une dimension subjective indéniable et une tonalité ironique participant à la construction de l'*ethos discursif* (Amossy, 1999) du journal. Par là même, le type de titre illustré par (12) appelle à un élargissement des critères de classification de l'appareil titulaire. On remarquera également la fonction d'appel de ce titre qui éveille la curiosité du lecteur.

En diachronie, le rubriquage tend à s'assortir d'un appareil titulaire. De nombreuses rubriques qui, dans les débuts de parution du *Petit Comtois*, hébergent des articles non-titrés se dotent peu à peu d'un système titulaire. C'est notamment le cas des *Faits Divers*, où la pratique du titrage s'impose en diachronie. Si, en 1883, seulement 11% des faits-divers sont titrés, l'utilisation des titres se généralise à partir de 1884 (47% d'articles titrés), pour devenir systématique à partir de 1892. Comme l'exemple (13) l'atteste, les fonctions du titre de l'article se révèlent fort différentes de celle du titre-rubrique, en l'occurrence « faits-divers » qui se limite à spécifier la forme générique de l'article :

Ch.III (13), « Ils vont bien les sujets de Léopold », Faits divers, (1898-02-09)

Le tirage au sort est, tous les ans, en Belgique, l'occasion de scènes d'orgie et de sauvagerie, qui se terminent trop souvent de façon tragique.

A Matines, des batailles en règle ont été livrées sur différents points de la ville; on cite un des combattants qui a reçu un coup de poignard dans le dos.

A Bincba, il a fallu le concours de la gendarmerie pour disperser les batailleurs qu'il a fallu charger sabre au clair.

À Garnières, des miliciens sont entrés dans un cabaret où se trouvait un vieillard de quatre vingt-un ans qu'ils ont entrepris. Un des assaillants a saisi l'octogénaire à la gorge et l'a serré si fort que lorsqu'il l'a lâché ce n'était plus qu'un cadavre.

De même que (12), le titre (13) échappe à une classification reposant sur des critères essentiellement sémantiques prenant forme dans une logique dichotomique, où un titre est pourvu exclusivement d'une valeur donnée et dépourvu d'une autre. Au contraire, (13), auquel sont assignées diverses fonctions, pointe la nécessité d'envisager le titre comme un énoncé caractérisé par l'intrication de diverses fonctions, à lire en termes de gradients. En premier lieu, « Ils vont bien les sujets de Léopold », considéré indépendamment de l'article qu'il surplombe, peut être envisagé comme un énoncé à valeur informationnelle. Le sujet fait en effet l'objet d'une prédication. Mis en perspective avec l'article qu'il surplombe, l'exemple (13) est une assertion par laquelle est exprimée une opinion en réaction à l'information principale de l'article. Qui plus est, ce commentaire se révèle empreint d'ironie, en cela que le titre constitue une mention d'une évaluation contradictoire avec celle de l'instance citante qui se donne à lire dans le corps de l'article. Le titre de l'article, porteur d'une affirmation positive, est dissonant avec l'énumération des actes de sauvagerie narrés dans l'article. Témoignant d'une visée de commentaire, l'exemple (13) est aussi caractérisé par une tonalité humoristique, renforcée par l'illusion d'oralité créée par l'emploi d'un pronom clitique dans une construction disloquée.

L'adoption progressive d'un appareil titulaire n'est pas le fait exclusif de la rubrique des *Faits Divers*. La même évolution est sensible en diachronie dans les articles de la rubrique *Chronique régionale*. Pendant les premières années de parution, la grande majorité des articles sont simplement titrés en fonction du lieu d'origine des nouvelles, comme l'illustre la figure suivante :

<p>Haute-Saône, où, à présent, il n'est plus rien du tout.</p>	<h2>Haut-Rhin</h2>
<p>Lure. — Le 1^{er} janvier 1884, à neuf heures du matin, un accident s'est produit à la gare de Lure. Deux wagons de marchandises se sont heurtés à la suite d'une fausse manœuvre d'aiguille et l'un d'eux a été brisé. Les pertes sont purement matérielles.</p>	<p>Belfort. — A propos de la lettre de M. Juster, et non Inster, comme les typographes nous l'ont fait dire, nous avons essayé d'établir que la campagne ouverte contre le monument Thiers-Denfert avait été menée surtout par le parti clérical, et nous avons trouvé étonnant qu'une municipalité républicaine comme celle de Belfort fit opposition à l'érection de ce monument, dont la signification n'est point hostile aux opinions politiques qu'elle avoue, faisant ainsi le jeu du parti clérical.</p>
<h2>Jura</h2>	<p>Le pieux <i>Journal de Belfort</i>, qui publie aujourd'hui la lettre de M. Juster, justifie absolument nos appréciations par ses commentaires, et il déclare même expressément que plusieurs des souscripteurs du monument ont entendu ne « s'associer qu'à l'hommage rendu à M. Thiers, mais qu'ils refusaient tout témoignage d'admiration » à Denfert.</p> <p>Nous avons donc eu raison de dire qu'en refusant le monument Thiers-Denfert sous des prétextes plus ou moins vraisemblables, la municipalité de Belfort fait le jeu du parti clérical.</p>
<p>Dole, 1^{er} janvier. — Notre gardo champêtre, M. Desgrange, a arrêté aujourd'hui, au hameau du Poiset, un jeune soldat de la dernière classe, nommé Dayonnard, cavalier au 1^{er} régiment de dragons en garnison à Gray; ce militaire n'avait aucun titre pouvant justifier de sa présence dans cet endroit. Il a d'ailleurs déclaré que, s'ennuyant après ses parents et ayant encore quelques effets à aller prendre chez lui, il n'avait rien trouvé de mieux que de partir sans tambour ni trompette le 31 décembre au matin. Il a été remis à la gendarmerie pour être ramené à son corps à Gray.</p>	

Figure III-24 : Extrait de « *Chronique régionale* » (1884-01-02), p.2.

En 1883, seulement 5% des articles de cette rubrique sont surmontés d'un titre à proprement parler. Progressivement, cette pratique se développe, jusqu'à concerner 17% des articles de la rubrique en 1886 et 35% des articles en 1903. Un rapide examen des titres suivants permet de préciser les caractéristiques syntaxiques et les fonctions de ce système titulaire :

Ch.III (14), « Instruction primaire », (1888-02-21)

À la suite d'un rapport de M. l'inspecteur primaire de Pontarlier, M. l'inspecteur d'académie a adressé une lettre de félicitation à M. Simon, instituteur à Gilley, pour les excellents résultats qu'il obtient dans son école. Nos félicitations personnelles à M. Simon, qui déploie tant de zèle à l'instruction de la jeunesse.

Ch.III (15), « Le dégel », (1893-01-25)

Les Bisontins sont véritablement à plaindre ; si leurs affaires les obligent à sortir, ils trouvent des rues transformées en lacs et des trottoirs couverts de verglas ; quand on n'a pas de l'eau jusqu'à la cheville, on risque de se tuer ou se casser un membre ; voilà une jolie perspective.

Aussi sommes-nous de l'avis de quantité de nos lecteurs : le service de la voirie est mal fait. Nous savons bien qu'actuellement, il est débordé. Mais c'est sa faute,

pourquoi n'a-t-il pas engagé ou plutôt créé des équipes d'ouvriers qui, au moment du dégel, se seraient vraiment mises vivement à l'ouvrage et nous auraient évité l'inconvénient de ne pas pouvoir sortir ?

Il ne manque pas de pauvres diables qui n'auraient pas demandé mieux de gagner, de cette façon, du pain pour eux et leur famille. [...]

Ch.III (16), « Un voleur qui l'a échappé belle », (1892-01-23)

Dans la nuit du 17 au 18 janvier, vers minuit, le sieur Collet, Antoine, cultivateur, était réveillé par les aboiements de son chien.

Au même instant, le fils Collet, entendant du bruit au grenier, s'arma d'un fusil et alla droit au voleur soupçonné. Sommé deux fois de se faire connaître, il ne répondit pas. Collet le mit en joue et tira. Le coup ayant raté, le voleur en profita pour sauter à terre de plusieurs mètres de hauteur et prendre la fuite. Il était accompagné d'un complice, qui a également fui. Tous les deux ont été reconnus par le sieur Collet.

Ces sont les nommées M...et B..., de la commune d'Ouvans.

L'un d'eux avait laissé sa casquette au grenier.

Ces malfaiteurs avaient escaladé un mur pour pénétrer dans l'habitation.

Les titres (14) et (15) peuvent tout deux être envisagés comme des titres référentiels : (14) renvoie à un découpage du monde thématique, tandis que (15) renvoie à un procès historique, qui peut être assimilée, dans son contexte d'énonciation, à une famille événementielle. Ces titres donnent donc une information sur la matière de la brève, sans condenser son information principale. Le titre (16) participe d'une autre classe de titres, puisqu'il livre une information par laquelle est condensée une des informations jugées principales de l'article.

Enfin, un dernier cas de figure reste à signaler. Dès les débuts de parution du *Petit Comtois*, certaines rubriques sont structurées par un appareil titulaire. Il s'agit des brèves que l'on peut les rencontrer dans les rubriques *Dernière heure*, *Dépêches de nuit*, *Dépêches de la journée*. On remarquera que tous les titres de ces rubriques présentent la particularité d'avoir pour critère sous-jacent une référence aux supports et à la temporalité. La compatibilité entre le titre-rubrique et le titre nous semble se justifier par là-même : le titre-rubrique de ces unités rédactionnelles est, comme nous l'avons dit plus haut, une sorte de case fourre-tout, qui héberge une multitude d'informations n'ayant d'autres points communs que celui d'avoir été réceptionnées par dépêches.

L'appareil titulaire permet ici dès lors de distribuer l'information et de dissocier les brèves qui se succèdent, comme on peut l'observer sur la figure ci-dessous.

<p>auraient besoin d'un prince pour jouer le rôle de la mouche du coche.</p> <p>Nos lecteurs remarqueront aussi, sans doute, que notre correspondant de Paris se demande si tout le tapage fait par les officieux de M. de Bismark n'aurait pour but d'excuser par avance des mesures draconiennes contre l'Alsace-Lorraine.</p>	<p>d'écourter les vacances au 1^{er} octobre est écartée.</p>
<p align="center">DÉPÊCHES DE NUIT</p> <p align="center"><i>Service de notre correspondant spécial.</i></p>	<p>M. Jules Ferry ne quittera pas Paris avant huit jours.</p>
<p align="center">Paris, le 28 août, 9 h. 15 soir.</p> <p align="center">Promotion de généraux.</p>	<p align="center">Epuration.</p> <p>M. le ministre de la justice, Martin-Feuillée, rentrera vendredi prochain.</p>
<p>Dans le conseil des ministres qui a eu lieu aujourd'hui, comme nous l'avions annoncé, on s'est occupé d'abord des affaires courantes.</p>	<p>Le mouvement judiciaire sera présenté et étudié dans le prochain conseil, et la loi sur la magistrature sera promulguée le 31 août.</p>
<p>M. le général Thihaudin, ministre de la guerre, a soumis un projet de promotion de généraux de division. Mais le mouvement n'est pas encore arrêté.</p>	<p>La semaine suivante, un premier mouvement sera publié. Ce mouvement portera sur douze postes de premiers présidents, notamment ceux de Lyon, Aix, Douai et Poitiers.</p>
<p align="center">Les élus.</p> <p>Parmi les noms mis en avant, on signale ceux des généraux Tricoche, Casseron de Villenoisy et Comte.</p>	<p align="center">Les ministres n'y vont plus.</p>
<p>Les nominations de ces officiers généraux seraient les seules arrêtées à ce jour. Il reste deux choix à faire, pour lesquels on parle de MM. les généraux Galland et Lambert, ce dernier commandant à Tunis.</p>	<p>M. le ministre de l'intérieur, Waldeck-Rousseau, renonce à se rendre au Puy pour l'inauguration de la statue de Lafayette.</p>
<p align="center">Généraux de brigade.</p> <p>Cette promotion comprendra également neuf généraux de brigade.</p>	<p align="center">Tenez-vous bien, Malgaches.</p> <p>M. le contre-amiral Galiberti quitte Paris pour prendre le commandement de l'escadre de Madagascar, en remplacement de l'amiral Pierre, qui revient très souffrant.</p>
<p>Les colonels déjà choisis seraient MM. Boudet, Raison, Voisin, Logerot et Cary.</p>	<p align="center">Désappointement.</p> <p>Le correspondant du <i>Temps</i>, à Vienne, lui télégraphie qu'il apprend de source certaine que le testament de M. de Chambord ne contient pas un mot ni sur la succession au trône ni sur la politique.</p>
<p align="center">Pas de nouvelles du Tonkin.</p> <p>Le gouvernement n'a reçu aucune nouvelle de l'Indo-Chine.</p>	<p align="center">Cuisine de famille.</p> <p>La <i>Gazette nationale de Berlin</i> répand le bruit d'une prochaine entrevue, à Swinemunde, entre l'empereur Guillaume et le czar.</p>
<p>Le conseil a discuté aujourd'hui la question des renforts à envoyer au Tonkin.</p>	<p align="center">Tenons-nous bien, l'Espagne s'en mêle.</p>
<p align="center">Formation d'un régiment de marche.</p>	<p>Le gouvernement allemand fait de grands efforts pour amener l'Espagne à la triple alliance, et par là isoler complètement la France.</p>
<p>Il est question de former un régiment de marche avec deux bataillons de la légion étrangère et deux bataillons de tirailleurs algériens.</p>	<p align="center">Nous ne le verrons pas !</p> <p>Le roi Alphonse XII se montre très sensible aux avances de M. de Bismark et paraît décidé à se rendre en Allemagne sans passer par la France à son retour, comme il en avait été question.</p>
<p align="center">La Chambre ne sera pas convoquée.</p> <p>Toute idée de convoquer la Chambre ou</p>	<p align="center">Pas d'accord.</p> <p>Une crise ministérielle a éclaté à Madrid. M. le maréchal Martinez Campos se montre décidé à la retraite.</p> <p>M. Sagasta reformerait le ministère. Les</p>

Figure III-25 : « Dépêches de nuit » (1883-08-29), p. 1.

Au-delà de cette première remarque, on observe que ces titres assurent parallèlement d'autres fonctions. Ce sont précisément les fonctions de ces titres, mais aussi leur forme et les attitudes du quotidien qu'ils concourent à structurer, que nous nous proposons maintenant d'étudier.

3.3.3. Analyse des titres de la rubrique des « *Dépêches de nuit* »

Dans la continuité de notre observation de l'évolution des pratiques journalistiques et de définition de l'identité du *Petit Comtois*, notre intérêt s'est tourné vers un examen de l'appareil titulaire de ce quotidien pendant la diachronie 1883-1903. Pour ce faire, nous avons construit un sous-corpus d'étude composé de toutes les brèves de la rubrique *Dépêches de nuit* de 1883 à 1903, soit 6910 titres.

Le choix de ce corpus a été déterminé en premier lieu par nos objectifs méthodologiques. La nature systématique de l'analyse a exigé de disposer d'un ensemble de données péritextuelles assez substantiel. Notre volonté d'étudier le titre dans son environnement textuel en diachronie, découlant de l'option théorique selon laquelle le titre de presse constitue une entité inséparable de l'ensemble textuel plus vaste qu'il subsume, a eu quant à elle pour pendant de restreindre le volume de nos données à une case rédactionnelle publiée sur la totalité de la période de notre corpus.

Le choix de la rubrique des *Dépêches de nuit* s'est ainsi très vite imposé, car celle-ci constituait un terrain d'observation particulièrement propice pour notre projet de recherche. Tout d'abord, cette case est publiée sans interruption de 1883 à 1903 tandis qu'elle est pourvue d'un appareil titulaire permanent. Cette unité rédactionnelle présente de surcroît des conditions de production tout à fait stimulantes : à savoir que le locuteur responsable du titre et le rédacteur de la dépêche sont deux personnes séparées. En effet, comme le décrivent les travaux de M. Palmer (1983), les dépêches du XIX^e siècle sont rédigées par un correspondant ou un service de presse, que ce dernier envoie à tous les

quotidiens ayant souscrit un abonnement à ce service, indépendamment de leur sensibilité politique. Le secrétaire de rédaction de chaque quotidien se charge alors de titrer ces dépêches, qu'il ne réécrit que très rarement. Les signatures des dépêches et les archives du *Petit Comtois* ne manquent pas de confirmer ce fonctionnement rédactionnel. De 1883 à 1884, le quotidien recourt ainsi aux services d'un correspondant spécial, M. Sylvin. Le quotidien renonce aux services de ce dernier en 1884 suite à des difficultés financières : Jules Gros, séjournant alors à Paris s'occupe de la rédaction des brèves qu'il envoie quotidiennement à Besançon²⁰. Par la suite, le quotidien recourt aux services de l'agence Havas. Comme nous l'avons déjà signalé, la rubrique des *Dépêches de nuit*, exposée à la Une du *Petit Comtois*, correspond à une unité rédactionnelle axée sur le fait brut, par laquelle le quotidien se targue d'accomplir une mission d'information. Il devient dès lors très intéressant d'examiner comment *Le Petit Comtois*, quotidien d'opinion et de sensibilité républicaines, met en scène l'information brute des dépêches dont il n'est pas le rédacteur. Pour ces multiples raisons, les titres des *Dépêches de nuit* nous semblent ainsi un niveau textuel particulièrement sensible pour interroger les stratégies privilégiées par chaque quotidien.

3.3.3.1. Méthode de classification des titres

Notre classification des titres repose en premier lieu sur une approche fonctionnaliste des titres, que nous affinerons par une analyse sémantique, syntaxique et polyphonique des titres.

Notre démarche a pour clé de voûte méthodologique une distinction entre la *visée* explicite et première d'un titre, et sa (ses) *dimension* (s), entendue(s) comme la (les) caractéristique(s) non hiérarchiquement prédominante(s) du titre. Cette option méthodologique découle du postulat selon lequel le titre, en tant qu'objet discursif nécessairement complexe, peut témoigner d'une visée

²⁰ Lettre du 22 janvier 1884 de Jules Gros à un actionnaire inconnu, extraite des archives de la famille Millot.

explicite donnée et être simultanément marqué par d'autres dimensions. Nous considérerons que quatre visées peuvent prévaloir dans un titre journalistique du XIX^e siècle : il peut viser à commenter l'information de l'article, à amuser le lecteur, à (re)produire une information, ou à défaut, n'avoir d'autre visée première que de classifier l'information. Ce sont ces visées que représentent les quatre axes du schéma III-26. Notre grille de lecture se propose de donner à lire ces visées en termes de gradients, à partir desquels sera évaluée la fonction première du titre observé. Ainsi, sur le schéma III-26, chaque axe est à lire comme une échelle ayant à chacune de ses extrémités une valeur maximale et minimale. Cette première grille de lecture engendre la création de quatre macro-catégories de titres : les titres-index, les titres ludiques, les titres-commentaires, et les titres informationnels.

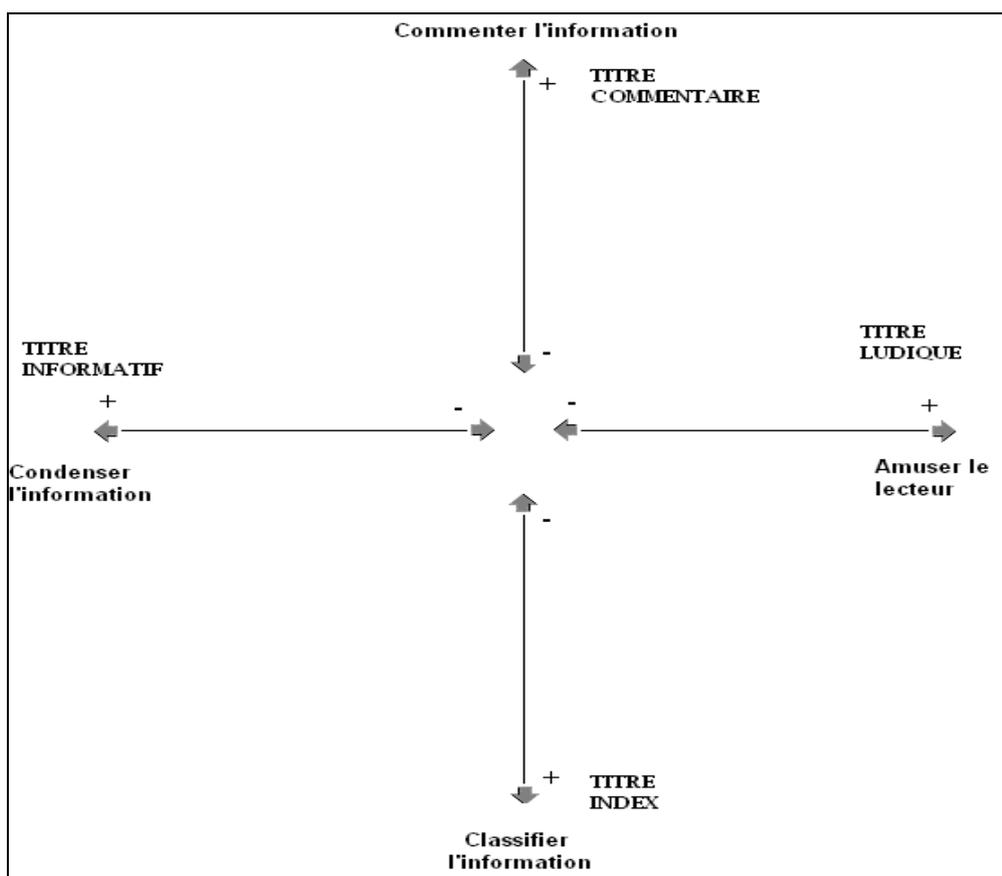


Figure III-26 : Les différentes visées du titre.

Notre analyse s'efforce toutefois de ne pas réduire le titre à une simple visée, qui impliquerait que la complexité de cet objet discursif soit figée et dénaturée. Dans cette perspective, nous recourons donc aux caractères non hiérarchiquement prédominants du titre, que nous entendrons comme ses dimensions, pour construire des catégories à l'intérieur des macro-catégories. Au-delà de sa visée première, le titre peut être chargé d'une ou de plusieurs dimensions : ainsi, un titre qui condense l'information principale d'un texte peut être chargé d'une dimension subjective explicite et/ou d'une dimension ludique. Nous distinguerons trois dimensions potentielles pour le titre : une dimension ludique, une dimension subjective, et une dimension informationnelle, par laquelle nous entendons la capacité d'un énoncé à livrer une information sur le monde.

Nous postulons que chacune des visées et des dimensions précédemment énumérées participe dans une mesure plus ou moins forte à alimenter la fonction incitative du titre, c'est-à-dire sa capacité à inviter le lecteur à lire l'article qu'il surplombe. Il s'agira donc pour nous d'observer comment la rédaction suscite l'intérêt du lecteur. Il nous semble cependant dès à présent essentiel d'observer que l'appareil titulaire des *Dépêches de Nuit*, du moins dans les premières années de parution, n'a pas initialement pour fonction d'assouplir le parcours de lecture fortement contraint du journal. Comme on peut l'observer sur la figure ci-dessous, les titres n'offrent pas systématiquement au lecteur la possibilité d'un parcours de lecture sélectif. Les éléments anaphoriques encadrés en rouge par nos soins montrent que les dépêches présupposent d'être lues dans l'ordre de leur succession :

<p>DÉPÊCHES DE NUIT (Service de notre correspondant spécial)</p> <p>Paris, 15 janvier, 7 h. 56 soir.</p> <p>Deuxième tirage de la loterie coloniale.</p> <p>Le numéro 1,001,165 gagne 50,000 francs.</p> <p>Les numéros 549 911 711,874 gagnent chacun 10 000 francs.</p> <p>Les numéros 33,807, 1,411,257 chacun 5,000 francs.</p> <p>Les numéros 191,841, 255,799, 340,915, 453,625, 548,087, 944,928, 1,129,625, 1,226,908, 1,243,159, 1,547,318 gagnent chacun 1,000 francs.</p> <p>Les numéros 428,666, 436,746, 537,948, 722,750, 806,686, 807,692, 1,029,642, 1,126,367, 1,213,716, 1,531,656 gagnent chacun 500 fr.</p> <p>Les salaires.</p> <p>La deuxième commission d'initiative a pris aujourd'hui en considération la proposition de M. Brugère, tendant à assimiler les salaires des ouvriers aux pensions, au point de vue de l'insaisissabilité d'une certaine quotité.</p> <p>Chambre des appels commerciaux.</p> <p>La même commission a pris également en considération la proposition de M. Faure, relative à la création de chambres d'appel commerciales et de chambres d'appel des décisions des prud'hommes.</p> <p>Rejet d'une proposition.</p> <p>Cette commission a repoussé la prise en considération d'une proposition de M. Le-fèvre-Pontalis, tendant à décider qu'aucune invalidation d'élection ne pourra être prononcée, sinon à la majorité des deux tiers des votants. M. Jules Gros, député du Doubs, a été nommé rapporteur.</p> <p>Sagesse de la commission.</p> <p>La deuxième commission n'a pas voulu statuer sur la proposition de MM. Laguerre et Camelinat, tendant à rechercher les ori-</p>	<p>gines de la guerre du Tonkin, et sur la mise en accusation des ministres.</p> <p>Cette commission a décidé qu'une convocation spéciale serait adressée à tous ses membres pour étudier cette affaire.</p> <p>Une prochaine réunion aura lieu lundi.</p> <p>Appréciation par la presse du message présidentiel.</p> <p>Tous les journaux républicains apprécient favorablement le message du président de la République. Certains journaux monarchistes sont grossiers. Le <i>Gaulois</i>, sous la signature de M. de Péne, connu autrefois pour sa courtoisie, dit que c'est là le message d'un pleutre.</p> <p>Nomination probable.</p> <p>La <i>Justice</i> dit que M. Thomson ne retournera pas en Cochinchine.</p> <p>Demande de révocation.</p> <p>Le même journal demande la révocation de M. Fouchon, secrétaire général du ministère de la justice, qui est allé trouver le ministre, accompagné de tous les directeurs du ministère, pour faire opposition à la nomination de M. Brousse, en qualité de sous-secrétaire d'Etat à la justice.</p> <p>Les Franc-Comtois à Paris.</p> <p>La <i>Nouvelle Revue</i> publie un long article de M. Rambaud sur un nouveau volume des <i>Lettres d'un proscrit</i>, publié par M^{me} Edgard Quinet.</p> <hr/> <p>LES ASSOCIATIONS OUVRIÈRES</p> <p>M. Yves Guvot, député de la Seine, a lu hier son rapport à la commission chargée d'examiner la proposition Nadaud, ayant pour objet d'étendre à certains travaux des villes les dispositions de la loi du 21 juin 1865 sur les associations syndicales.</p> <p>Ce rapport a été approuvé et sera déposé sur le bureau de la Chambre.</p> <p>La proposition réalise une utile réforme. La loi de 1865 ne s'appliquait qu'aux syndicats agricoles et, en outre, distinguait deux sortes de syndicats : les syndicats li-</p>
--	---

Figure III-27 : Extrait des « Dépêches de nuit » (1886-01-16), p. 1.

La macro-catégorie des titres-index

Nous traiterons en premier lieu des titres qui présentent la caractéristique de ne pas condenser le contenu de l'article qu'ils surplombent, mais de pré-catégoriser l'information en rattachant la singularité de celle-ci à une classe référentielle plus large. De fait, ces titres comptent de multiples occurrences au fil des numéros du journal. Résultant d'une opération de classification et non pas de condensation de l'information, les *titres-index* sont des entrées qui ne

prédisent du contenu de l'article que son horizon thématique ou géographique. Parce qu'ils renvoient à un découpage géographique ou thématique du monde en familles événementielles (Adam, 1997) ou à un procès historique, ces titres sont des liens vers des éléments contextuels et intertextuels. Cette catégorie de titre mobilise en effet un ensemble de référents, à entendre comme autant de connaissances encyclopédiques et culturelles (Eco, 1990), que le lecteur est supposé partager. Par là même, ces titres renvoient au journal en tant que collection, et donc à ses précédents numéros. Ce qu'il nous semble important de souligner, c'est que ces titres répondent essentiellement au besoin de la rédaction d'ordonner l'information. Le fonctionnement de ces titres est en cela similaire à celui des titres des rubriques. Pour éviter toute confusion avec le niveau du rubriquage et pour insister sur les fonctions de ces titres, nous référons aux titres précédemment définis en tant que *titres-index*.

Les titres (17) et (18), formés respectivement à partir des constructions syntaxiques *Préposition + Nom propre* et *Préposition + syntagme nominal* témoignent d'un système de classification à partir d'un paradigme géographique. Tous deux relèvent de la catégorie des titres-index puisqu'ils ne disent nullement l'information mais se bornent à en spécifier l'horizon géographique.

Ch.III (17), « A Souakim », (1889-01-04)

Londres, 3 Janvier.

On télégraphie de Souakim au *Morning Post* de ce matin que, dans l'après-midi d'hier, le général Grenfell a distribué des récompenses aux hommes des bataillons noirs, pour leur conduite courageuse à l'attaque des tranchées du mahdi.

Un déserteur venu de Handoub raconte qu'Osman Digma serait dans l'intention de se séparer de ses partisans, et qu'il aurait essayé de renvoyer son harem de Souakim.

Cette tentative a été empêchée par les patrouilles et les femmes et les enfants ont dû réintégrer Handoub.

Ch.III (18), « Au ministère de l'intérieur », (1888-01-03)

M. Duval, préfet de la Côte-d'Or, est nommé directeur de l'administration départementale et communale, en remplacement de M. Bourgeois, nommé récemment préfet de police.

Dépourvus de prédicat, (19) et (20) renvoient à un procès historique, dont la connaissance est présupposée et condensée par le substantif de tête. La valeur de catégorie de ces exemples est sensible au travers de l'absence de déterminant, traduisant un traitement antérieur du procès historique. Nous considérerons ainsi ces deux exemples comme des titres-index :

Ch.III (19), « Affaires d'Egypte », (1885-01-14)

Le *Standard* dit que la Russie et l'Allemagne refusent de garantir le nouvel emprunt égyptien.

Ch.III (20), « Événements d'Egypte », (1885-01-29)

Deux correspondants de journaux anglais à la suite de l'armée du général Wolseley ont été tués dans le combat du 19 janvier.

(21), (22), et (23), qui sont construits à partir d'un syntagme nominal précédé d'un déterminant défini, ont également valeur de renvois à des procès historiques en cours, devenus catégories événementielles. Le déterminant défini reflète la valeur anaphorique de ces titres à un précédent traitement dans les numéros antécédents du quotidien, que confirme la lecture des corps d'article. La première phrase de (21) présuppose ainsi elle-même des éléments contextuels disponibles dans l'intertexte, en l'occurrence le traitement de la crise ministérielle en Espagne dans le numéro précédent du *Petit Comtois* :

Ch.III (21), « La crise ministérielle en Espagne », (1902-03-15)

Madrid, 14 mars.

La situation n'a pas changée. La régente a commencé à conférer avec les chefs des différents groupes parlementaires et les anciens présidents des Chambres. Plusieurs listes ministérielles sont en formation, mais aucune n'est encore définitivement arrêtée.

On croit généralement que M. Sagasta conservera la présidence du conseil.

La régente désirerait que la crise reçoive une solution demain.

Ch.III (22), « La guerre avec la Chine », (1885-01-08)

Les agents diplomatiques et consulaires français de Chine réfugiés à Shanghai seront autorisés à quitter le sol de l'empire chinois.

Après avoir donné cette nouvelle, le *National* ajoute que c'est là un signe que les hostilités vont entrer dans une nouvelle période.

Ch.III (23), « L'élection de la Seine », (1899-01-11)

On voit sur les murs de Paris une affiche portant ces mots : « Général Boulanger. — Dissolution, revision, une Constituante. » C'est probablement l'affiche plus

spécialement destinée aux réactionnaires, car on y cherche vainement le mot de République.

Enfin, l'appartenance de (24) et (25) à la macro-catégorie des titres-index se justifie par le fait que ces énoncés ne résument pas tant l'information de l'article qu'ils rattachent la singularité de celle-ci à une famille événementielle. L'absence de déterminant traduit ici le non embrayement de ces énoncés sur le monde, au sujet duquel ils ne disent rien. On notera de surcroît l'absence de toute prédication spatio-temporelle dans ces titres. La fonction première de ces titres est référentielle : ils indiquent de quelle famille événementielle relève l'information que l'article introduit. La récursivité de ces titres dans notre corpus confirme leur statut d'index.

Ch.III (24), « Tremblement de terre », (1902-03-13)

Constantinople, 12 mars.

Un terrible tremblement de terre a été ressenti à Kiankari (Turquie d'Asie). 3.000 maisons de la ville, qui compte 20,000 habitants, ont été détruites.

Deux musulmans et deux chrétiens ont été tués ; il y a plus de cent blessés.

Ch.III (25), « Voyage ministériel », (1888-01-17)

Paris, 16 janvier. M. Loubet, ministre des travaux publics, appelé par un deuil de famille, s'est rendu dimanche à Montélimar. Il sera de retour mardi matin à Paris pour assister au conseil des ministres.

De même, les énoncés (26), (27) et (28) prédisent « de quoi » parle l'article. Sur le plan sémantique, nous concédons, à la suite de J. de Bonville, que le substantif de chacun de ces exemples évoque un schéma actanciel, portant en germe une certaine valeur informationnelle. Néanmoins, celle-ci est compromise par l'actualisation incomplète du substantif liée à l'emploi d'un déterminant indéfini et à l'absence de toute prédication spatio-temporelle. Ces titres catégorisent la singularité de l'information de l'article par le renvoi à une famille événementielle. La multiplicité des occurrences de ces titres vient encore confirmer leur statut d'index.

Ch.III (26), « Une insurrection », (1899-01-01)

[...] Le *New-York Herald* publie une dépêche d'El Paso, d'après laquelle une foule immense, conduite par des prêtres, aurait attaqué, vendredi dernier, le palais présidentiel à Mexico. Elle aurait été repoussée par les troupes du gouvernement

qui auraient fait 2,000 prisonniers.72 prêtres auraient été tués ; 200 autres, au nombre desquels serait l'archevêque, auraient été arrêtés le lendemain et l'ordre aurait été donné de les fusiller. On croit que cet ordre sera exécuté malgré les prières des femmes.

Ch.III (27), « Un cyclone », (1899-01-11)

New-York, 10 janvier.

Un cyclone vient de passer sur Reading, brisant tout sur son passage.

Une filature de soie a été renversée par le vent. Non loin de cette filature, une explosion suivie d'incendie se produisait, tuant huit personnes. [...]

Ch.III (28), « Un décès », (1903-03-02)

6 h. 55 soir.

Le vice-amiral Roustan, préfet maritime de Brest, est mort cette nuit à Paris.

Les titres-index à dimension subjective

Selon nous, une dimension subjective explicite est compatible avec le fonctionnement des titres-index. Certains titres, à l'instar de (29), renvoient anaphoriquement à un procès historique et journalistique, tout en portant les traces d'une présence explicite de la subjectivité du locuteur, ici sensible par l'emploi d'un adverbe de gradation quantitative. De même, (30) et (31), qui exemplifient la classe des titres-index caractérisés par l'absence de prédication spatio-temporelle et par une actualisation incomplète, témoignent d'une forte dimension subjective, qui est explicitée par des adjectifs affectifs.

Ch.III (29), « Encore la grève générale », (1902-03-06)

Alais, 5 mars.

Le congrès national des mineurs a adopté, dans sa séance de ce matin, une importante résolution. En voici le texte :

Le congrès poursuivra les décisions du congrès de Lens par la grève générale immédiatement, et sans nouvelles négociations avec le gouvernement.

La première partie a été votée à l'unanimité. La seconde partie, à partir du mot «immédiatement », a été adoptée par 124 voix contre 105, après une vive discussion.

Ch.III (30), « Terrible incendie », (1900-01-14)

Un incendie terrible s'est déclaré à Newark (New Jersey), dans une grande maison occupée par de nombreuses familles italiennes.

Quinze personnes, en majorité des enfants, ont péri. On croit que le sinistre est dû à la malveillance.

Ch.III (31), « Lugubre découverte », (1900-03-24)

Une dépêche de New-York au *Daily Telegraph* annonce que le cadavre d'une jeune femme, nommée Boon, a été trouvé carbonisé et attaché à un poteau près d'une ville de la Caroline.

Quatre blancs ont été arrêtés sous l'inculpation d'être les auteurs de ce crime.

La macro-catégorie des titres informationnels

Nous regrouperons sous la macro-catégorie *titres informationnels* les titres (re)produisant une information de l'article. Nous distinguerons à l'intérieur de cette macro-catégorie les titres pleinement informationnels des titres à dimension informationnelle.

Les titres pleinement informationnels

Par *titres pleinement informationnels*, nous référerons aux titres fondamentalement singuliers et autonomes sur le plan référentiel qui (re)produisent l'information et (pré)disent le contenu de l'article.

Appartiennent à cette classe les exemples (32), (33) et (34) qui, sur le plan syntaxique, correspondent à des syntagmes nominaux²¹ pourvus d'une prédication spatiale ((32) et (33)) ou d'un complément du nom actanciel (34). L'absence de déterminant reflète ici le caractère initial d'un procès historique, qui se vérifie à la lecture de l'article.

Ch.III (32), « Explosion de grisou en Allemagne », (1889-01-05)

Une explosion de grisou, qui a tué vingt-sept ouvriers et blessé un grand nombre d'autres, a eu lieu dans la mine d'Esperanza, dans la province d'Oviédo.

Ch.III (33), « Révocation du Maire de Carcassonne », (1899-01-10)

M. Jourdane, maire de Carcassonne, déjà suspendu par l'autorité supérieure, vient d'être révoqué.

Ch.III (34), « Mort d'un agent de change », (1902-03-13)

Agen, 12 mars.

M. Emmanuel Chaumié, agent de change à Agen, frère de M. Chaumié, sénateur de Lot-et-Garonne, est mort hier soir à l'âge de cinquante-sept ans.

Entre autres occurrences de titres informationnels, on citera les exemples (35) et (36). Ces derniers se caractérisent par l'emploi de la forme passive, résultant

²¹ Issus du passage d'une construction verbale à la structure nominalisée correspondante.

de la transformation d'une proposition de l'article (35), ou de la reprise d'un syntagme verbal à la forme passive de l'article avec mise en ellipse de l'auxiliaire (36). On notera que si (36) est marqué par une réduction des circonstances spatiales et temporelles par rapport à l'article, il apporte néanmoins une nouvelle information *via* l'ajout de l'adjectif « vivants ». Dans les deux exemples suivants, le titre correspond ainsi à une extraction et à une condensation de l'information principale délivrée dans l'article. Aussi, leur catégorisation en tant que titre informationnel nous semble ainsi se justifier.

Ch.III (35), « Mordue par un rat », (1897-02-01)

On a, ces jours derniers, à Nouzat (Dordogne), enterré une pauvre femme, Mme D..., qui est morte des suites de la morsure d'un rat.

Ch.III (36), « Deux hommes ensevelis vivants », (1899-01-05)

Un éboulement s'est produit ce matin à la sablonnière de Boursay. Deux hommes ont été ensevelis. Les travaux de sauvetage ont commencé ; on espère retirer les victimes vivantes.

De même, les énoncés (37) et (38) condensent l'information majeure délivrée dans l'article, épurée néanmoins de tout complément circonstanciel, tout en présentant la particularité d'être formulés par une phrase verbale complète :

Ch.III (37), « Les sucriers se syndiquent », (1897-02-01)

Les représentants de 151 fabriques de sucre viennent de fonder, à Berlin, la société du syndicat des sucres allemands, à responsabilité limitée.

Ch.III (38), « Le Rhône décroît », (1897-02-06)

Lyon. — Le Rhône décroît sensiblement.

Les titres pleinement informationnels à dimension subjective

Condenser l'information principale de l'article, en tant qu'elle procède d'une opération de sélection, n'est pas une pratique dénuée de subjectivité. Cette subjectivité peut également être plus ou moins explicite à la surface du discours. Ce que confirme l'énoncé (39), dont la valeur informationnelle s'accompagne d'une dimension subjective. Cette dernière est sensible à travers la désignation par métonymie de l'Empire d'après le nom de la porte d'honneur

monumentale du Grand Vizirat à Istanbul, et par l'utilisation d'une expression de style familier (« mettre les pouces »).

Ch.III (39), « La sublime Porte met les pouces », (1900-01-22)

Le sultan vient de céder aux demandes russes sur tous les points essentiels. Un iradé va être publié incessamment.

Huit dames du palais ont été exilées pour machinations avec Mahmoud pacha, le beau-frère du sultan, actuellement en fuite. [...]

Les titres à dimension informationnelle

Nous nommerons *titre à dimension informationnelle* tout titre doté d'une certaine valeur informationnelle, néanmoins compromise par un manque d'autonomie référentielle et/ou un défaut d'actualisation. Sur le plan syntaxique, cette classe d'énoncés s'accomplit tout particulièrement par des phrases déclaratives ou exclamatives.

Les énoncés (40), (41), (42) et (43) à la différence des titres répondant à la même construction syntaxique regroupés dans la classe des titres informationnels, se caractérisent par un défaut d'autonomie référentielle. Ces quatre énoncés ont une valeur informationnelle certaine, reposant sur l'ajout d'apport notionnel supplémentaire à propos du sujet, c'est-à-dire sur un couple thème/rhème. Cependant, ces énoncés ne peuvent être considérés en toute rigueur comme informationnels du fait de leur défaut d'autonomie, postulé comme trait définitionnel de cette catégorie. La compréhension de ces énoncés est en effet dépendante de la lecture de l'ensemble textuel plus large qu'ils condensent, en raison de l'emploi d'expressions déictiques. Pour connaître l'affaire à laquelle renvoie le sujet composé d'un syntagme nominal introduit par un déterminant défini cataphorique de (40), le lecteur doit en effet se reporter au texte-article. On notera cependant que l'article lui-même ne livre pas d'éléments sur la nature des faits ayant engendré la procédure signalée, et qu'il sollicite ainsi de la part du lecteur des éléments contextuels disponibles dans l'intertexte. De même, pour identifier les auteurs et les circonstances de « l'invitation » dans (41), il est nécessaire de prendre connaissance de l'article. En ce qui concerne (42) et (43), l'identification du référent des pronoms

personnels cataphoriques de troisième personne n'est accessible que par la lecture des articles correspondant. Nous formulons l'hypothèse que la valeur référentielle déficiente de ces énoncés est un procédé qui renforce la fonction incitative du titre.

Ch.III (40), « L'enquête continue », (1903-03-02)

M. Leydet, juge d'instruction a confronté cet après-midi Mme Humbert avec Parmentier.

Ch.III (41), « Le président accepte l'invitation », (1902-03-15)

Le président de la République a envoyé, hier soir, sa réponse au tsar Nicolas II par laquelle il accepte l'invitation de se rendre en Russie.

Ch.III (42), « Ils partent pour la Crète ! », (1897-02-13)

Le cuirassé <sic> Dévastation et le croiseur Bugeaud sortiront, de Toulon le 16 courant et mettront en disposition de rallier le pavillon du contre-amiral Pothier, à la Canée. Le cuirassé grec Speizià est parti ce soir pour la Canée. On se tient prêt à recevoir de nouveaux ordres dans l'arsenal de Toulon.

Ch.III (43), « Ils ont capturé plusieurs fermiers », (1900-01-02)

Rensburg, 28 février

Londres. — Lord Roberts télégraphie :

« Paardeberg, 28 févr. 8 h. 50 soir.

Le général Clément annonce qu'ayant appris que Colesberg avait été évacué, il se mit en marche pour l'occuper et entra à Colesberg, où il fut accueilli avec enthousiasme.

Il s'est emparé de quantité de munitions et a capturé plusieurs fermiers.

Il est retourné ensuite à Rensburg ; il annonce que la voie ferrée est libre.

La macro-catégorie des titres-commentaires

Relève de cette classe tout titre au travers duquel le locuteur exprime une opinion sur l'information de l'article sans prédire celle-ci. Ce qui transparaît à leur lecture ne relève pas de l'assertion d'un horizon d'attente, pas plus qu'il ne s'apparente à un énoncé à valeur informationnelle. Tout au contraire, le titre-commentaire, chargé d'une dimension subjective et argumentative fortes, exprime une position socio-discursive.

Ainsi, les titres (44), (45) et (46) n'apportent aucune information sur le monde et ne précisent nullement l'horizon d'attente thématique de l'article. Ces énoncés titulaires partagent cependant la caractéristique d'exprimer l'opinion de la rédaction à l'égard des événements présentés dans les brèves qu'ils

surplombent. L'énoncé (44) hiérarchise ainsi l'importance de la dépêche en signalant, par l'emploi d'un syntagme hautement axiologique, l'intérêt de l'information qu'elle délivre, à savoir que l'administration allemande procède à la fabrication des mêmes types d'armes qu'utilise la France. Dans le titre (45), le locuteur critique la ligne politique de *L'Intransigeant*, organe rallié au boulangisme et au nationalisme, en retournant contre lui les accusations précédemment adressées au ministre de l'Intérieur, qui aurait entraîné l'aliénation mentale de Louise Michel. En décrivant avec sarcasme la politique de cette feuille par un complément du nom relevant de l'isotopie des désordres mentaux dont les connotations sont éminemment péjoratives, *Le Petit Comtois* critique la position politique de la feuille parisienne. Dans (46), le locuteur dénie toute efficacité à l'adoption d'un serment politique par les membres du Reichstag dont la proposition est l'information principale délivrée par la brève correspondante.

Ch.III (44), « Bon à constater », (1903-03-16)

On mande de Berlin au *XIXe Siècle* :

Obéissant à un mot d'ordre, les journaux militaires ne cessent de répéter que le système français de protéger les canons par une cuirasse est inutilisable; ils contestent également la valeur du frein adopté en France.

Pendant ce temps, l'administration militaire allemande fait fabriquer dans le plus grand silence des canons munis de cuirasses d'une épaisseur de 4 millimètres et d'un frein qui n'est qu'une imitation plus ou moins bien réussie de celui de l'artillerie française. Il va de soi que la commande a été donnée à la maison Krupp.

Ch.III (45), « Politique de névropathes », (1885-01-04)

L'Intransigeant continue ses attaques contre les hommes du gouvernement.

Il accuse aujourd'hui le ministre de l'intérieur d'être cause que Louise Michel est atteinte d'aliénation mentale.

Ch.III (46), « Précautions inutiles », (1883-08-31)

La *Gazette de Francfort* dit qu'à la suite de l'affaire du député de Metz, M. Antoine, le gouvernement proposera l'adoption du serment politique pour les membres du reichstag, sur le modèle de ce qui se passe pour le landtag prussien.

La macro-catégorie des titres ludiques

Enfin, nous distinguerons les *titres à dimension ludique* et les *titres ludiques*. Nous entendrons par ludique toute instauration d'un jeu entre le

journal et le lecteur, pouvant concrètement s'opérer à travers un jeu avec la langue, l'interdiscours ou le co-texte. Il nous semble en effet essentiel de dépasser une conception du ludique rivée aux seuls jeux de mots alors que les titres observés dans notre corpus et les brèves correspondantes ne relèvent pas du même énonciateur. D'où un décalage énonciatif exploité par le locuteur chargé des titres, que nous assimilerons à un jeu avec le lecteur. Si un énoncé ne vise qu'à être ludique au détriment des autres fonctions, il sera qualifié comme tel. Précisons, qu'à nos yeux, le ludique n'est pas systématiquement associé à une charge humoristique, qui peut être plus ou moins existante.

Les titres (47) et (48) illustrent la macro-catégorie des titres ludiques. Ces énoncés ne servent aucune visée argumentative ni ne tendent à apporter une quelconque information sur le monde. L'énoncé (47) se limite à un jeu de mot, allusion à l'utilisation de la locution figée « le théâtre des opérations » utilisée dans la brève, qui peut d'ailleurs être envisagée comme une métaphore guerrière. Le titre (48) repose également sur le défigement d'une locution figée, par la substitution d'un nom propre homonyme du substantif « diète » tel qu'il apparaît dans la locution. Visant ici à définir les catégories des titres, nous n'approfondirons pas l'analyse du défigement et ne présenterons pas maintenant la réflexion existant autour de cette figure, telle que développée notamment par P. Fiala et B. Habert (1989) et par F. Sullet-Nylander (1998). Nous aborderons ultérieurement cette question dans le cadre de l'examen des différents ressorts du ludique et de leurs valeurs pragmatiques dans notre corpus.

Ch.III (47), « A chacun sa manière d'aller au théâtre », (1900-01-16)

Pretoria, 15 mars. Le général Joubert est parti hier soir pour le théâtre des opérations.

Ch.III (48), « On va les mettre à la Diète », (1900-03-21)

La *Gazette officielle* publiera demain la convocation de la Diète, dont la réunion est fixée au 26 mars.

Les titres à dimension ludique

Si le ludique n'intervient que pour orner ou affirmer la visée du titre, qu'elle soit de commenter ou de condenser l'information, nous parlerons alors de *titre à dimension ludique*. Pour clarifier notre propos, considérons l'énoncé (49) pourvu d'une indéniable dimension ludique : il contient en effet un jeu de mot sur la polysémie de la graphie « orange », à la fois nom de pays et substantif, et sur la diasémie de « pépin ». Cet énoncé peut en effet être lu sur une double isotopie //fruit// et //complications//. Au-delà de cet aspect ludique, ce titre vise principalement à commenter l'information délivrée par la brève, en l'occurrence l'avancée des troupes britanniques dans l'État d'Orange. De fait, nous envisagerons un tel énoncé comme un titre-commentaire à dimension ludique.

Ch.III (49), « Méfiez-vous de l'orange sans pépin », (1900-01-18)

Wan-Zyl, 16 mars.

Les troupes du général Clément se sont avancées dans l'Etat d'Orange jusqu'à une distance considérable, sans rencontrer aucun Boer.

Tous les Boers semblent s'être retirés du front des troupes anglaises.

Les communications sont rétablies entre Bloemfontein et Norvais-Pont.

Simulant l'oralité dans l'expression de la négation, le titre (50) relève très nettement de la macro-catégorie des titres-commentaires. Il vise à critiquer la politique d'expansion coloniale de l'Allemagne. La visée argumentative de cet énoncé est amplifiée par une tonalité humoristique reposant sur l'emploi d'un élément d'ordre stéréotypique, à savoir une référence d'ordre culinaire. Par la mise en contraste d'un élément alimentaire symbole de l'exotisme et du raffinement au XIX^e siècle et d'un plat culinaire traditionnel populaire stéréotypé, (50) dénie au peuple allemand toute prétention à recevoir une exposition coloniale.

Ch.III (50), « C'est pas des figues, c'est de la choucroute qu'il leur faut », (1902-03-14)

Berlin, 13 mars.

On organise, sous les auspices du duc de Mecklenbourg, une exposition coloniale qui se tiendra à Berlin au mois d'octobre ; elle sera précédée d'un congrès colonial.

3.3.3.2. Aperçu sur les caractéristiques formelles générales des titres en diachronie

À présent, notre propos est de dresser une première esquisse de classification des titres des *Dépêches de nuit*. À cet effet, nous décrivons leurs caractéristiques formelles générales, indépendamment de leur macro-catégorie d'appartenance.

Des titres très courts

Examinons tout d'abord la longueur moyenne des titres, année par année, en prenant pour unité de mesure le mot. *Grosso modo*, la longueur moyenne du titre tend à être limitée à 3,5 mots sur l'ensemble de la tranche diachronique (cf. tableau III-1). Bien que les indications chiffrées présentées ci-dessous ne soient que des moyennes, elles révèlent néanmoins le très court format des titres. À titre d'information, nous préciserons qu'un titre du journal contemporain, tel que *Le Figaro*, compte à peu près 8 mots en 1997, selon une étude réalisée par F. Sullet-Nylander (1998 : 37). Si on note un accroissement de la longueur des titres de façon assez régulière entre 1896 à 1900, il se révèle hasardeux de donner une signification à l'évolution de la longueur des titres sur la période 1883-1895, dont les oscillations empêchent toute esquisse de tendance.

	1883	1884	1885	1886	1887	1888	1889
Longueur moyenne(en nombre de mots)	3,9	3,9	3,7	3,8	3,8	3,6	3,4
	1890	1891	1892	1893	1894	1895	1896
Longueur moyenne(en nombre de mots)	3,4	2,9	3,7	3,2	3,4	3,7	3,5
	1897	1898	1899	1900	1901	1902	1903
Longueur moyenne(en nombre de mots)	3,6	3,9	3,7	4,1	4	4	4,1

Tableau III-1 : Longueur moyenne des titres (en nombre de mots).

Des titres reposant majoritairement sur un syntagme nominal

Considérons à présent la structure syntaxique des titres des *Dépêches de nuit*. La nominalisation, se révèle très nettement le mode privilégié de

construction des titres observés dans notre corpus. D'un point de vue général, le corpus ne compte que très peu de phrases verbales complètes, correspondant *a minima* au prototype *sujet+ verbe*. Si l'on observe la figure III-24, retraçant l'évolution du pourcentage de phrases verbales dans le corpus, année par année, il s'avère en effet que les phrases verbales ne caractérisent que moins d'1/5 des titres de l'année 1900 où figurent pourtant le taux le plus élevé de tels titres.

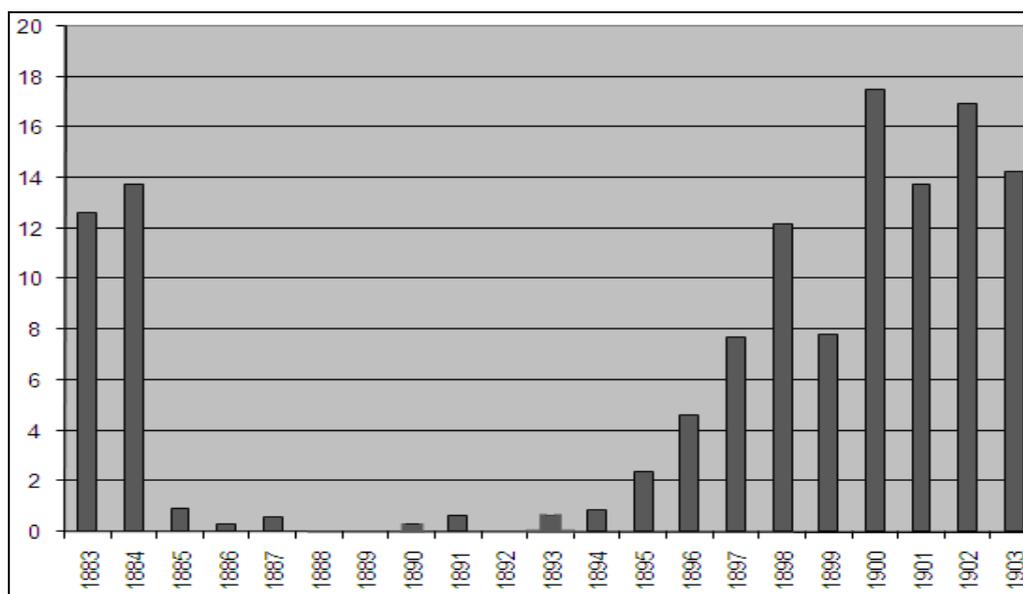


Figure III-28 : Les phrases verbales complètes en diachronie (en pourcentage).

Il nous semble intéressant de noter le contraste entre d'une part, les deux premières années de parution du quotidien et la période 1895-1903, et d'autre part la tranche diachronique 1885-1894, où l'utilisation des phrases verbales se révèle nulle ou extrêmement faible (0,3% en 1886, 0,27% en 1890).

Après cet examen des caractéristiques de surface, il convient d'approfondir nos résultats en procédant à une analyse croisée des faits de nature syntaxique avec les fonctions discursives des titres.

3.3.3.3. Répartition des fonctions des titres dans le sous-corpus

Pour analyser l'évolution des titres dans notre corpus, chacune des 6910 occurrences a été catégorisée comme relevant d'une macro-catégorie, et si le cas se présentait, d'une dimension dominante. Cet inventaire a ensuite été établi sous forme de pourcentage, pour plus de lisibilité, et rassemblé dans le tableau III-2.

La macro-catégorie des titres-index prédomine très nettement dans l'ensemble du corpus. Les titres-commentaires viennent en second, loin devant les titres-informationnels. L'emploi des titres à dimension informationnelle et surtout des titres ludiques est beaucoup plus rare. En diachronie, deux grandes tendances sont perceptibles. La première concerne l'évolution des titres-index : l'utilisation de cette macro-catégorie est extrêmement massive de 1883 à 1892, où elle représente près de 80% en moyenne, pour ensuite s'amoinrir régulièrement jusqu'à ne plus atteindre 50% de 1899 à 1903.

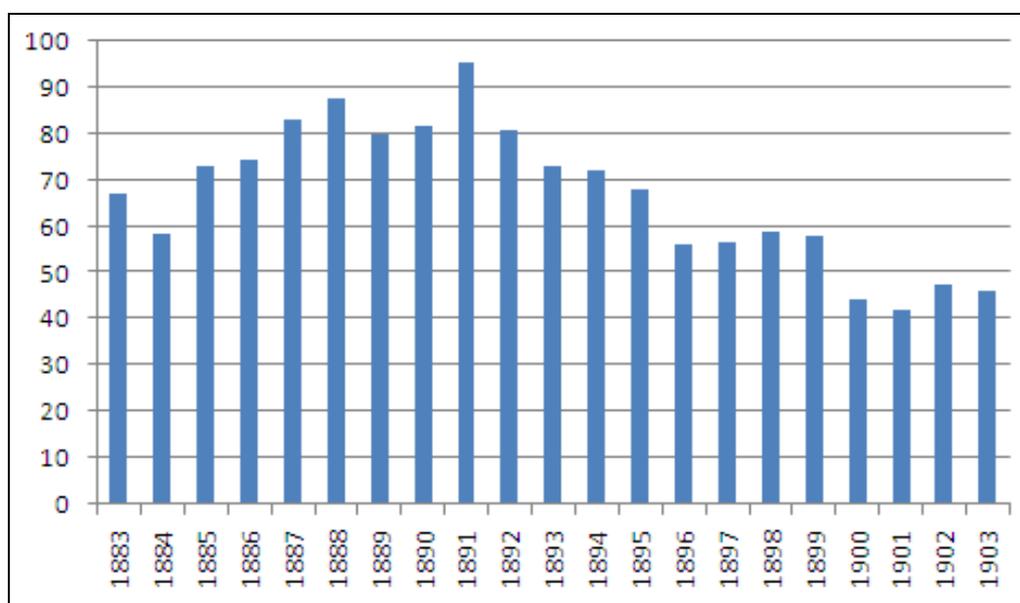


Figure III-29 : Pourcentage de titres-index (en diachronie).

La tendance est globalement inverse pour les titres-commentaires : après 1885, ils sont assez rares jusqu'en 1895, puis s'imposent sur la dernière tranche diachronique 1896-1903, jusqu'à représenter 31,7% des titres en 1897, 43,1% en 1900 et 35,4% en 1903.

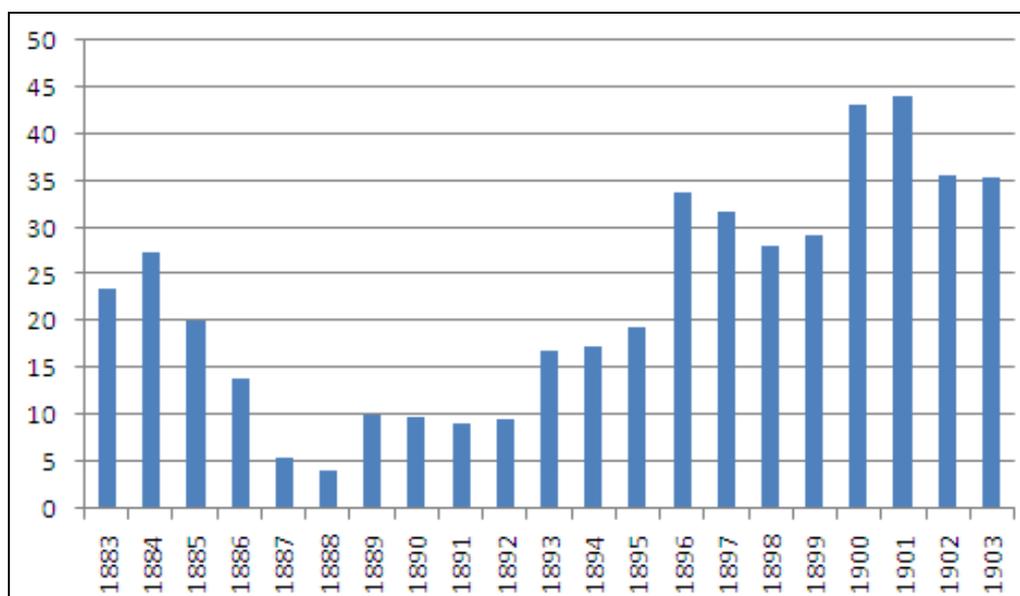


Figure III-30 : Pourcentage de titres- commentaires (en diachronie).

L'évolution des titres à dimension informationnelle et titres informationnels est beaucoup moins nette. Plus encore, il nous semble délicat de lui donner un sens. Si l'on considère l'évolution des titres informationnels tout d'abord (figure III-31), aucune tendance lourde ne s'esquisse : après avoir avoisiné 14% lors de la première année de parution du quotidien, le pourcentage de titres informationnels ne cesse d'osciller sur la période 1884-1901 : 5,2% en 1885 puis 10% en 1886, près de 9, 5% en 1889 et 7,2 % en 1890 pour ne représenter que 5,7% de l'ensemble des titres en 1891, etc. Si l'on observe que deux des trois plus fortes fréquences des titres-commentaires se concentrent sur les années 1902 et 1903, nous nous garderons bien d'analyser ces données comme une tendance significative.

Comme on peut l'observer sur la figure III-32, l'emploi des titres à dimension informationnelle est tout aussi fluctuant en diachronie : de 1883 à 1899, le

pourcentage de titres à dimension informationnelle passe de 0,9% en 1883 à 6,3% en 1884, pour ne pas excéder les 1,9% en 1885. Le pourcentage d'emploi de ce type de titre tend à se stabiliser sur la tranche diachronique 1900-1903, représentant entre 4,2% en 1900 et 5,5% en 1903.

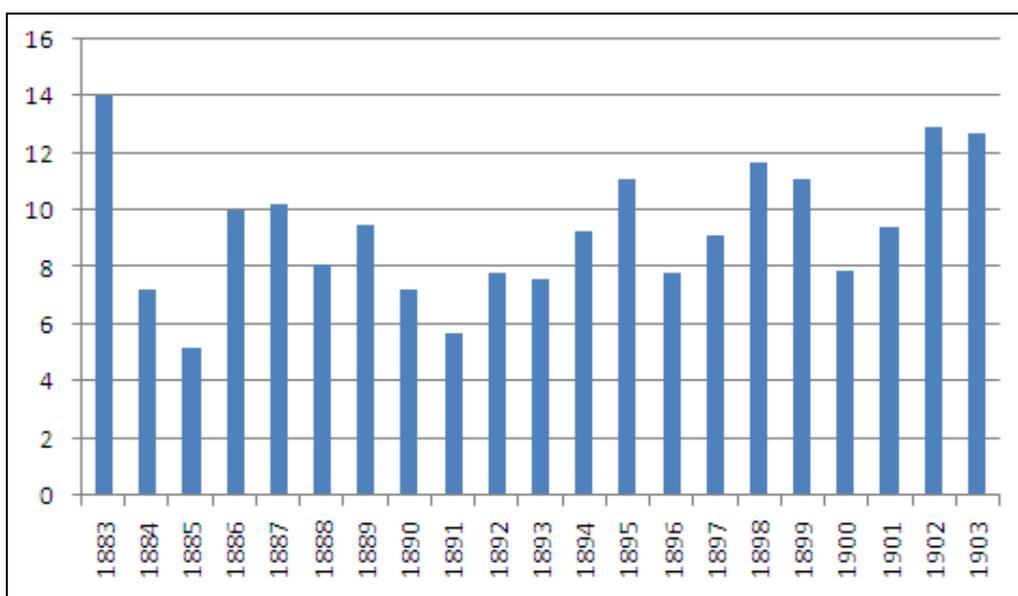


Figure III-31 : Pourcentage de titres informationnels (en diachronie).

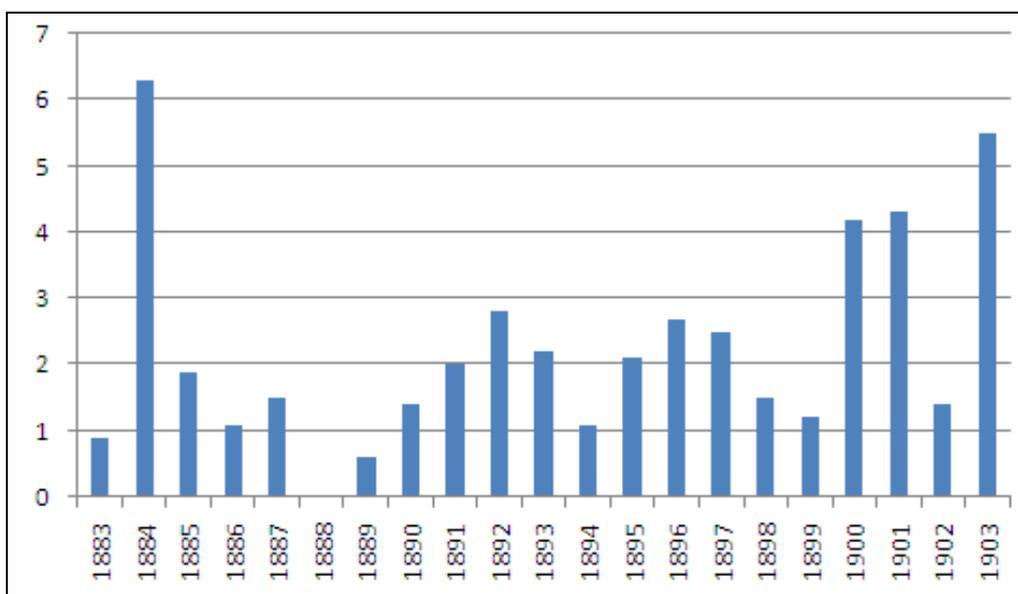


Figure III-32 : Pourcentage de titres à dimension informationnelle (en diachronie).

	1883	1884	1885	1886	1887	1888	1889
INFORMATIONNEL	47=14%	24=7,2%	16=5,2%	28=10%	28=10,2%	22=8,1%	30=9,5%
Dont à dimension subjective	1=2,1%	1=4,2%	1=6,2%	1=3,6%	0	0	0
DIMENSION INFORMATIONNELLE	3=0,9%	21=6,3%	6=1,9%	3=1,1%	4=1,5%	0	2=0,6%
Dont à dimension ludique (%)	1=33,3%	2=9,5%	1=16,6%	0	0	0	0
Dont à dimension subjective	1=33,3%	1=4,8%	0	0	0	0	0
INDEX	225=67,2%	194=58,6%	225=72,8%	208=74,5%	226=82,8%	238=87,5%	253=79,8%
Dont à dimension subjective (%)	5=2,2%	15=7,7%	1=0,4%	3=1,4%	3=1,3%	0	5=2%
Dont à dimension ludique (%)	1=0,4%	6=3%	2=0,9%	0	0	0	3=1,2%
COMMENTAIRE	79=23,6%	91=27,5%	62=20%	39=14%	15=5,5%	11=4%	32=10,1%
Dont à dimension ludique	8=10,1%	11=12,1%	4=6,4%	2=5,1%	3=20%	1=9,1%	3=9,4%
LUDIQUE	0						
	1890	1891	1892	1893	1894	1895	1896
INFORMATIONNEL	26=7,2%	17=5,7%	17=7,8%	24=7,6%	33=9,3%	37=11,1%	23=7,8%
Dont à dimension subjective	0	0	0	0	0	0	0
DIMENSION INFORMATIONNELLE	5=1,4%	6=2%	6=2,8%	7=2,2%	4=1,1%	7=2,1%	8=2,7%
Dont à dimension ludique (%)	0	0	0	0	0	0	2=25%
Dont à dimension subjective	0	0	1=16,6%	0	0	0	1=12,5%
INDEX	296=81,5%	253=85,5%	175=80,6%	232=73,2%	257=72,2%	225=67,8%	165=56,1%
Dont à dimension subjective (%)	8=2,7%	2=0,8%	3=1,7%	3=1,3%	10=3,9%	9=4%	6=3,6%
Dont à dimension ludique (%)	3=1%	1=0,4%	2=1,1%	4=1,7%	0	0	2=1,2%
COMMENTAIRE	36=9,9%	27=9,1%	21=9,7%	54=17%	62=17,4%	64=19,3%	99=33,7%
Dont à dimension ludique	6=16,6%	0	2=9,5%	7=13%	10=16,1%	10=15,6%	25=25,2%
LUDIQUE	0						
	1897	1898	1899	1900	1901	1902	1903
INFORMATIONNEL	36=9,1%	46=11,7%	36=11,1%	34=7,9%	33=9,4%	49=12,9%	44=12,7%
Dont à dimension subjective	3=8,3%	3=6,5%	5=13,9%	2=5,9%	1=3%	4=8,2%	2=4,5%
DIMENSION INFORMATIONNELLE	10=2,5%	6=1,5%	4=1,2%	18=4,2%	15=4,3%	16=4,1%	19=5,5%
Dont à dimension ludique (%)	1=10%	1=20%	0	0	0	0	1=5,2%
Dont à dimension subjective	1=10%	1=20%	0	1=5,5%	1=6,6%	2=12,5%	1=5,2%
INDEX	225=56,7%	230=58,7%	188=58%	190=44,3%	147=42,1%	181=47,4%	160=46,1%
Dont à dimension subjective (%)	4=1,8%	15=6,5%	9=4,8%	8=4,2%	12=8,2%	2=1,1%	12=7,5%
Dont à dimension ludique (%)	1=0,4%	4=1,7%	1=0,5%	2=1%	1=0,7%	9=5%	1=0,6%
COMMENTAIRE	126=31,7%	110=28,1%	95=29,3%	185=43,1%	154=44,1%	136=35,6%	123=35,4%
Dont à dimension ludique	14=11,1%	27=24,3%	19=20%	36=19,4%	15=9,7%	24=17,7%	22=17,9%
LUDIQUE	0	0	1=0,3%	2=0,5%	0	0	0

Tableau III-2 : Répartition des titres selon leur fonction en diachronie.

Affinons à présent les premières remarques relatives aux caractéristiques formelles des titres en croisant les fonctions des titres et leurs caractéristiques syntaxiques. Pour faciliter la lecture de nos observations, nous décrirons ici successivement les caractéristiques inhérentes à chaque macro-catégorie.

Les titres-commentaires

Une macro-catégorie recourant dans une mesure significative aux phrases verbales complètes

Le tableau III-3 représente la ventilation des phrases verbales complètes selon la macro-catégorie des titres dans lesquelles elles s'inscrivent. Il pointe une forte corrélation entre les phrases verbales et les titres-commentaires. Mis à part en 1883 et en 1884, où les titres informationnels et les titres à dimension informationnelle recourent également aux phrases verbales complètes, on ne rencontre ces dernières que dans les titres commentaires dans la première tranche diachronique du corpus. À partir de 1896, si toutes les macro-catégories de titres, à l'exception des titres-index, tendent à employer des phrases verbales complètes, les titres-commentaires restent la catégorie où cette construction est la plus largement employée.

Année	1883	1884	1885	1886	1887	1888	1889
TITRE A DIMENSION INFORMATIONNELLE	6,7	15,2	0	0	0	0	0
TITRE COMMENTAIRE	60	65,3	100	100	100	0	0
TITRE INDEX	0	0	0	0	0	0	0
TITRE INFORMATIONNEL	33,3	19,5	0	0	0	0	0
TITRE LUDIQUE	0	0	0	0	0	0	0
Année	1890	1891	1892	1893	1894	1895	1896
TITRE A DIMENSION INFORMATIONNELLE	0	0	0	0	0	0	6,7
TITRE COMMENTAIRE	100	100	0	100	100	100	86,6
TITRE INDEX	0	0	0	0	0	0	0
TITRE INFORMATIONNEL	0	0	0	0	0	0	6,7
TITRE LUDIQUE	0	0	0	0	0	0	0
Année	1897	1898	1899	1900	1901	1902	1903
TITRE A DIMENSION INFORMATIONNELLE	15,6	6,3	4	12,7	13,7	12,9	14,8
TITRE COMMENTAIRE	68,8	74,9	61,5	77,5	76,3	75,7	70,3
TITRE INDEX	0	0	0	0	0	0	0
TITRE INFORMATIONNEL	15,6	18,8	34,5	8,4	8,5	11,4	14,8
TITRE LUDIQUE	0	0	0	1,4	1,5	0	0

Tableau III-3 : Répartition des phrases verbales complètes par macro-catégorie (en pourcentage).

Des phrases verbales majoritairement conjuguées au présent

Examinons les temps et les modes employés dans ces phrases verbales complètes en observant le tableau ci-dessous.

	1883	1884	1885	1886	1887	1888	1889
PRESENT	92,5	80	100	100	100	0	0
FUTUR	7,5	10	0	0	0	0	0
PASSE-COMPOSE	0	3,4	0	0	0	0	0
IMPARFAIT	0	6,6	0	0	0	0	0
CONDITIONNEL	0	0	0	0	0	0	0
	1890	1891	1892	1893	1894	1895	1896
PRESENT	100	100	0	100	66,6	100	84,61
FUTUR	0	0	0	0	0	0	15,39
PASSE-COMPOSE	0	0	0	0	0	0	0
IMPARFAIT	0	0	0	0	33,4	0	0
CONDITIONNEL	0	0	0	0	0	0	0
	1897	1898	1899	1900	1901	1902	1903
PRESENT	90,47	80	66,6	76,9	84,2	75	75
FUTUR	9,53	14,3	20	5,8	10,5	16,6	12,5
PASSE-COMPOSE	0	2,85	13,4	5,8	0	3,3	0
IMPARFAIT	0	2,85	0	7,7	5,3	3,3	10,0
CONDITIONNEL	0	0	0	3,8	0	1,8	2,5

Tableau III-4 : Emploi des temps de l'indicatif et le conditionnel des titres-commentaires (en pourcentage).

Les phrases verbales de cette macro-catégorie sont massivement associées au présent, et ce, tout spécialement sur la tranche diachronique 1885-1893 et en 1895, où le présent est d'ailleurs le seul temps employé. En 1883 et 1884, ainsi que de 1896 à 1903, si les temps et les modes sont plus diversifiés, il reste que le présent demeure le temps le plus fréquemment utilisé.

Les titres (51) et (52) sont représentatifs de l'emploi de ce temps dans les titres-commentaires. Le présent dans l'énoncé (51) sert le dessein ironique de l'énoncé, en feignant un rapport de conséquence entre la décision de constituer une commission chargée d'évaluer les différentes méthodes d'enseignement de la prononciation du grec, rapportée dans la dépêche en tant qu'événement passé, et l'avenir de la nation française. Renforcée par la valeur temporelle ici

étendue du présent, la tonalité hyperbolique de l'énoncé (51) vient dès lors créer un effet de dissonance pinçant, au travers duquel est dénigrée l'importance de l'information présentée dans l'article.

Ch.III (51), « La France est sauvée », (1896-01-31)

Le ministre de l'instruction publique vient de charger une commission d'étudier les moyens de substituer dans l'enseignement la prononciation moderne du grec à la prononciation dite Erasmiennne, en usage aujourd'hui dans les établissements scolaires.

Par un jeu avec l'interdiscours, le titre commentaire à dimension ludique (52) vise à faire résonner, par exploitation d'une paronymie entre *oisifs* et *oiseaux*, une vérité générale caractérisée par une valeur permanente du présent, et la singularité d'un procès historique : en l'occurrence, la libération d'un écrivain anarchiste et d'un journaliste socialiste. Au détour de ce jeu de mot axiologisé, *Le Petit Comtois* tourne en dérision les anarchistes et les socialistes :

Ch.III (52), « Les oisifs captifs prennent leur vol », (1895-02-03)

L'écrivain anarchiste, M. Jean Grave, et le journaliste socialiste révolutionnaire, M. Breton, qui bénéficient de l'amnistie, ont été mis en liberté aujourd'hui de la maison de Clairvaux.

Les titres-commentaires se caractérisent également par des constructions verbales au futur. Ce temps appuie la visée de commentaire de cette catégorie de titre en mettant l'accent sur les conséquences de l'événement rapporté par l'article. Ainsi, les titres (53) et (54), qui témoignent d'une dimension subjective explicite, expriment la réaction de la rédaction du quotidien face à l'annonce successive de la création d'une « alliance républicaine progressiste » et de la détérioration de l'état de santé de C. Rhodes. Par ces titres, sont respectivement signalés une méfiance à l'égard de la création de l'alliance républicaine et la violente indifférence du *Petit Comtois* à la mort prochaine du Premier ministre de la colonie du Cap.

Ch.III (53), « On les verra à l'œuvre », (1902-03-14)

Un nouveau groupement vient de se former au Sénat, quatre-vingts sénateurs environ, se sont unis et ont constitués un parti auquel ils ont donné le titre d'alliance Républicaine progressiste. M. Pevet a été nommé président à l'unanimité.

Ch.III (54), « Nous n'en porterons pas le deuil », (1902-03-05)

Le Gap, 24 mars. Après une nuit agitée, M. Cecil Rhodes est plus faible ; il ne s'est réellement produit aucune amélioration dans l'état du cœur qui reste grave.

Proportionnellement à l'emploi du présent et du futur, l'utilisation d'autres temps et d'autres modes reste anecdotique dans la catégorie des titres-commentaires. On notera ainsi quelques rares cas d'emploi du mode conditionnel. Loin de correspondre à l'utilisation d'un conditionnel dit journalistique, l'occurrence du conditionnel dans les titres-commentaires renvoie, non pas à une prise de distance face à la véracité des dires rapportés en article, mais à un futur hypothétique. Dans le titre (55), le *Petit Comtois* adhère à l'hypothèse selon laquelle la démission du chancelier serait prochaine.

Ch.III (55), « Ils pourraient avoir raison », (1902-03-14)

Une dépêche de Munich annonce que le comte de Bulow, chancelier de l'Empire, arrivera vers la fin de la semaine prochaine dans cette ville, se rendant en Italie.

Ce voyage est motivé par son mauvais état de santé.

Toutefois, plusieurs journaux de Top-position, veulent y voir le prélude de la démission du chancelier.

De même, l'utilisation de l'imparfait dans l'ensemble de notre corpus est tout à fait occasionnelle, comme le démontre la consultation du tableau III-4. L'emploi de l'imparfait vise alors à signaler l'antériorité d'un fait par rapport à celui présenté dans l'article. En ancrant l'information dans l'imparfait, le titre dénie à celle-ci sa valeur principale, c'est-à-dire l'actualité, comme dans l'exemple ci-dessous :

Ch.III (56), « On le savait déjà », (1903-02-21)

Le Caire, 20 mars.

Sur l'avis des médecins, le prince impérial d'Allemagne et son frère ont décidé de renoncer à leur voyage en Palestine et à Constantinople.

Aussitôt que la guérison du prince impérial sera complète, les deux princes se rendront directement à Rome pour se joindre à l'empereur Guillaume II.

Un lieu de concentrations des occurrences des modalisations exclamatives, interrogatives et impératives

L'assertivité, tout comme sur l'ensemble des autres macro-catégories du corpus, est la forme la plus courante dans la construction syntaxique des titres-

commentaires. De nombreuses traces de la subjectivité du locuteur viennent alors supporter la visée de commentaire comme l'illustrent les exemples suivants :

Ch.III (57), « Il n'ose pas avouer ses pertes », (1900-01-13)

Le *War Office* publie la dépêche suivante de lord Roberts datée d'un village près Bloemfontein :

« Notre marche en avant, hier dimanche, n'a pas rencontré d'opposition.

Les officiers que j'ai laissés en arrière pour relever la liste de nos morts et de nos blessés n'ont pas encore rejoint la colonne.

Nous avons à ajouter à nos pertes, en officiers, deux tués et trois blessés. »

Ch.III (58), « Il n'est pas assez réactionnaire pour être élu », (1903-03-26)

Le bruit s'accrédite que M. Waldeck-Rousseau, sur les instances de ses admirateurs et de ses amis, posera sa candidature à l'Académie française pour le siège laissé vacant par la mort de M. Gaston Paris.

Si l'assertivité demeure la modalité privilégiée dans la production des titres-commentaires, il est à remarquer que les modalités impératives, interrogatives et exclamatives sont également employées dans des mesures significatives dans cette macro-catégorie de titres. Les titres-commentaires concentrent d'ailleurs presque toutes les modalités portant les traces d'une subjectivité explicite du locuteur.

	1883	1884	1885	1886	1887	1888	1889
MODALITE INTERROGATIVE	8,9	11	0	0	0	0	0
MODALITE IMPERATIVE	8,9	0	0	0	0	0	0
MODALITE EXCLAMATIVE	3,8	6,6	8	2,5	0	0	0
MODALITE ASSERTIVE	78,4	82,4	82	97,5	100	100	100
	1890	1891	1892	1893	1894	1895	1896
MODALITE INTERROGATIVE	0	7,4	0	0	1,6	0	4
MODALITE IMPERATIVE	2,7	0	0	0	0	0	2
MODALITE EXCLAMATIVE	0	0	0	0	3,2	1,5	3
MODALITE ASSERTIVE	97,3	92,6	100	100	95,2	98,5	91
	1897	1898	1899	1900	1901	1902	1903
MODALITE INTERROGATIVE	1,6	3,6	6,3	5,4	5,8	5	4,9
MODALITE IMPERATIVE	2,4	0	1	5,4	0,6	7	1,6
MODALITE EXCLAMATIVE	4,8	7,2	4,2	5,6	4,6	12,1	2,5
MODALITE ASSERTIVE	91,2	89,2	88,5	83,6	89	75,9	91

Tableau III-5 : Les modalités des titres-commentaires (en pourcentage).

En diachronie, notons que les titres-commentaires se caractérisent par une utilisation significative des modalités autres qu'assertive pendant les toutes premières années de parution. L'utilisation des modalités impératives, interrogatives et exclamatives oscille au contraire entre une tendance nulle ou anecdotique de 1886 à 1893 : la modalité interrogative connaît un point culminant à 7,4% en 1891, mais il convient de nuancer ce pourcentage en le rapportant au faible nombre de titres-commentaires répertoriés à cette date. Si l'assertivité reste encore très largement le mode privilégié dans les titres-commentaires de 1894 à 1895, on observe une tendance à la diversification des modalités de 1897 à 1903, où les titres exclamatifs, impératifs et interrogatifs sont employés dans des mesures significatives.

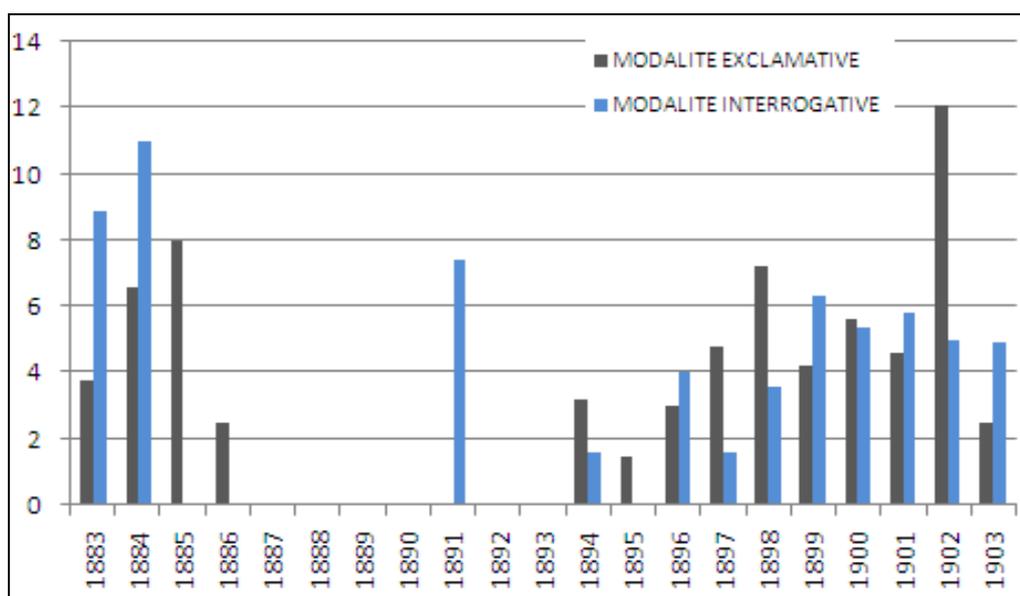


Figure III-33 : Les modalités exclamative et interrogative dans les titres-commentaire (en pourcentage).

Par ordre de fréquence, la seconde modalité la plus fortement représentée dans cette catégorie est l'exclamatif. La coloration subjective et l'affectivité que marque ce type d'énoncés la prédestinent tout naturellement à s'inscrire dans la classe des titres-commentaires. S'ils ne sont caractérisés par aucune valeur informative, les titres (59), (60), (61) commentent l'information délivrée par les articles qu'ils « coiffent », dont ils orientent la réception. Le mode exclamatif vient renforcer l'illusion d'oralité de ces énoncés au style familier. Renvoyant à une expression familière, (59) constitue ainsi une antiphrase par laquelle est exprimé le manque de moralité des actes rapportés. Le titre (60) témoigne de la même dimension subjective. Par cet énoncé exclamatif, c'est l'opinion de la rédaction du quotidien sur le projet de vote de J. Fabre qui nous est présenté dans le corps de l'article.

Ch.III (59), « C'est du propre ! », (1903-03-31)

Lisbonne, 27 février. — Sur l'ordre des autorités militaires, on vient d'arrêter le lieutenant Teixeira Malheiro. Il est accusé d'avoir été une des principales causes de la révolte d'Angelo, ayant fait subir des atrocités aux indigènes. Mais le lieutenant ne se laisse pas démonter. Il menace de faire des révélations et de

prouver, pièces en main, que de très hauts personnages de l'armée font la traite des noirs dans l'Afrique portugaise.

Ch.III (60), « Le sénat ne le votera pas ! », (1897-02-13)

Voici l'ordre du jour que déposera M. Joseph Fabre comme conclusion de sa prochaine interpellation : « Le Sénat invite le gouvernement à n'admettre aucune intervention dans la politique intérieure de la France et à appliquer avec vigilance et fermeté toutes les lois qui protègent la société civile contre les empiétements des autorités ecclésiastiques. »

Ch.III (61), « Pardi ! », (1901-03-07)

Pékin, 5 mars.

Une information de source officielle dément que l'empereur ait l'intention de quitter sa résidence actuelle.

Au contraire, si les puissances persistent dans leur projet relatif aux territoires des légations, il est probable que la cour ne retournera jamais à Pékin.

Un nombre conséquent de titres rencontrés dans notre corpus s'appuient également sur des tournures interrogatives. Parce qu'ils participent d'un simulacre d'oralité, soit qu'ils expriment une demande d'information adressée au lecteur soit qu'ils relèvent de la pure rhétorique, ces titres assurent pleinement une visée de captation. Dans (62), le pronom démonstratif « cela » renvoie à un référent dont l'identification est problématique, ce qui nous semble pouvoir être envisagé comme un ressort complémentaire de la captation. Aucun élément co-textuel ne permet de l'actualiser : dépourvu d'antécédents, le pronom fait appel aux connaissances culturelles et encyclopédiques du lecteur. En nous basant sur l'histoire du Venezuela, on peut voir dans ce titre une allusion aux événements de 1902, ayant abouti au blocus des ports vénézuéliens par la Grande-Bretagne et l'Allemagne en raison de dettes non remboursées par le Venezuela. L'actualité est ainsi commentée en étant mise en rapport avec l'histoire. Faussement naïves, les formes interrogatives dans les titres (63) et (64) ne font pas tant appel à l'opinion du lecteur qu'elles ne véhiculent l'opinion de la rédaction sur les événements rapportés. La forme interrogative ne doit pas faire illusion ; elle est le support d'une affirmation politique, devant laquelle est placé le lecteur avant même la réception de

l'événement lui-même. Dans ces énoncés, il s'agit bien de réagir à l'information, tout en interpellant le lecteur en imitant les conditions d'un dialogue :

Ch.III (62), « Est-ce que cela va recommencer ? », (1903-03-02)

New-York, 1^{er} mars.

Un télégramme de la Trinidad annonce que les révolutionnaires vénézuéliens ont pris Carupano, après six heures de combat ; ils pillèrent la ville et blessèrent plusieurs habitants.

Les résidents allemands ont demandé l'envoi d'un navire de guerre pour les protéger.

Ch.III (63), « Bien, mais où prendra-t-il les soldats ? », (1902-03-24)

Le War-Office a décidé de créer dix-huit nouvelles batteries d'artillerie de campagne.

Les pièces seront à tir rapide du dernier modèle.

Ch.III (64), « Pourquoi ne fait-il pas appeler Borne ? », (1900-01-04)

Au ministère de la guerre, on communique la note suivante :

« L'état de santé du général Galliffet continue à s'améliorer, mais avec une grande lenteur. Le ministre de la guerre est toujours alité et condamné à un repos absolu. »

Bien qu'elles ne comptent pas de verbes, les formes interrogatives nominales assurent les mêmes fonctions. Ainsi, les titres (65) et (66) ne s'enquière pas uniquement d'informations absentes du corps de l'article. (65) jette un regard critique sur les déclarations victorieuses des anglais en le faisant contraster avec le silence relatif aux pertes anglaises, tandis que (66) ironise sur la durée de la guerre des Boers.

Ch.III (65), « Et les pertes anglaises ? », (1902-03-02)

Lord Kitchener télégraphie de Harrismith, 1^{er} mars :

« Les pertes des Boers, dans les récentes opérations, s'élèvent à 800. »

Ch.III (66), « Au mois de juin de quelle année ? », (1900-01-31)

Londres. — Lord Roberts écrit à un de ses amis que la guerre ne pourra, vraisemblablement, se terminer avant le mois de juin.

L'impératif, tout comme le mode interrogatif et le mode exclamatif, assure une fonction incitative. Le titre (67) sollicite ainsi directement le lecteur, en lui recommandant de prendre ses distances avec l'information délivrée dans la brève correspondante. Dans (68), on note l'adresse feinte aux malgaches, par laquelle est sous-entendue une évolution du conflit au bénéfice des français,

l'énoncé ayant également pour corollaire une tonalité humoristique voire ironique.

Ch.III (67), « Méfiez-vous toujours », (1900-01-04)

Lourenço-Marquez, 1er mars.

21 hommes et 1 officier ont été faits prisonniers par les Boers vendredi, à Colenso. Le général boer Botha, a télégraphié à Pretoria que les mouvements des Anglais depuis dimanche étaient inexplicables.

Ch.III (68), « Tenez-vous bien, Malgaches », (1883-08-29)

M. le contre-amiral Galibert quitte Paris pour prendre le commandement de l'escadre de Madagascar, en remplacement de l'amiral Pierre, qui revient très souffrant.

Plus que les marques d'une subjectivité explicite, les modalités interrogatives, impératives et exclamatives, qui se posent comme une des caractéristiques formelles des titres-commentaires, assurent une fonction incitative. Feignant de prendre le lecteur à partie, ces modalités donnent l'illusion d'un dialogue dynamique avec le lecteur. Rompant avec le factuel des brèves, elles permettent au quotidien d'influencer la réception de l'information, en la colorant affectivement et politiquement.

Franc-parler pour parler vrai

Ce qu'il nous semble intéressant de noter au sujet des titres-commentaires, c'est qu'au-delà de la modalité adoptée, un ensemble de caractéristiques formelles sont dévolues à imiter l'oralité. Bien qu'ils relèvent de la modalité assertive, les titres suivants simulent les conditions d'un dialogue virtuel avec le lecteur, en s'adressant à lui comme s'il s'agissait d'un échange discursif ayant cours dans une sphère privée. Suppression des marques de la négation, à l'instar de ce qu'on observe dans les titres (69), (70) et (71), s'accompagnant d'une omission du sujet dans (69) et (70), contraction et utilisation du pronom « ça », comme dans (72) : tels sont les traits linguistiques que les titres du *Petit Comtois* empruntent à l'oralité pour capter l'attention du lecteur tout en alimentant la connivence. Entretenir l'accointance avec son

lectorat, tel nous semble bien l'objectif de la tournure syntaxique relevée dans (73) :

Ch.III (69), « C'est pas fini », (1903-03-02)

M. Cattauï a formé aujourd'hui appel du jugement de la neuvième chambre de la police correctionnelle de Paris, qui a acquitté la famille Humbert. Conformément à la jurisprudence établie en pareil cas, le greffier, au nom du parquet, a fait aussitôt appel à minima.

Ch.III (70), « Manquait plus que ça », (1898-02-06)

Les Dardannelles. — Une forte secousse de tremblement de terre a été ressentie aujourd'hui.

Ch.III (71), « Voyons pas d'inconvénient », (1901-03-08)

Le différend qui a mis aux prises M. Paul Déroulède et André Buffet se rouvre sur de nouvelles bases à la suite du télégramme adressé à M. André Buffet et conçu en termes violents.

M. André Buffet a prié deux de ses amis de Paris de s'aboucher avec les nouveaux témoins de M. Déroulède, qui sont MM. Maurice Barrés et le docteur Devilliers.

Ch.III (72), « V'là qu'ça recommence », (1900-01-04)

Kimberley, 1^{er} mars. Les Boers ont marché sur Barkly-West, refoulant les petits postes anglais.

Ch.III (73), « C'est le tour à Chamberlain », (1902-03-20)

On téléphone de Londres qu'une dépêche arrivée cette après-midi annonce la mort de Cécil Rhodes.

Traces du franc-parler qu'affecte le quotidien pour se poser comme le journal du petit peuple, les titres-commentaires sont le site de prédilection d'un style familial, comme l'illustrent les exemples (74), (75), (76) et (77). Si les expressions familières sont légion dans les titres commentaires, les surnoms injurieux, pourvus d'une dimension argumentative certaine, apparaissent également dans cette macro-catégorie de titres. La désignation des bonapartistes par le surnom axiologisé « badingueusards » employée dans le titre (77) reflète ainsi de façon explicite l'aversion républicaine que porte le quotidien à ces derniers, dans le même temps qu'il agit comme ressort de l'accointance avec le lectorat républicain.

Ch.III (74), « Ils n'y vont pas de main morte », (1902-03-14)

D'après une dépêche de Constantinople, Delatrlo, le chef brigand albanais, a enlevé dernièrement huit officiers turcs du vilayet de Janina.

Ces officiers avaient encouragé leurs troupes à commettre toutes sortes d'atrocités. Delatrio les fit pendre. Leurs corps furent ensuite coupés en morceaux et les morceaux envoyés à la garnison turque la plus proche accompagnés d'une lettre déclarant que c'était ainsi que les Albanais entendaient traiter les tyrans.

Ch.III (75), « Diable, ça se corse », (1901-02-31)

Paris, 30 mars 1891

Saint-Pétersbourg, 30 mars.

Une soixantaine d'étudiants avaient entrepris de se rendre en députation pour exposer au tsar les doléances et les réclamations de la jeunesse scolaire; mais ils en ont été empêchés par la police qui a arrêté plusieurs d'entre eux.

Ses deux collègues, pour avoir protesté, ont été conduits aux travaux forcés en Sibérie.

Ch.III (76), « C'est leur droit d'être grincheux », (1883-08-31)

On signale plusieurs manifestations réactionnaires : une à la cour d'Angers, qui a laissé un avocat qualifier la réforme de la magistrature de loi de proscription, l'autre à Périgueux, où le président du tribunal, en prenant sa retraite, a prononcé un discours inconvenant dans un banquet où les avocats et les avoués républicains s'étaient abstenus de paraître.

Ch.III (77), « Les badingueusards ont oublié Sedan », (1899-02-27)

Sur la place Vendôme devait avoir lieu une manifestation organisée par les comités impérialistes ; des mesures de police avaient été prises depuis ce matin.

On a arrêté quelques manifestants qui, malgré la défense faite, jetaient des bouquets de violettes près de la colonne.

Les titres informationnels

Des titres en moyenne plus longs que ceux relevant des autres macro-catégories du corpus

La première caractéristique formelle des titres informationnels que nous voudrions souligner est que ces titres sont en moyenne plus longs que ceux qui relèvent des autres macro-catégories répertoriées, ce que montre le graphique III-34.

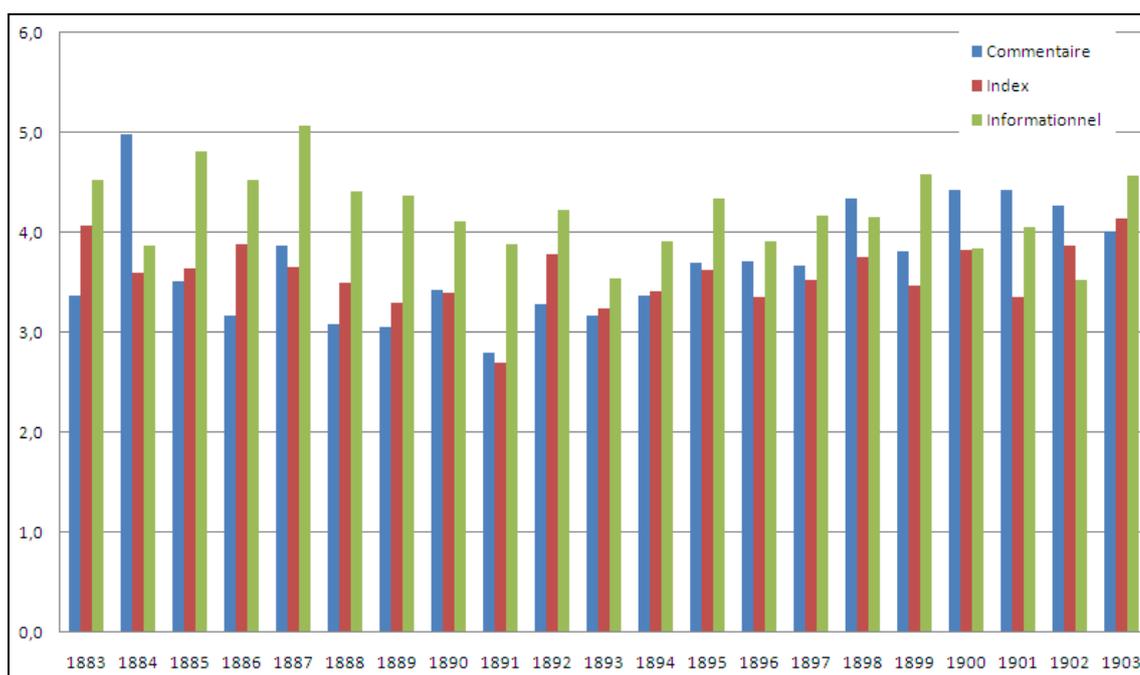


Figure III-34 : La taille moyenne des titres (en nombre de mots).

Que traduit la longueur supérieure de ces titres en comparaison avec la longueur moyenne observée dans le corpus ? Il paraît légitime de formuler l'hypothèse selon laquelle elle est l'indice de structures syntaxiques complexes, ce que nous allons examiner à présent.

Des titres informationnels généralement construits à partir d'un syntagme nominal suivi d'un syntagme prépositionnel

Pour observer les types de constructions caractéristiques des titres-informationnels, nous avons répertorié dans le tableau ci-dessous les quatre formes de constructions les plus fréquemment rencontrées dans cette macro-catégorie : au type (A) correspondent ainsi les titres basés sur la séquence *Syntagme nominal + syntagme prépositionnel* ; au type (B) les constructions passives tronquées, que nous avons distinguées des phrases verbales complètes désignées pour leur part en tant que type (C). Enfin, le type (D) regroupe les titres où une partie de la phrase est mise en ellipse.

	1883	1884	1885	1886	1887	1888	1889
TYPE A	26=59,2%	17=70,9%	17=89,5%	28=100%	17=78,9%	19=95%	21=70%
TYPE B	1=2,2%	0	2=10,5%	0	6=26,1%	1=5%	9=30%
TYPE C	15=34,1%	6=25%	0	0	0	0	0
TYPE D	2=4,5%	1=4,1%	0	0	0	0	0
	1890	1891	1892	1893	1894	1895	1896
TYPE A	31=39,9%	15=83,4%	17=85%	19=86,4%	23=76,7%	34=91,9%	20=95,2%
TYPE B	2=6,1%	2=11,1%	1=5%	3=13,6%	7=23,3%	0	0
TYPE C	0	0	0	0	0	0	1=4,8%
TYPE D	0	1=5,5%	2=10%	0	0	3=8,1%	0
	1897	1898	1899	1900	1901	1902	1903
TYPE A	23=69,7%	32=71,1%	24=66,7%	24=70,6%	22=66,6%	33=67,2%	33=71,7%
TYPE B	3=9,1%	4=8,9%	4=11,1%	4=11,8%	3=11,1%	8=16,4%	5=10,9%
TYPE C	5=15,1%	9=20%	8=22,2%	6=17,6%	8=22,3%	8=16,4%	7=15,2%
TYPE D	2=6,1%	0	0	0	0	0	1=2,2%

Tableau III-6 : Les différents types de construction syntaxiques des titres-informationnels en diachronie.

Le type (A) : Syntagme nominal + syntagme prépositionnel

Comme on peut l'observer très nettement sur le tableau ci-dessus, la construction la plus massivement utilisée dans les titres informationnels se révèle de type (A), correspondant au patron *Syntagme nominal + syntagme prépositionnel*.

La prédominance de ce type de construction dans les titres informationnels nous semble se justifier par sa capacité à condenser l'information. Dans les exemples (78), (79) et (80), l'énoncé est focalisé sur le syntagme nominal que spécifie le syntagme prépositionnel.

Ch.III (78), « Tremblements de terre en Amérique », (1889-01-05)

Des tremblements de terre ont eu lieu à San-José et à Alajuela. Huit personnes ont été tuées et un grand nombre blessées. Des édifices publics sont endommagés.

Ch.III (79), « Accident à l'Elysée », (1889-01-20)

Le jardinier de l'Elysée, M. Lebon, a été tué ce matin par la chute d'un arbre qu'il abattait dans le jardin du président.

Mme Carnot a fait immédiatement remettre 300 fr. pour venir en aide à la famille de ce malheureux ouvrier.

Ch.III (80), « Tentative d'insurrection en Serbie », (1902-03-06)

Belgrade, 6 février.

Alavantitch, parent du prétendant Karageocevitch, accompagné de plusieurs hommes, est arrivé ce matin par le bateau à Schabatz. Il portait l'uniforme de général Serbe et il a exhorté les gardes-frontières à le suivre comme général. Il s'est rendu ensuite à la caserne de gendarmerie, où il a harangué les gendarmes. [...]

Trois officiers supérieurs ont été arrêtés. On aurait saisi chez eux des papiers très compromettants qui prouveraient l'existence d'un complot.

L'absence de déterminants ajoute à l'effet de condensation de l'information liée à la nominalisation. Cette configuration n'est cependant pas la règle constante, puisqu'un nombre conséquent de titres informationnels construits à partir d'une nominalisation sont précédés d'un article indéfini, à l'instar des exemples suivants :

Ch.III (81), « Une accalmie en Macédoine », (1903-03-14)

Sofia, 12 mars. — Le retour du froid et la célébration des fêtes du Baïram ont produit une interruption des hostilités en Macédoine.

Ch.III (82), « Un éboulement dans un tunnel », (1889-01-24)

Londres, 23 janvier. Un terrible accident a eu lieu la nuit dernière à Aberggwynfi, dans le tunnel qu'on construit sur la ligne de Khonda et Swansea. Il a coûté la vie, craint-on, à sept personnes.

Une bande d'ouvriers travaillaient à la construction du tunnel, lorsqu'une grande quantité de terre et de pierres roula sur ces ouvriers, dont un certain nombre furent ensevelis sous l'avalanche.[...]

Nombreux sont également les syntagmes nominaux constitués autour d'un nom propre ou se référant à la fonction occupée par l'actant « agent » de l'énoncé, comme dans (83) et (84).

Ch.III (83), « Le ministre de la marine à Saint-Malo », (1903-03-18)

Saint-Malo, 17 mars.

Le ministre de la marine s'est rendu à Saint-Malo pour assister au départ du vapeur Burgundia, transportant à Miquelon les vivres nécessaires pour les pêcheurs de Terre-Neuve.

M. Pelletan est rentré ce soir à Paris.

Ch.III (84), « M. Rochefort à Paris », (1895-02-04)

M. Henri Rochefort est arrivé à Paris à quatre heures quarante, par le train de Calais. Dès deux heures, une foule énorme se pressait autour de la gare du Nord ; la salle des Pas-Perdus était bondée. On n'accédait sur les quais qu'avec un billet de parcours. Cinq cents gardiens de la paix maintiennent la circulation sur le

boulevard Denain et dans la rue de Dunkerque. Beaucoup de députés socialistes, parmi lesquels MM. Gérault-Richard, Rouanet, Jaurès, etc. La rédaction et l'administration de l'Intransigeant et de la Cocarde attendent l'arrivée du train. Lorsque Rochefort arrive, les acclamations retentissent. On crie : « Vive Rochefort ! » Les délégués des comités socialistes et révisionnistes sont massés devant l'église Saint-Vincent-de-Paul. [...]

De façon tout à fait représentative de ce que l'on peut observer sur l'ensemble du corpus des titres informationnels, les syntagmes prépositionnels rencontrés dans les exemples précédemment cités ont une valeur de localisation géographique. *A contrario*, la valeur temporelle est systématiquement effacée des titres dans ce type de construction, alors que la nominalisation elle-même situe l'information dans un hors-temps : seul le corps de l'article précise alors les circonstances temporelles de l'événement présenté en titre.

Le type (B) : les constructions passives tronquées

Le deuxième type syntaxique que nous avons répertorié dans notre corpus de titres informationnels correspond aux constructions passives tronquées²² (type B). S'il se révèle malaisé de donner une signification à l'évolution diachronique d'emploi des constructions passives tronquées, comme le suggère le graphique suivant, on se risquera cependant à noter la stabilisation de leur fréquence à partir de 1897.

²² Au sujet des transformations liées au passif et aux règles d'effacement de la copule, on se référera à Dubois (1969).

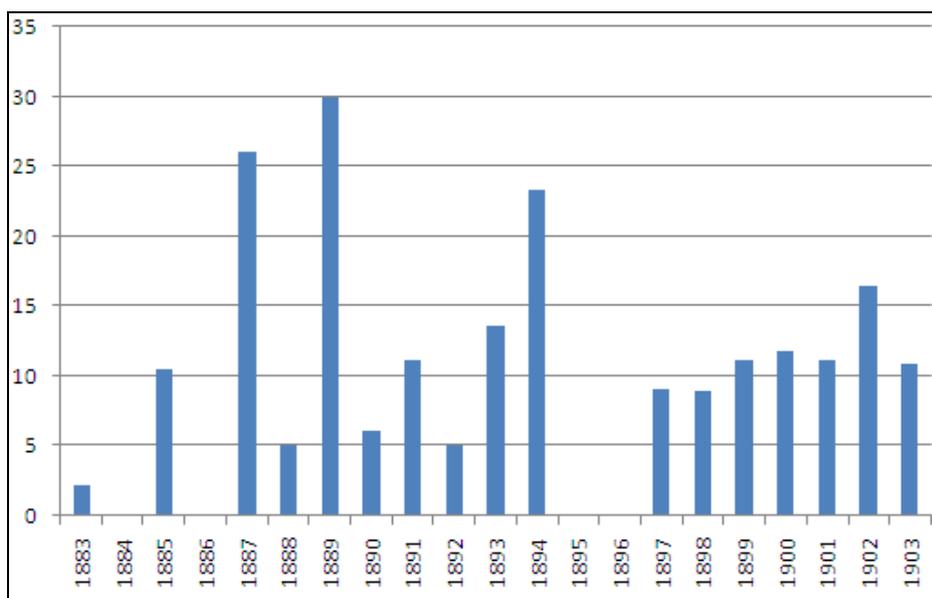


Figure III-35 : Les constructions de type B dans les titres informationnels (en pourcentage).

Dans ce type de construction, l'effacement de l'auxiliaire de passivation s'accompagne de celui du complément d'agent. Les titres de type (B) se caractérisent par un effacement des marques du discours : aucun des exemples répertoriés ne comporte ainsi un complément circonstanciel spatio-temporel. Se résumant à l'actualisation d'un thème par un rhème, ce type de construction se caractérise de fait par une forte condensation du contenu informationnel. Dans les exemples (85), (86), (87), (88), on relèvera la tendance à la typification du syntagme nominal, pourvu ou non d'un déterminant, relatif aux catégories socio-professionnelles ou, plus largement à l'activité du sujet. Plus qu'un personnage précis, dans les exemples, ce sont des parangons, des représentations identitaires et culturelles contemporaines du moment de l'énonciation, qui sont les thèmes des énoncés.

Ch.III (85), « Rentière étranglée », (1903-03-30)

Mme veuve Lachèze, rentière, âgée de cinquante-six ans, a été trouvée étranglée dans son hôtel, à Neuilly; la mort remontait à dix-huit heures environ.

Au premier étage se trouvait un coffre-fort que le ou les assassins avaient, sans y réussir, tenté d'ouvrir.

Ch.III (86), « Espion fusillé », (1902-03-13)

Cracovie, 11 mars.

On apprend de Varsovie que le 7 mars, à huit heures, on a fusillé le colonel Grimm, en garnison dans la circonscription militaire de Varsovie. Ce colonel avait avoué avoir livré, depuis dix ans, des secrets militaires à l'Allemagne.

Ch.III (87), « Un révérend fusillé », (1894-02-09)

Le révérend père Castaneda, accusé de complot contre le gouvernement, a été fusillé près de Talaueingo.

Son exécution a causé un vif sentiment d'affliction.

Ch.III (88), « Trois cents hommes massacrés », (1902-03-17)

Des passagers arrivés de Bocas-del- Toro racontent que les trois cents hommes de la garnison de David, ville colombienne du Costa-Rica, ont été massacrés par les insurgés. Cinquante partisans du gouvernement ont été tués en défendant leurs retranchements, et le reste cerné, puis massacré à coup de machette sans aucun quartier.

Le type (C) : les phrases verbales complètes

Hiérarchiquement, le type C, par lequel nous avons répertorié les occurrences de phrases verbales complètes, est la seconde forme syntaxique la plus fréquemment rencontrée dans les titres informationnels. Comme on peut l'observer sur le graphique ci-dessous, les phrases verbales complètes représentent près de 35% des titres-informationnels en 1883, puis 25% en 1884. Cependant, de 1885 à 1895, l'emploi de cette construction est nul : toute interprétation de ce résultat nous semble devoir être prudente, puisque sur cette même période, les effectifs des titres informationnels sont légèrement inférieurs à ceux de 1883 ou de la tranche 1897-1903. Il reste que ce profil déficitaire est cohérent avec la très faible utilisation des phrases verbales observée sur l'ensemble du sous-corpus sur cette période (cf. figure III-28). Nous nous limiterons à signaler que les phrases verbales complètes sont utilisées dès 1883 dans notre sous-corpus et qu'elles tendent à se développer sur la période 1896-1903.

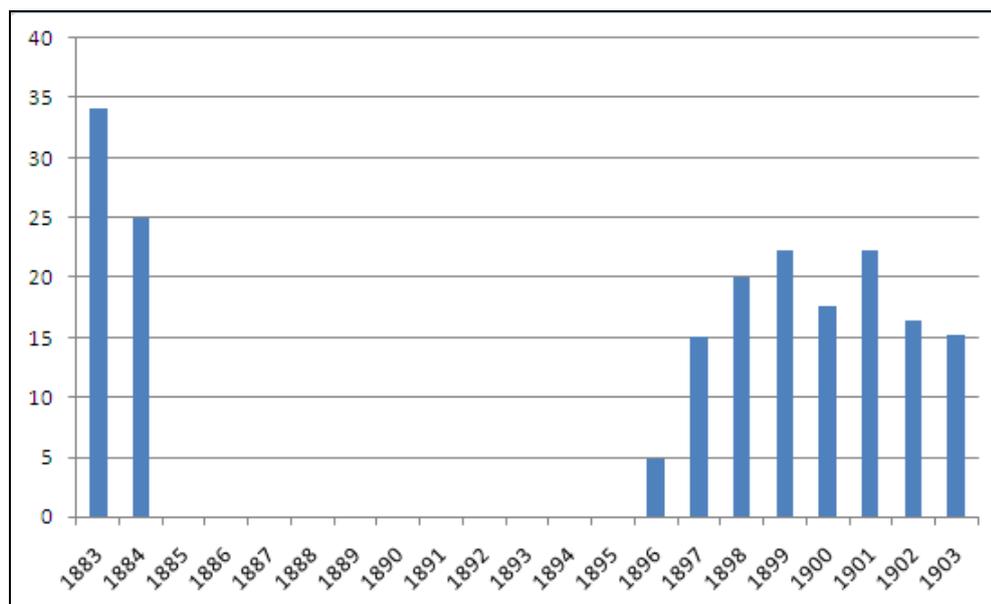


Figure III-36 : Les phrases verbales complètes dans les titres informationnels en diachronie (en pourcentage).

Examinons les temps et les modes employés dans ces énoncés. Dans la mesure où les titres reposant sur une construction nominale estompent les marques de la temporalité, il nous paraît intéressant d'observer quels rapports à l'actualité sont instaurés par les phrases verbales.

	1883	1884	1885	1886	1887	1888	1889
PRESENT	86,6	0	0	0	0	0	0
PASSE-COMPOSE	0	0	0	0	0	0	0
FUTUR	13,4	0	0	0	0	0	0
IMPARFAIT	0	0	0	0	0	0	0
CONDITIONNEL	0	0	0	0	0	0	0
	1890	1891	1892	1893	1894	1895	1896
PRESENT	0	0	0	0	0	0	0
PASSE-COMPOSE	0	0	0	0	0	0	0
FUTUR	0	0	0	0	0	0	100
IMPARFAIT	0	0	0	0	0	0	0
CONDITIONNEL	0	0	0	0	0	0	0
	1897	1898	1899	1900	1901	1902	1903
PRESENT	57,1	90,9	62,5	83,4	100	80	57,1
PASSE-COMPOSE	0	0	0	16,6	0	0	0
FUTUR	28,5	9,1	37,5	0	0	20	28,5
IMPARFAIT	0	0	0	0	0	0	0
CONDITIONNEL	14,4	0	0	0	0	0	14,4

Tableau III-7 : Temps des phrases verbales complètes relevant des titres informationnels (en pourcentage).

Il ressort clairement de la consultation du tableau III-7 que le temps privilégié dans le type de construction considéré est le présent. On notera, sur ce point, le contraste de traitement temporel de l'événement entre le titre et l'article. Alors que l'article situe explicitement l'événement dans le passé, comme en témoignent l'utilisation de déictiques et des temps du passé simple et du passé composé, le titre lui, donne à lire l'événement au présent. En d'autres termes, le point de référence des événements ainsi mis en scène se révèle contemporain de l'acte d'énonciation et de réception de l'énoncé. Le présent, en tant que « temps de l'actualité » donne ainsi l'illusion d'une actualisation spatiale et temporelle plus forte de l'information²³, qui est à relier à la recherche constante du discours du quotidien à coïncider avec le temps du monde. Parce qu'il vise à

²³ Ce que P. Charaudeau décrit notamment sous le terme d'*effet d'actualisation* (Charaudeau, 1992 : 464).

donner l'impression au lecteur qu'il vit l'événement, on pourrait considérer cette utilisation du présent comme relevant de l'hypotypose.

Ch.III (89), « Les insurgés franchissent la frontière », (1883-08-07)

On dit que 21 officiers espagnols et 22 soldats passèrent la frontière et furent conduits à Perpignan.

Parmi ces officiers se trouvaient un colonel et huit commandants.

Ch.III (90), « La terre tremble de nouveau », (1903-03-30)

Brest, 29 mars. Une nouvelle secousse de tremblement de terre a été ressentie hier à Brest ; la première secousse s'était fait sentir le 7 février.

Il convient de mettre en parallèle cette observation avec la très faible utilisation du passé composé, qui exprime l'aspect accompli d'un procès, tout en pouvant marquer l'antériorité par rapport à l'acte d'énonciation.

Ch.III (91), « Les Boers sont rentrés chez eux », (1900-03-07)

Londres. — Le War Office communique un télégramme de Lord Roberts daté d'Osfonteia, 5 mars : « Le général Buller annonce que le Natal est virtuellement débarrassé des Boers. »

Lorsque les titres informationnels construits à partir d'une phrase verbale complète ne sont pas conjugués au présent, ils le sont généralement au futur : 13,4% en 1883, 28,5% en 1897, 37, 5% en 1899. Il est intéressant de constater que l'utilisation du futur s'inscrit principalement dans des phrases négatives : ce temps est ainsi convoqué pour nier une information attendue ou prédite comme fortement probable. Dans les titres (92), (93) et (94), c'est respectivement l'internement de M. Zorilla qui est démenti, présupposant des déclarations contradictoires antécédentes, l'accompagnement du président annulé en raison de l'insistance de députés ministériels, l'emménagement du président Loubet retardé en raison d'une demande de Mme Faure.

Ch.III (92), « M. Zorilla ne sera pas interné », (1883-08-19)

On dément le bruit que l'ambassadeur d'Espagne à Paris aurait demandé au gouvernement français l'internement de M. Zorilla.

Ch.III (93), « M. Combes n'accompagnera pas le président », (1903-03-29)

D'après *la Liberté*, M. Combes aurait décidé, sur l'insistance d'une délégation des députés ministériels, de ne pas accompagner le président de la République dans son voyage en Algérie, sa présence devant être nécessaire à Paris pour

l'application de la loi sur les congrégations. Pour les mêmes raisons, M. Vallé et le général André resteraient à Paris.

Ch.III (94), « M. Loubet n'entrera à l'Elysée que la semaine prochaine », (1899-02-23)

Mme Félix Faure a demandé au président de la République si elle pouvait occuper jusqu'à samedi les appartements de l'Elysée.

M. Loubet a immédiatement acquiescé à ce désir.

Le président de la République ne s'installera donc, à l'Elysée que dans les premiers jours de la semaine prochaine.

Alors que les quotidiens qui nous sont contemporains usent du conditionnel dans de larges mesures, il est stimulant de noter que, sur l'ensemble du corpus des *Dépêches de nuit*, le conditionnel n'est que très rarement employé. Mis à part en 1897 et en 1903, on ne dénombre ainsi aucune occurrence de ce mode, alors que derrière les pourcentages observés en 1897 et 1903, ne se dissimule qu'une seule occurrence. Avant de nous interroger sur la signification de la rareté du conditionnel dans notre corpus de presse, déterminons tout d'abord la valeur de cette modalisation.

Dans les exemples (95) et (96), l'emploi du conditionnel correspond manifestement à une non-prise en charge de la nouvelle par le locuteur responsable de la conception du titre. Il s'agit bien ici de prendre des précautions avec l'information rapportée et de se décharger de la responsabilité de sa véracité sur le locuteur d'origine. En cela, ces deux exemples d'usage du conditionnel relèvent bien de ce que L. Abouda (1997 : 563) nomme « conditionnel de non-prise en charge » qu'il pose comme caractéristique du discours journalistique, ou de ce que K. Togeby (1982 : 388) nomme « conditionnel de la rumeur ». Modalisant l'acte d'énonciation du locuteur originel, ces titres suppriment la marque de ce dernier, que le corps de l'article seul explicitera : plus qu'une information, ces titres disent « un bruit » en tant que nouvelle non confirmée, dont le quotidien refuse de prendre la responsabilité énonciative. Notons les rapports entre l'usage du conditionnel du titre et la mise en scène de l'information dans la brève. Dans la brève (96), l'information selon laquelle le prétendant Bou-Hamars serait capturé est mise en scène *via* le discours rapporté : la dépêche prend acte de la réception d'un

courrier adressé au gouverneur de Tanger selon lequel Bou-Hamars est capturé. La nouvelle, dans ce cadre, ne fait l'objet d'aucune modalisation au conditionnel. Pour sa part, l'absence de réponse du sultan aux ambassadeurs est elle-même modalisée en second dans le corps de la brève (96) : elle est alors rapportée au passé composé de l'indicatif. Seule la seconde partie de l'énoncé, apportant un éclairage sur les conséquences de cette première information, se révèle modalisée au conditionnel. Or, quelle est l'information que condense le titre (96) ? Ce titre ne rapporte pas uniquement l'absence de réponse du sultan à l'invitation de proclamer l'indépendance de la Crète : il combine les deux informations en adjoignant au thème « le sultan » un rhème à dimension subjective, issu de l'analyse de la deuxième partie de la brève. Ce rhème « psychologise » le silence de l'actant en le rapportant au climat d'antagonisme signalé en fin de dépêche : par le fait, c'est également la dimension incertaine de ce segment qui est transférée dans le titre.

Ch.III (95), « Bou-Hamars serait capturé », (1903-03-08)

Tanger, 6 mars. — Le gouverneur de Tanger reçoit à l'instant un courrier de Fez lui annonçant la capture du prétendant.

Ch.III (96), « Le sultan ne voudrait rien entendre », (1897-02-27)

Selon le *Tagblatt*, le sultan n'a pas répondu à l'invitation des ambassadeurs de proclamer lui-même l'autonomie de la Crète ; les relations entre la Porte et les représentants des puissances seraient tendues.

Les trois types de construction précédemment évoqués suffisent à éclairer la quasi-totalité des structures des titres-informationnels de notre corpus. Signalons néanmoins que certains titres de cette macro-catégorie relèvent d'une structure syntaxique où une forme actualisante (*ici, il y a*) peut être considérée comme en ellipse, comme en (97) et (98). On remarquera que ces titres sont empreints des marques du discours et que (98) circonscrit l'événement dans une temporalité explicite :

Ch.III (97), « Toujours pas de nouvelles de la « Gascogne », (1895-02-10)

A sept heures, la compagnie transatlantique n'avait reçu aucune nouvelle de la Gascogne. La Teutonie, qui est arrivée dans le port de New-Yorck ce matin, à deux heures, avait un retard de quatre jours. Peut-être le retard de la Gascogne

est-il simplement causé par le mauvais état de la mer, sans même que le paquebot n'ait subi aucune avarie. La Teutonie n'en avait aucune.

Ch.III (98), « Pas de conseil des ministres demain », (1895-02-26)

Paris, 25 février 1895.

La réunion habituelle du conseil des ministres, qui devait avoir lieu demain mardi, à l'Élysée, est reportée au lendemain mercredi.

Les titres informationnels à dimension subjective

De façon très majoritaire, les traces de la subjectivité du locuteur qui affleurent à la surface des titres informationnels sont des expressions relevant d'un style familier, comme dans les exemples (99) et (100).

Ch.III (99), « Les Autrichiens se remuent », (1903-03-16)

Vienne, 15 mars.

Le *Tageblatt* annonce que dans les derniers jours de mars, la lourde division de l'escadre austro hongroise, composée de trois croiseurs cuirassés et d'un torpilleur, partira pour la Méditerranée orientale et visitera, à cette occasion les côtes de l'archipel grec et de l'Asie Mineure.

Il y a cinq ans qu'on n'a pas vu d'escadre austro-hongroise dans le Levant.

Ch.III (100), « Le Reichstag se rebiffe », (1898-02-04)

Un vote du Reichstag a réduit de 6,000 marcs les émoluments des secrétaires d'Etat. Ce vote est motivé par le refus du gouvernement d'améliorer les appointements des employés subalternes, alors que l'augmentation du traitement du chancelier et de tous les ministres a été proposée.

On remarquera que derrière l'emploi des prédicats formulés ci-dessus, le journal procède à une lecture subjective de l'information mise en scène dans la brève que le titre surplombe. Ainsi (100) relève de la personnification, définie à la suite de P. Fontanier, comme suit :

La personnification consiste à faire d'un être animé, insensible ou d'un être abstrait et purement idéal, une espèce d'être réel et physique, doué de sentiment et de vie, enfin ce qu'on appelle une personne (Fontanier, 1977 : 112).

Cette personnification s'accompagne en l'occurrence d'une entreprise de *psychologisation* des faits à la suite de P. Charaudeau en cela qu'on observe bien ici une tentative du locuteur de « prêter une intention à des instances collectives ou à des entités abstraites » (Charaudeau, 1997 : 246).

Les titres index

Comme nous l'avons signalé précédemment, les titres-index sont la macro-catégorie la plus abondamment représentée dans le corpus étudié, et ce tout spécialement dans une première tranche diachronique s'étendant de 1883 à 1896 (cf. figure III-29). Examinons les caractéristiques formelles de ces titres qui ne disent de l'information que son horizon d'attente.

Nous avons observé cinq grands types de constructions syntaxiques prégnantes dans cette macro-catégorie. Le titre-index peut ainsi reposer sur un groupe circonstanciel comme l'exemple ci-dessous, codé dans le tableau suivant en tant que type A :

Ch.III (101), « En Roumanie », (1899-02-18)

Bucharest. — Le président de la Chambre a exprimé la sympathie du Parlement roumain pour la France. La Chambre a décidé que son bureau télégraphierait ses condoléances au gouvernement français.

Les types B (exemple 102), C (exemple 103) et D (exemple 104) se rapportent pour leur part à des constructions nominales, respectivement dépourvues de déterminant, comportant un déterminant défini, et intégrant un article indéfini :

Ch.III (102), « Affaires de Portugal », (1890-01-30)

Lisbonne, 29 janvier.

Quelques journaux républicains menacent de publier les noms d'un ou de deux commerçants de Lisbonne et de Porto qui, tout en protestant contre l'ultimatum de lord Salisbury, désirent continuer des rapports commerciaux avec la nation anglaise.

Le ministre de la marine a visité aujourd'hui les navires de guerre portugais en rade de Lisbonne. La réception a été très sympathique. Les coups de canon d'ordonnance ont été tirés. [...]

Ch.III (103), « Les interpellations », (1894-02-02)

L'interpellation Faberot sur l'interdiction de la Bourse du travail viendra à la première séance. M. Touchard, député du Sénégal, renonce à questionner le gouvernement sur l'entrée de nos troupes à Tombouctou.

Ch.III (104), « Un ultimatum », (1893-01-12)

Il est vaguement question d'un ultimatum que l'Angleterre aurait envoyé au sultan du Maroc. On affecte de dire que cette nouvelle n'est pas sans gravité.

Les rares constructions échappant à ces quatre premiers types ont pour leur part été regroupées sous le type « autres ».

	1883	1884	1885	1886	1887	1888	1889
TYPE A	1,8	4,4	2,7	1,0	21,3	10,3	12,3
TYPE B	60,9	35,6	22,2	40,9	19,1	22,5	15,5
TYPE C	36,4	57,0	69,3	51,2	56,9	65,2	66,3
TYPE D	1,3	3,0	5,8	6,9	2,7	2,0	6,0
AUTRES	0,0	0,0	0,0	0,0	0,0	0,0	0,0
	1890	1891	1892	1893	1894	1895	1896
TYPE A	5,2	8,7	2,3	15,4	11,4	8,7	12,0
TYPE B	28,3	24,6	19,9	24,1	28,0	23,0	29,1
TYPE C	61,7	43,7	72,5	57,9	57,5	62,8	53,8
TYPE D	4,8	2,4	5,3	2,6	3,1	5,5	5,1
AUTRES	0,0	20,6	0,0	0,0	0,0	0,0	0,0
	1897	1898	1899	1900	1901	1902	1903
TYPE A	8,4	9,1	13,9	13,4	7,2	12,7	13,4
TYPE B	22,2	21,3	21,4	19,3	18,3	13,5	18,9
TYPE C	63,1	60,9	57,8	61,5	69,3	65,5	55,5
TYPE D	5,8	8,7	6,4	5,3	5,2	7,4	11,6
AUTRES	0,4	0,0	0,5	0,5	0,0	0,9	0,6

Tableau III-8 : Types de construction syntaxique des titres-index (en pourcentage).

Des titres majoritairement construits sur des syntagmes nominaux accompagnés d'un déterminant défini

Le tableau III-8, qui présente les résultats de la classification syntaxique des titres-index, indique de façon très nette que le type C, référant aux constructions nominales dotées d'un déterminant défini, est la construction syntaxique privilégiée dans la production de titre-index. Le titre (105) qui obéit à une telle construction, fixe le cadre thématique de l'information mise en scène dans la brève, en renvoyant à un processus historique par la valeur anaphorique du déterminant défini :

Ch.III (105), « Les remaniements ministériels », (1885-01-08)

Paris, 7 janvier, 7 h. 30 soir.

Le *National* dit qu'il ne sera pas nommé de sous-secrétaire d'Etat à la guerre. Le même journal ajoute que le rattachement des colonies au ministère du commerce est proche et que le décret sera signé incessamment.

Le type B : des titres construits sur un syntagme nominal sans déterminant

Par ordre descendant de fréquence, les syntagmes nominaux dépourvus de déterminant sont la seconde construction syntaxique dans les titres-index. Aucune tendance diachronique nette n'affleure à la surface : construction prégnante en 1883 où elle représente près de 60% des titres-index, le type B ne présente aucune évolution significative de 1884 à 1903. On remarquera cependant que cette construction caractérise moins de 30% des titres-index de 1887 à 1903.

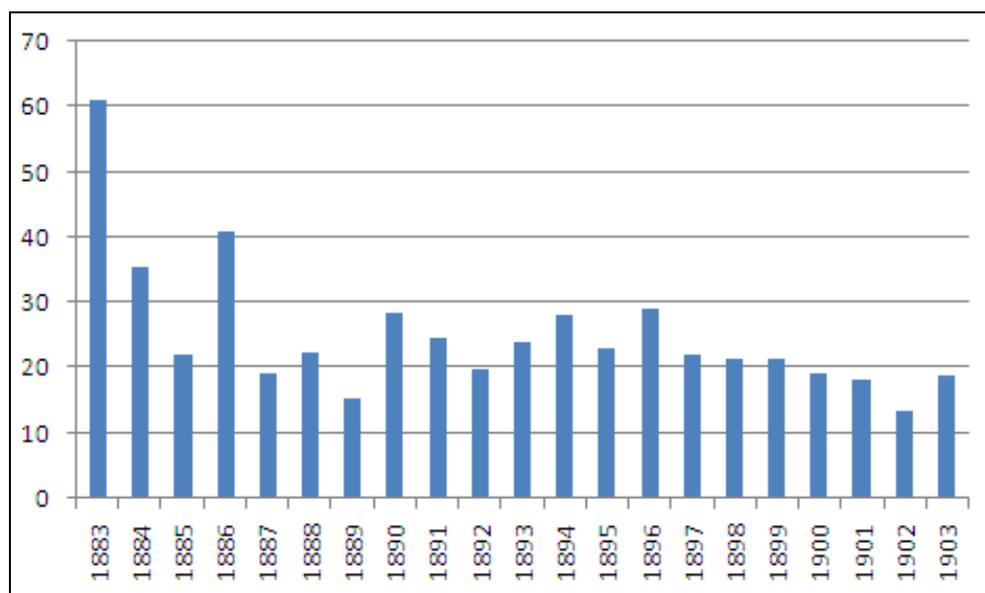


Figure III-37 : La construction syntaxique B dans les titres-index en diachronie (en pourcentage).

L'horizon d'attente que ces titres instaurent est plus ou moins précis : ainsi, (106) ne dit rien sur le cadre géographique ou thématique des nouvelles correspondantes, pas plus que (107). Ce dernier, qu'on peut assimiler à un titre

de rubrique, circonscrit la singularité de la nouvelle dans une classe thématique. Enfin, (108) renvoie à un processus historique : le titre ne condense pas l'information principale de la brève, qui ne traite d'ailleurs pas de l'exécution des sergents à Santo-Domingo, mais des conséquences de cet événement.

Ch.III (106), « Poignée de nouvelles », (1899-02-18)

Paris est très calme, le public s'arrache les journaux du soir criés sur le boulevard. Les groupes de gauche du Sénat, réunis avant la séance en réunion plénière, ont acclamé à l'unanimité la candidature de M. Loubet à la présidence de la République.

La commission sénatoriale du projet de dessaisissement a nommé M. Eugène Guérin, rapporteur.

Ch.III (107), « Nécrologie », (1899-03-25)

On annonce la mort de M. Ch. Nutter, archiviste de l'Opéra, et du général Grimaudet, ancien ministre de la guerre.

Ch.III (108), « Exécution des sergents de Santo-Domingo », (1883-08-15)

L'exécution des sergents appartenant aux troupes insurgées à Santo-Domingo a produit une déplorable impression dans le peuple et un très vif mécontentement dans l'armée.

La haine de cette dernière pour le maréchal Martinez Campos, ministre de la guerre, est pour beaucoup dans le mouvement insurrectionnel. La loi martiale a été proclamée dans les principales villes d'Espagne.

Pour sa part, le recours au groupe circonstanciel, envisagé en tant que type A, connaît une utilisation beaucoup plus discrète que les types B et C.

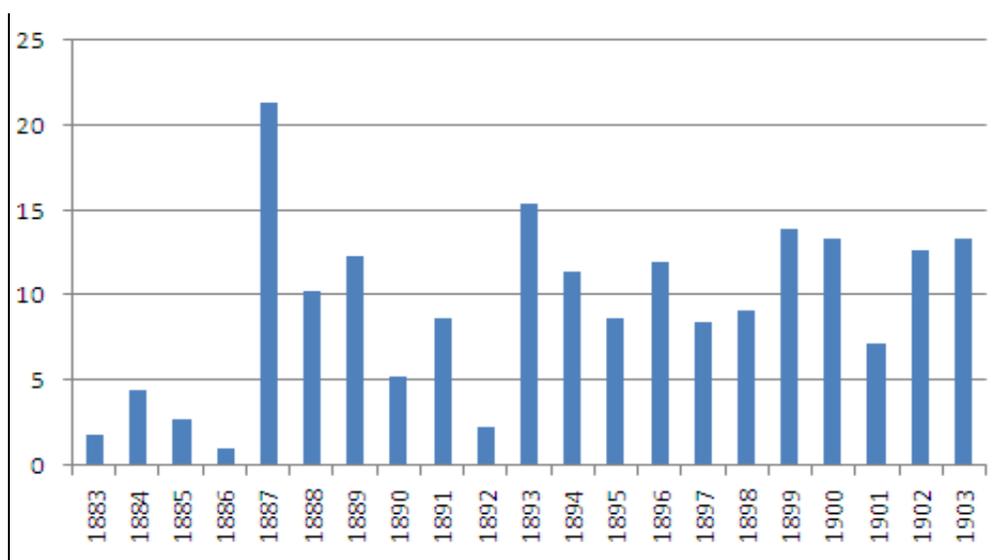


Figure III-38: La construction syntaxique A dans les titres-index en diachronie (en pourcentage).

De 1883 à 1892, l'emploi de groupe circonstanciel est très fluctuant et oscille entre 1,8% en 1883 et 21.3% en 1887, tout en avoisinant les 10,3% en 1888 et les 5,2% en 1890. Le recours à cette construction tend ensuite à se stabiliser à partir de 1893, où les écarts de fréquence d'une année à l'autre se révèlent beaucoup plus faibles : entre 7.2% en 1901 et 13.4% en 1893. Dans les exemples ci-dessous, tout comme sur l'ensemble des occurrences de cette construction dans notre corpus, les syntagmes prépositionnels renseignent uniquement sur le cadre géographique de la nouvelle. Tout en constituant un embrayage spatial, ce type de construction établit une différenciation thématique vis-à-vis des autres dépêches et instaure un classement proche de celui assuré par la rubrique :

Ch.III (109), « Au Père-Lachaise », (1899-02-26)

On a scellé ce matin, au-dessus du caveau des familles Lebault-Faure, qui garde la dépouille du président Félix Faure, la pierre qui en ferme l'entrée.

Ch.III (110), « A Athènes », (1893-01-06)

M. le comte de Montholon, ministre de France en Grèce, a été reçu en audience solennelle par le souverain. M. de Montholon a remis au prince Georges les

insignes de grand-croix de la Légion d'honneur, ainsi qu'une lettre autographe de M. Carnot, président de la République française.

Un faible pourcentage de titres-index construit sur un groupe nominal accompagné d'un déterminant indéfini

Les constructions nominales pourvues d'un déterminant indéfini (ou type D) sont peu utilisées. On remarquera sur l'histogramme suivant que cette construction ne caractérise en moyenne pas plus de 6% des titres-index, à l'exception de certains points saillants en 1886, 1898, 1902 ou 1903.

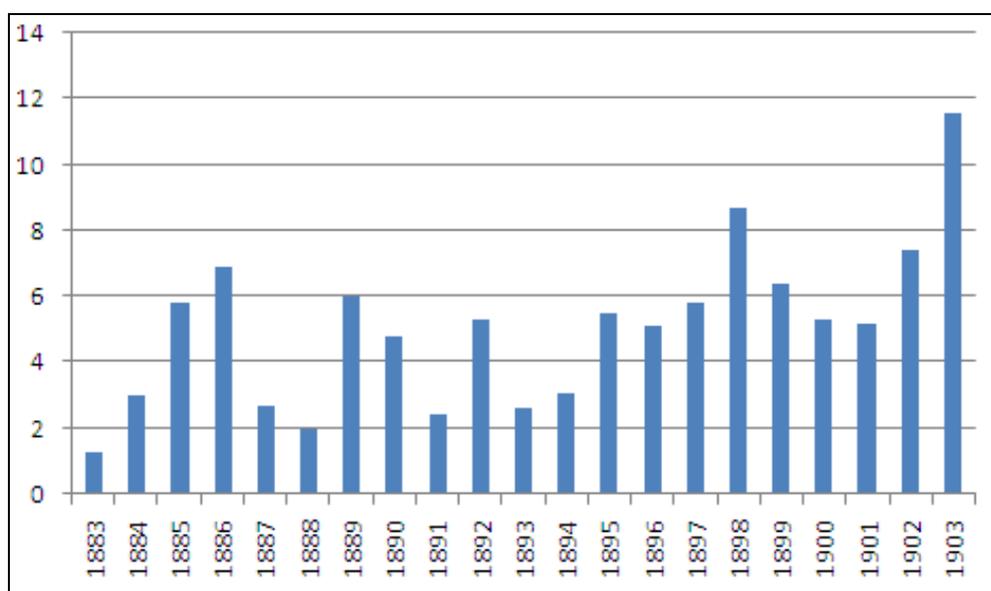


Figure III-39 : La construction syntaxique D dans les titres-index en diachronie (en pourcentage).

Les exemples (111) et (112) illustrent ce type de construction :

Ch.III (111), « Un ultimatum », (1893-01-12)

Il est vaguement question d'un ultimatum que l'Angleterre aurait envoyé au sultan du Maroc. On affecte de dire que cette nouvelle n'est pas sans gravité.

Ch.III (112), « Un démenti » (1897-02-01)

La Epoca déclare sans fondement les bruits d'après lesquels il se serait élevé des divergences entre le ministre des affaires étrangères et M. Taylor, ministre des Etats-Unis.

Sur le plan sémantique, nous concédons, à la suite de J. de Bonville (2008) que le substantif de chacun de ces exemples évoque un schéma actanciel, qui porte en germe une certaine valeur informationnelle. Néanmoins, celle-ci est compromise par l'actualisation incomplète du substantif liée à l'emploi d'un déterminant indéfini et à l'absence de toute prédication spatio-temporelle. De surcroît, les nombreuses occurrences de ces titres viennent confirmer leur statut d'index.

Les titres ludiques et les titres à dimension ludique

Comme l'indique le tableau III-2, les titres ludiques, ayant pour visée principale de « jouer avec le lecteur », sont extrêmement rares dans notre corpus. Nous n'en avons en effet dénombré que trois occurrences, circonscrites à 1899 et 1900, que nous avons déjà observé *supra*.

Le ludique est cependant fortement exploité dans les titres, où il se révèle plutôt convoqué comme dimension. La dimension ludique est tout particulièrement sollicitée dans les titres-commentaires et dans une moindre mesure, dans les titres-index, comme on peut l'observer sur les histogrammes suivants.

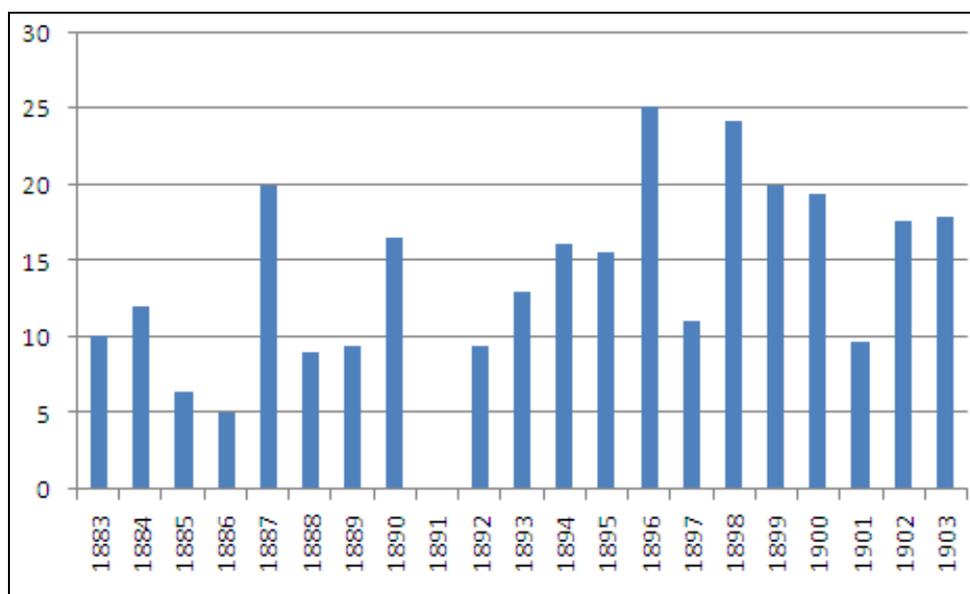


Figure III-40: Pourcentage des titres-commentaires à dimension ludique en diachronie.

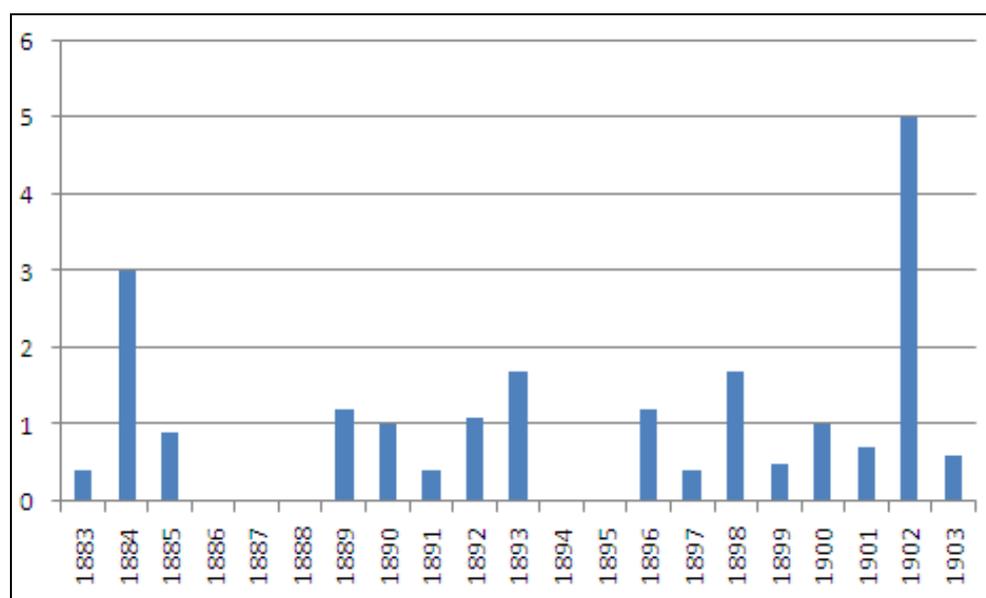


Figure III-41: Pourcentage des titres-index à dimension ludique en diachronie.

L'allusion

Nous poserons l'allusion, définie à la suite de J. Authier-Revuz (1995) comme une forme non-signalée de citation, dont l'interprétation nécessite le recours à l'extralinguistique et à l'interdiscours, et comme l'une des figures rhétoriques principales de l'aspect ludique d'un énoncé. On remarquera d'ailleurs qu'étymologiquement, l'allusion renvoie au jeu, puisque le latin *allusio*, de *ludere*, signifie en effet *jouer*. Le constat de la fréquence des allusions dans les titres du *Petit Comtois* entraîne les questions suivantes : quelles sont les raisons de cette fréquence soutenue, alors que celles-ci, puisant dans un interdiscours plus ou moins maîtrisé par le récepteur, peuvent causer un échec de compréhension ? Quelles sont les valeurs pragmatiques de ces allusions ?

Si l'allusion est à ce point utilisée dans les titres du *Petit Comtois*, c'est tout d'abord parce qu'elle est un jeu avec l'interdiscours que la rédaction propose au lecteur. Les risques de non-reconnaissance de l'allusion (Authier-Revuz, 2000) ont pour pendant le plaisir qu'éprouve le lecteur méritant qui aura décrypté le renvoi allusif. L'allusion s'impose comme l'« indice d'une culture et l'instrument d'une connivence » (Kerbrat-Orechioni, 1977 : 126) : ce procédé

flatte le récepteur détenteur de connaissances encyclopédiques et culturelles tout autant qu'il montre à celui-ci que le journal les partage. Créant un effet de complicité entre les deux acteurs de l'échange et travaillant en toile de fond les « airs » que le quotidien se donne, l'allusion est un facteur de construction de l'*ethos* du *Petit Comtois*.

Tels sont les facteurs qui justifient selon nous l'existence d'une profusion d'allusions dans les titres-commentaires. Pour exemplifier les caractéristiques et les fonctions de ces allusions, considérons les titres (113), (114) et (115). Ces titres partagent le trait de constituer des renvois allusifs à des connaissances culturelles et de relever de la classe des titres-commentaires à dimension ludique. Supposant la perception de l'hétérogénéité du segment, ces titres exploitent le dédoublement du sens et disent plus que ce qu'ils énoncent. Ainsi, (134) renvoie à un célèbre vers de Malherbe, rédigé pour reconforter M. de Perrier de la perte précoce de son enfant. Dans le même temps, par cet énoncé, *Le Petit Comtois* raille la brève carrière du Maréchal Serrano en tant qu'agent diplomatique.

Ch.III (113), « Et rose il a vécu, etc... », (1884-01-21)

M. le maréchal Serrano, duc de la Torre, quitte aujourd'hui Madrid pour se rendre à Paris, où il ne séjournera que fort peu de temps, bien que le nouveau cabinet manifeste le désir de le voir rester encore quelques jours à son poste, jusqu'à son remplacement par un agent diplomatique plus en communion d'idées avec le nouveau ministère.

Mobilisant les connaissances littéraires du lecteur, le titre (114) fait référence à une célèbre phrase d'Alexandre Dumas extraite des *Mohicans de Paris*²⁴.

Ch.III (114), « Cherchez la femme », (1902-03-26)

Il y a une histoire de femme au fond de l'affaire de Fouad-Pacha.

Voici, du moins, ce qu'on se raconte dans les salons du quartier de Pera.

Une très belle femme du harem du général de division de Nazim-Pacha était courtisée par le maréchal Fouad Pacha et le général Echmi Pacha, chef de la police secrète du sultan. Malgré son âge, le maréchal fut préféré par la belle.

De rage, le policier résolu de perdre le maréchal et le mari trompé. Il échafauda toute une histoire de complot et l'on sait le reste.

²⁴ *Il y a une femme dans toutes les affaires; aussitôt qu'on me fait un rapport, je dis : « Cherchez la femme ».*

Quant à la belle qui aurait été la cause de l'affaire, elle a disparu et l'on ignore ce qu'elle est devenue.

Pour sa part, (115) renvoie au titre d'une fable de La Fontaine, pour en appliquer la moralité aux déclarations de M. de Rochefort. Celles-ci sont jugées vaines par *Le Petit Comtois*, car selon lui, M. de Rochefort n'accédera jamais à la fonction de sénateur.

Ch.III (115), « Le renard et les raisins », (1885-01-12)

Des électeurs du quartier de la Folie-Méricourt ont offert la candidature au Sénat à M. de Rochefort, pour protester contre la politique de M. Jules Ferry. M. de Rochefort refuse en disant qu'il n'a aucune aptitude pour le mandat parlementaire et que, s'il était sénateur, il serait moins que rien.

Un segment discursif hétérogène non-signalé peut être dépourvu d'énonciateur original défini. C'est notamment le cas des proverbes, dont l'utilisation est massive dans les titres des dépêches du *Petit Comtois*. Si toute définition du proverbe est malaisée, comme l'ont souligné P.-J.-L. Arnaud (1991) et J.-C. Anscombe (1994), nous nous limiterons à mentionner ces caractéristiques principales: il s'agit d'un énoncé figé, caractérisé par une certaine notoriété dans une communauté linguistique et culturelle donnée, dont la référence énonciative est effacée, et dont l'interprétation repose sur le figuré (Shapira, 1999 : 55-64). En raison des conditions génériques figuratives de son interprétation et parce qu'il est un renvoi à l'interdiscours, nous envisagerons le proverbe comme un des moteurs du ludique. Notons que le calibrage du titre s'accommode mal de la relative longueur observée des proverbes : de fait, un des segments du proverbe est mis en ellipse. Le rédacteur laisse alors le soin au lecteur de le reconstruire. Loin de perturber la réception de l'énoncé, présent dans la mémoire collective, cette mise en ellipse est déjà un jeu avec le lecteur : à lui de finir la phrase du locuteur voire d'en assumer partiellement la responsabilité. Considérons le fonctionnement énonciatif et argumentatif des proverbes présents dans les titres de notre corpus. Nous voudrions insister sur ceci que les proverbes ne sont pas destinés à être interprétés littéralement. Ils servent au contraire une stratégie argumentative comme le souligne Jean-Claude Anscombe :

Un proverbe n'est pas destiné à fournir de l'information par lui-même. Il sert au contraire de cadre et de garant à un raisonnement (Anscombe, 1994 : 106).

Le proverbe constitue un cadre stéréotypique par lequel le locuteur caractérise une situation et se positionne à l'égard de cette dernière. La voix du locuteur qui le profère se mêle ainsi à toutes celles qui ont énoncé avant lui ce même fragment du fonds parémique, au « on » de la « sagesse » commune. De fait, la « sagesse des nations » donne autorité à la voix du locuteur, ce qui nous semble en faire un ressort de prédilection des titres-commentaires. Dans l'énoncé (116), la parole proverbiale est utilisée pour émettre un commentaire implicite sur une action à partir de la projection des caractéristiques d'une situation intrinsèquement absurde : « vendre la peau de l'ours avant de l'avoir tué ». Ce titre disqualifie ainsi, sans formulation axiologique explicite, le traité entre l'Allemagne et l'Italie prévoyant que celle-ci annexerait des régions françaises en cas de guerre qui se solderait par une victoire de cette double alliance.

Ch.III (116), « La peau de l'ours », (1884-01-27)

Paris, 27 janvier, 9 h. soir.

Une dépêche de Toulouse donne comme émanant de bonne source la nouvelle que, par son traité avec l'Allemagne, l'Italie, au cas où elle participerait à une guerre contre la France, devrait avoir Nice et la Savoie, et, de plus, une portion de la Provence jusqu'à Toulon inclusivement.

Si cette nouvelle n'est pas de pure fantaisie, il est édifiant de voir comment M. de Bismark s'y prend pour nous créer des ennemis.

L'exemple (117), donnant à recomposer le proverbe « Charité bien ordonnée commence par soi-même », est également convoqué dans une visée de commentaire de l'information. En prenant les airs du « bon sens », le locuteur exprime son opinion sur la distribution du produit issu des bonnes actions de la presse parisienne, qui doit être selon lui attribué aux seuls nécessiteux français. De même, dans (118), par l'énonciation d'un principe doxique, le locuteur assoit son opinion dépréciative sur la famille Crispi en l'inscrivant sur le « on-vérité » du proverbe (Grésillon, Maingueneau, 1984 : 113). Insistons sur l'*ethos* du journal esquissé par l'utilisation de ces proverbes, par lesquels celui-ci met en

lumière son bon sens populaire, à mettre en rapport avec le lectorat ciblé par *Le Petit Comtois*, c'est-à-dire les petites gens, paysans et ouvriers.

Ch.III (117), « Charité bien ordonnée », (1885-01-12)

On connaît la décision prise par la réunion de la presse parisienne, d'organiser en faveur des pauvres de Paris et de la France une représentation à l'Opéra, un bal costumé à la Bourse, la publication d'un volume d'œuvres inédites constituant le livre d'or de la littérature française, et une loterie de cinq millions à 25 centimes le billet. Depuis plusieurs jours, de grandes discussions existaient au sein des bureaux de rédaction sur la question de savoir si le produit de la fête serait attribué aux pauvres de Paris et aux étrangers. Plusieurs journalistes ont déclaré qu'ils ne prendraient aucune part à l'organisation de cette œuvre, si le produit en devait être porté à l'étranger.

Ch.III (118), « Tel père, Tel fils », (1900-03-26)

On sait que le fils de M. Crispi avait été emprisonné pour un vol de 30,000 fr. au préjudice de la comtesse Celléri, chez laquelle il avait introduit des voleurs qui dévalisèrent tout l'appartement. Le fils de M. Crispi vient d'être condamné à quatre ans de réclusion.

Ch.III (119), « Tant va la cruche à l'eau... », (1896-01-02)

Le *Daily Graphic* énumère les difficultés du cabinet anglais actuel. «La plus grande, dit il, est l'hostilité qui règne partout contre l'Angleterre » ; et il conseille de prendre des moyens de défense en conséquence.

On dénombre parmi les proverbes employés dans notre corpus un certain nombre de proverbes latins. De la même façon que pour les exemples précédents, le locuteur dote sa voix de l'autorité de la « sagesse des nations », et plus, de l'histoire. Si *le proverbe est toujours vieux*²⁵, le latin accentue l'intemporalité de l'assertion. Par ailleurs, comme le latin reste un des piliers de l'enseignement dont ont bénéficié les rédacteurs et (supposément) les lecteurs du *Petit Comtois*, l'utilisation de cette langue affermit bien entendu l'ethos du quotidien, dans le même temps qu'il alimente la connivence avec le lecteur. Ce qu'il nous semble plus important encore de remarquer, c'est que par le recours au latin, le locuteur enracine l'actualité dans une histoire elle aussi classique. Les proverbes « si vis pacem para bellum » et « verba volant, scripta manent »

²⁵ Ce que notent très justement D. Maingueneau et A. Grésillon (1984 : 114) en citant Beaumarchais : « voilà notre imbécile avec ses vieux proverbes ».

résonnent ainsi comme des leçons de sagesse universelles, qui cautionnent ou accusent les événements présentés dans les brèves (120) et (121).

Ch.III (120), « Si vis pacem », (1903-03-12)

Nisch, 12 mars. Les accumulations de troupes turques à la frontière de Bulgarie ont forcé le gouvernement serbe à se préparer à toute éventualité. Ainsi, on étudie en ce moment, à Belgrade, la mobilisation des divisions du Danube et de la Morava, afin de parer à tout événement.

Ch.III (121), « Verba volant, scripta manent », (1903-03-27)

Le bruit court que M. Brogier, inspecteur général du commissariat de la marine, serait sur le point d'être l'objet d'une mesure de rigueur pour avoir adressé, à l'insu du ministre de la marine, une circulaire confidentielle à ses subordonnés au sujet du décret relatif à l'organisation du corps d'administrateurs de l'inscription maritime.

Pour se parer de bon sens, le locuteur responsable du titrage des brèves a même recours au pastiche de proverbes. Ces pastiches ne consistent pas tant à détourner un proverbe existant qu'à détourner les conditions génériques de l'énonciation proverbiale, pour reprendre A. Grésillon et D. Maingueneau :

Nous appellerons *conditions génériques* de l'énonciation proverbiale la réunion des *conditions formelles* (syntaxiques, prosodiques...) [...] et de certains traits pragmatiques, les *conditions d'emploi*. Ces dernières exigent qu'il s'agisse d'une vérité de « bon sens » adressée à un allocataire universel, d'une vérité qui puisse être soutenue par le ON de la sagesse des nations (Grésillon & Maingueneau, 1984 : 115).

Cette définition nous semble caractériser parfaitement le fonctionnement des titres (122) et (123), que nous envisagerons comme des pseudo-proverbes, créés par le locuteur à partir d'un détournement des conditions génériques du proverbe. De celui-ci, (122) et (123) présentent en effet les caractéristiques formelles. La fonction poétique de ces titres est manifeste, reposant sur la prosodie et les rimes en fin des mots *perpétuité* et *propriété* dans (122) ainsi que sur la mise en rapport de deux termes de sens opposés dans (123), en l'occurrence *démentis* et *affirmation*. De plus, chacun de ces exemples recourt à un présent et des syntagmes nominaux génériques, qui lui permettent de prétendre à un statut de vérité générale. S'attribuant les caractéristiques formelles du proverbe, ces pastiches exploitent les traits pragmatiques de ce

dernier : ils feignent d'être investis d'une autorité doxique pour asseoir une entreprise argumentative. Dans (123), le pastiche confère ainsi l'autorité naturelle « du bon sens » à l'opinion exprimée par le locuteur, à savoir que l'état de santé du pape est dégradé, malgré ce qu'en disent les nombreux démentis, confirmant à leur façon, « qu'il n'y a pas de fumée sans feu ».

Ch.III (122), « Jouissance à perpétuité ressemble beaucoup à la pleine propriété », (1903-03-08)

Châteaudun, 7 mars.

Le tribunal civil de Châteaudun a rendu son jugement dans le procès intenté par la ville de Châteaudun aux sœurs de Saint-Jean qui, à la suite de la laïcisation de leur école, avaient refusé de quitter l'immeuble, arguant d'un traité passé en 1829, lequel, prétendaient-elles, leur donnait la jouissance à perpétuité dudit immeuble.

La ville de Châteaudun gagne son procès et les sœurs sont invitées à quitter les lieux.

Ch.III (123), « Trop de démentis valent une affirmation », (1903-03-10)

Rome, 9 mars. On dément de nouveau les bruits alarmants qui ont été répandus sur l'état de santé du pape.

La figure du défigement

La figure du défigement, que nous définirons provisoirement comme la dislocation d'une expression, qui pour le locuteur d'une communauté linguistique et culturelle donnée, constitue un bloc, se rencontre à de multiples reprises dans l'appareil titulaire. Avant de nous concentrer sur l'analyse de ces défigements, il nous semble intéressant de faire dialoguer ce constat avec les réflexions de P. Fiala et B. Habert (1989) sur l'emploi des défigements dans la presse du XX^e siècle, dont la fréquence se justifierait en tant que réaction à la langue de bois des hommes politiques. Dans le cadre d'un examen des jeux de mots posés comme « saturant » la presse française d'opinion et d'information depuis 1970, P. Fiala et B. Habert établissent une définition du défigement insistant sur son fonctionnement proche de l'allusion, dans le sens où l'énonciation du défigement donne à entendre le sens d'un segment hétérogène non-signalé :

On peut définir le défigement dans les titres comme la reprise d'un figement, visant à faire ressurgir tout ou partie du sens originel de ses éléments (Fiala & Habert, 1989 : 89).

Pour faire entendre l'expression figée circulant dans la mémoire collective, le défigement tend à une similitude formelle maximale avec l'expression figée, tout en procédant à une altération de ses composants, de sorte que le double sens créé puisse être réactivé par le lecteur. Selon les auteurs, le défigement a en effet pour objet premier de « produire l'ambiguïté, deux ou plusieurs choses à la fois dans une même expression linguistique » (Fiala & Habert, 1989 : 91).

Mettant l'accent sur ce que P. Fiala et B. Habert nomment la dimension mémorielle du figement, renvoyant à l'ensemble des énoncés perçus comme indivisibles par une communauté linguistique donnée, F. Sullet Nylander (1998) distingue pour sa part les figements linguistiques, désignant « toute expression, locution ou phrase figée dont on ne peut pas changer les termes [...] sans changer le sens et *de facto* produire divers effets de sens » (Sullet-Nylander, 2005 : 117) et les figements culturels, d'autre part, ainsi définis :

Quant aux *figements* dits *culturels*, ce sont des énoncés mémorisés par les sujets d'une même communauté linguistique. Il s'agit de *références culturelles* : titres de livres, de films ou d'autres œuvres répertoriées, ainsi que des proverbes, des phrases entières extraites de chansons, de poèmes ou de divers textes connus d'un grand nombre de francophones (Sullet-Nylander, 2005 : 118).

En nous inspirant de ce *distinguo*, nous analyserons les différents cas de défigements répertoriés dans notre corpus selon qu'ils portent sur des segments mémorisés linguistiques ou culturels. Les cas de cassure de figements linguistiques tels que (124) sont présents dans notre corpus dans une proportion très légèrement inférieure à celles des figements culturels.

Dans l'exemple suivant, le défigement de la locution figée « gros Jean comme devant », dont la popularité est due à la fable *Pierrette et le pot de lait*²⁶ de J.

²⁶ « [...] Les diadèmes vont sur ma tête pleuvant : Quelque accident fait-il que je rentre en moi-même ; Je suis gros Jean comme devant. » La Fontaine, J. (de), *Fables*, VII, 10, éd. H. Régnier, T. 2, p. 154.

de la Fontaine repose sur l'homonymie. Le groupe désignatif « gros Jean²⁷ » se voit substituer le patronyme d'une des personnalités dont il est question dans la dépêche. Ce défigement sert une visée de commentaire : il s'agit bien ici de produire un jeu de mot, tout en exprimant une opinion à l'égard des événements rapportés. En l'occurrence, (124) exprime la déception du locuteur devant l'absence de faits nouveaux ayant une certaine importance dans la dépêche correspondante.

Ch.III (124), « Grosjean comme devant », (1899-02-27)

La cour de cassation, réunie en conseil supérieur de la magistrature, s'est occupée de l'incident Grosjean. M. Faure-Biguet a été nommé rapporteur.

S'il obéit au même principe de défigement par homonymie tel que mobilisé dans l'exemple précédent, (125) brise un figement culturel. Construit à partir d'une substitution de la graphie *Lagarde*, patronyme du ministre de France en Ethiopie en poste jusqu'en mars 1903, à la graphie originale *La garde*, (125) renvoie à la célèbre phrase de Pierre Cambronne. Cet exemple nous semble doté d'une certaine valeur informative, bien que relativement opaque : dans le même temps qu'il joue sur la mémoire discursive, cet énoncé informe d'un changement d'affectation, avec nouvelles perspectives. Il reste un jeu de mot relativement gratuit.

Ch.III (125), « Lagarde meurt mais ne se rend pas », (1903-03-06)

M. Lagarde, ministre de France en Ethiopie, sera prochainement appelé, assure-t-on, à un autre poste.

M. Delcassé entretiendra le prochain conseil des ministres du choix de son successeur.

Le titre (126) correspond également à un jeu sur un figement culturel qui est le proverbe « l'argent ne fait pas le bonheur ».

Ch.III (126), « La triplice ne fait pas le bonheur », (1898-03-02)

Le *Corriere della Sera* se livre à des réflexions amères, au sujet du rappel de la classe 1874, sur les conditions actuelles de l'armée italienne :

« Malgré les milliards dépensés, malgré les centaines de millions nécessaires à son entretien annuel, nous avons une armée qui, en temps de paix, ne suffit

²⁷ La locution figée *gros Jean*, désignant un homme rustre, trouve également une attestation sous la graphie *Gros-jean* dans le TLFi.

même pas aux besoins de la sûreté publique. En voulant maintenir douze corps d'armée par obstination ou obligation, nous obtenons des bataillons comptant à peine cent cinquante hommes. Avec douze corps d'armée nominaux, l'Italie est un pays faiblement constitué en temps de paix, et le cas de guerre présente un danger. Un tel pays ne peut être apprécié même par ses alliés et il est déprécié par les autres. »

Polysémie, paronymie

L'expression évoquée par l'énoncé (127) ne peut être considérée comme figée. Néanmoins, il y a bien une mémorisation de l'expression utilisée, qui crée un horizon d'attente. L'effet ludique joue de cette attente par la substitution d'un paronyme. L'exemple (127) constitue un défigement de la locution figée « Echec au roi » par laquelle un joueur avertit son adversaire qu'il met en danger son roi. C'est donc la proximité sémantique des termes *chancelier* et *roi* qui sont le ressort du jeu de mot, qu'on qualifiera de défigement par parasynomie.

Ch.III (127), « Echec au chancelier », (1885-01-11)

M. de Bismark vient encore d'être battu au reichstag sur la question des crédits pour la politique coloniale. Dans le cours de la discussion, M. de Bismark a dit qu'il était nécessaire de se dépêcher dans le choix d'une occupation, car les autres nations pourraient bien prendre l'avance.

Des jeux de mots pour dire l'autre

Dans le cadre d'une mise en lumière des ressorts de la dimension ludique des énoncés, il nous semble inévitable d'aborder les pratiques de désignation des personnalités et des peuples dans notre corpus. Cette classe d'observation se double d'un fort intérêt socioculturel en cela qu'elle renvoie au contexte d'impérialisme colonial exacerbant les nationalismes européens florissants et d'essor de l'immigration que connaît la fin du XIX^e siècle. C'est dans ce contexte que la référence et la description de l'Autre, donne à lire le Nous à travers la *doxa*, c'est-à-dire des représentations sociales prédominantes, qui sont autant d'informations sur la communauté linguistique, culturelle, voire socio-politique dont elles émanent.

Nombreux sont les désignants de l'autre qui reposent sur un élément stéréotypique en tant que déjà-dit idéologique embusqué dans le discours. Par ce présupposé, posé comme savoir de sens-commun, le journal, sphère discursive d'influence, véhicule une représentation de l'autre « déjà pensée », servant un projet argumentatif subreptice. Dans le même temps, en recourant au stéréotype, le journal raffermi son *ethos* et active sa connivence avec le lecteur, avec qui il partage ses représentations. En désignant l'autre, le journal joue sur les « évidences » qu'il partage avec le lecteur, derrière lesquelles s'inscrivent en palimpseste le culturel et le socio-politique. Le journal laisse le soin au lecteur de (re)construire ce stéréotype au fil des numéros, des pages et des colonnes²⁸.

Abordons dans cette perspective, l'examen des désignants des Allemands et des Italiens dans les titres des dépêches du *Petit Comtois*. Rappelons encore une fois que ces titres ont pour enjeu de colorer les informations présentées dans les brèves correspondantes, rédigées par un autre locuteur que le secrétaire de rédaction du journal. Les désignants des Allemands et des Italiens ont pour caractéristique de se baser dans notre corpus sur la mise en évidence d'un élément culturel ayant trait cette fois à leurs stéréotypes culinaires. La nation allemande et ses citoyens sont ainsi référés par le terme « choucroutmann », comme on peut l'observer dans les exemples (128), (129) et (130). Résultant de la contraction du lexème *choucroute* et du substantif emprunté à la langue allemande *mann*, signifiant *homme*, ce substantif est employé comme antonomase. Il nous est délicat de nous prononcer sur le statut de ce terme : dans nos recherches, nous n'en avons trouvé aucune autre occurrence, ce qui est cependant loin de nous permettre de décréter que ce terme est un néologisme, ou toute forme idiolectale propre au *Petit Comtois*. Il peut tout autant s'agir d'une réminiscence de chansons ou de poèmes populaires tels que produits après la défaite de 1870. Il reste que ce terme reflète et alimente l'anti-germanisme et son corollaire, le militantisme

²⁸ Nous abonderons de fait dans le sens de R. Amossy, selon laquelle le stéréotype est une construction de lecture (1991 : 21-22).

revanchard ambiant à mettre en rapport avec l'issue de la guerre franco-prussienne. Le terme « Choucroutmann » n'entretient pas uniquement la connivence entre le lecteur et le quotidien parce qu'il constitue un jeu de mots ; il alimente la connivence avec le lecteur en mettant en lumière une idéologie partagée.

Ch.III (128), « Choucroutmann prend ses précautions », (1898-02-04)

Deux importantes maisons de Yokohama ont reçu du gouvernement allemand de grosses commandes de fournitures de vivres pour le corps d'occupation de Kia-Tchéou. Le gouvernement voudrait par là éviter le risque d'empoisonnement que courraient ses soldats si les adjudications étaient remises à des Chinois, se souvenant en cela de l'histoire des premières années de l'occupation de Hong-Kong par les Anglais.

Ch.III (129), « Choucroutmann travaille ferme », (1903-03-27)

Kiel, 25 mars. — Une activité extraordinaire règne actuellement dans les chantiers maritimes allemands. Huit vaisseaux d'escadre sont en voie de construction. Tous, sauf le Schwaben, ont été commandés à l'industrie privée. Le chantier impérial, à Wilhelmshaven, ne fait que des réparations, et cela pour cause : l'amirauté n'a guère eu à se louer des constructeurs officiels.

Ch.III (130), « Choucroutmann se promène », (1902-03-02)

Washington, 1er mars.

Le prince Henri de Prusse a dîné, hier soir, à l'ambassade d'Allemagne.

Le monde diplomatique et officiel était représenté au banquet.

A onze heures trente, le prince a quitté Washington ; il se propose de visiter les États de l'Ouest et du Sud-Ouest.

Les Italiens ne sont pas plus épargnés dans les titres des dépêches du *Petit Comtois* que dans ses pages intérieures. Le gouvernement italien et les Italiens sont ainsi désignés de façon dépréciative par le substantif hautement stéréotypique *macaroni*, érigé en nom propre. Reposant sur une allusion à un élément culinaire posé comme typiquement national, les titres (131) et (132) témoignent de telles personnifications empreintes de sentiments xénophobes. Au travers de cette antonomase, plus qu'un jeu de mots, *Le Petit Comtois* cherche à instaurer une connivence avec son lecteur en exploitant une véritable injure raciste :

Ch.III (131), « Macaroni bien ennuyé », (1902-03-02)

Naples, 1^{er} mars.

Le gouvernement est entré en négociation avec la fille de Crispi, pour arriver à un arrangement à l'amiable dans l'affaire, des papiers laissés par l'ancien président du conseil.

La fille de Crispi s'est déclarée prête à restituer les documents d'Etat, mais elle ne veut pas que, sous prétexte de raison d'Etat, on fasse disparaître certaines lettres privées qui sont très compromettantes pour plusieurs hommes politiques.

Ch.III (132), « Ce que pense Macaroni », (1900-01-16)

L'*Opinione* critique sévèrement la réponse de l'Angleterre faite au président Kruger au sujet de la proposition de paix :

« Quelles victoires, dit-elle, ont donc obtenues les Anglais pour prétendre enlever aux deux républiques le bien suprême de leur indépendance et de leur liberté ? Le plus grand capitaine des temps modernes, après être entré deux fois victorieux à Vienne et une fois à Badin, laissa non seulement exister l'Autriche et la Prusse, mais il agrandit la première. »

Considérons encore les titres suivants. Aux antipodes de tout respect culturel, ces deux exemples comportent le terme *magot* désignant un « bibelot figurant un personnage plus ou moins grotesque, sculpté ou modelé, provenant ou imité de l'Extrême-Orient »²⁹. À partir d'un trope renforcé par une personnification, (133) et (134) réfèrent ainsi indifféremment à tous les peuples de l'Extrême-Orient. Ce désignant traduit et véhicule la xénophobie du quotidien. Cette violence verbale, inimaginable de nos jours, même dans la presse pamphlétaire, mériterait une analyse à part entière.

Ch.III (133), « Magot s'agite de nouveau », (1902-03-05)

Shang-haï. — La révolte qui a éclaté dans le district de Nan-Ning est plus grave qu'on ne le croyait, d'après les premiers renseignements. Environ mille soldats licenciés sont parmi les rebelles.

Nan-Ning est une ville de 50,000 habitants de la province de Quang Si ; elle est située à 100 lieues de la frontière française.

C'est dans la ville de Nan-Ning que l'insurrection des Taipings a pris naissance en 1849.

Ch.III (134), « Visite des magots », (1900-03-21)

De la *Politique coloniale* :

Nous croyons savoir que le gouvernement de l'Indochine a décidé le roi du Cambodge, l'empereur d'Annam, et deux rois du Laos à accepter l'invitation de se

²⁹ TLFi, en ligne, consulté le 10-05-2008.

rendre à Paris dans le courant de l'été, avec une suite nombreuse pour visiter l'Exposition.

Pour désigner les anglais, *Le Petit Comtois* recourt au nom de John Bull, personnage symbolisant l'Angleterre et ses citoyens. Parangon de l'Anglais typique créé par John Arbuthnot en 1712, ce désignant est très communément utilisé au XIX^e et XX^e siècle, contrairement à l'antonomase *Choucroutmann*, qui pour sa part, n'a pas traversé les frontières et les années et s'est inscrit dans la mémoire collective. Ce qu'il nous semble intéressant de relever dans les exemples (135) et (136), c'est le processus par lequel la rédaction laisse au soin du lecteur de (re) construire le stéréotype. Au travers d'une expression figée dépréciative et d'un présent à valeur permanente, le titre (135) dresse le portrait d'une Angleterre aux ambitions coloniales démesurées, que critique le quotidien. Le titre (136) dénonce pour sa part la fourberie de l'Angleterre, en procédant à une lecture subjective de la stratégie de ce pays, jugée déloyale par *Le Petit Comtois*. Dans (137), il faut entendre, dans un contexte où le nationalisme est exacerbé, la jubilation du *Petit Comtois* devant l'échec des anglais, qui se voient refuser l'autorisation d'exploitation d'une concession minière.

Ch.III (135), « John Bull a les dents trop longues », (1898-02-14)

A l'occasion d'une récente audience, accordée par M. Rhodes à des membres de la Ligue impériale, M. Rhodes a exprimé le vœu de voir les colonies du Cap, de Natal et de la Rhodesia, contracter une union apte à contrebalancer l'importance de la fédération entre le Transvaal et l'Etat libre d'Orange. Le *Volkstein* de Pretoria tourne en ridicule cette proposition, en faisant remarquer que la colonie du Cap, étant composée surtout d'éléments africanders, ne consentirait jamais à contracter une union qui serait créée contre les Etats boërs, tandis que la colonie de Natal ne désire pas prendre une attitude hostile à ces Etats, parce qu'elle dépend d'eux économiquement.

En ce qui concerne la Rhodésie, le *Volkstein* dit que ni la colonie de Natal, ni celle du Cap ne souhaitent une fédération avec une société qui, jusqu'ici, ne s'est pas encore développée au point de former un Etat.

Ch.III (136), « John Bull veut faire un coup de Jarnac », (1898-02-12)

D'après une dépêche de Singapour au *Daily Mail*, l'Amirauté a donné l'ordre à l'amiral Bridge d'acheter tout le charbon gallois qui se trouve dans les dépôts anglais d'Extrême-Orient, ainsi que dans les ports chinois et japonais.

Le résultat de cette mesure est que les autres escadres croisant dans les mers de Chine éprouvent de grandes difficultés à se ravitailler en charbon et sont pratiquement immobilisées. [...]

Ch.III (137), « Attrape John Bull ! », (1902-03-03)

La *Politische Correspondenz*, qui possède des sources spéciales d'information, annonce que l'empereur Menelik a décidé de refuser à la « British-Abyssinian Exploration Company limited » l'autorisation d'exploiter l'importante concession minière du Sud de l'Abyssinie, concession qui avait été payée par la Compagnie 2,125,000 fr. On prétend que la décision prise par Menelik est due à l'influence française.

Dire l'autre, c'est aussi référer à une personnalité. Pour dire la personnalité « adversaire », le locuteur responsable des titres recourt fréquemment à des surnoms qui circulent dans l'interdiscours. Ainsi, l'héritier Napoléon Joseph Charles Paul Bonaparte (1822-1891), est affublé du surnom *Plon-Plon* dans le titre (138), qui est le nom sous lequel il est moqué par les des républicains de la Troisième République. S'il est évocateur pour le lecteur contemporain de la parution du quotidien, l'emploi de ce surnom répond aux contraintes de condensation pesant sur le titre, tout en mettant en lumière des représentations partagées. Pour sa part, le surnom de *Touchatout*³⁰ attribué à Guillaume II semble une initiative propre au *Petit Comtois*, qui se limite à affubler ce personnage de l'un des surnoms les plus usités au XIX^e siècle³¹.

Ch.III (138), « Plon-Plon veut manifester », (1883-08-30)

On sait que les journaux ont prêté au prince Napoléon l'intention de se livrer à une nouvelle manifestation à l'occasion de la mort de M. de Chambord.

Le prétendant bonapartiste est rentré à Paris, arrivant de Hollande, et l'on affirme que la mort de M. de Chambord est l'unique cause du retour du prince.

Ch.III (139), « Les projets de Touchatout », (1898-02-19)

Le bruit court que l'empereur aurait l'intention de clôturer la session du Landtag le 15 mars, afin que les élections de la Chambre des députés aient lieu avant celles du Reichstag. Si les élections de la Chambre sont fixées avant le 15 juin, la Chambre ne sera dissoute que pour la forme. Le gouvernement procéderait ensuite aux élections du Reichstag, pour profiter de l'effet moral de l'élection d'une majorité réactionnaire au Landtag, élection qui est certaine avec le suffrage censitaire.

³⁰ Généralement orthographié : « Touche-à-tout ».

³¹ Le plus connu des « Touchatout » de la fin du XIX^e siècle et du début du XX^e siècle reste très certainement Léon-Charles Bienvenu (1835-1911), journaliste et écrivain français.

Ch.III (140), « Trop de zèle, ami Touchatout ! », (1900-01-18)

Londres. Le *Manchester Guardian* dit apprendre que Guillaume II, outre la dépêche de félicitations qu'il a adressée à la Reine Victoria, lors de la reddition du général Cronje, a envoyé un message direct à lord Robert, se félicitant de ce même événement.

3.3.3.4. Remarques conclusives sur l'appareil titulaire des *Dépêches de nuit*

Au terme de cet examen, nous proposerons quelques remarques synthétiques concernant les propriétés inhérentes à l'appareil titulaire de la rubrique *Dépêches de nuit*.

La première caractéristique du périphrase des brèves de cette rubrique est tout d'abord un calibrage très court : un titre est en moyenne composé de 3,5 mots. La longueur limitée de ces titres a pour corollaire un fort emploi de la nominalisation et une faible actualisation des syntagmes nominaux, indices de la fonction d'indexation du périphrase de l'article, qui prédomine dans le corpus. La majorité des titres des dépêches, et ce, tout spécialement sur la période 1885-1896, est en effet vouée à catégoriser l'information, comme en témoigne la prégnance des titres-index. Ces derniers inscrivent la singularité de la nouvelle dans la continuité d'un processus historique ou la typifient en la rattachant à une famille événementielle, voire ne précisent que son horizon géographique. Cette façon de mettre en scène l'actualité dans des cadres stabilisés nous semble une caractéristique essentielle de la presse du XIX^e siècle.

Au contraire des titres-index, les titres informationnels, exhibant l'actualité dans sa dimension singulière en condensant l'information principale de l'article, constituent une part minoritaire de l'appareil titulaire observé. Il est très intéressant de noter qu'aucune tendance lourde ne se dégage concernant leur emploi en diachronie : de 1883 à 1903, les titres informationnels restent minoritaires. En moyenne plus longs en raison de leur complétude, les titres de cette classe reposent le plus couramment sur un syntagme nominal accompagné d'un syntagme prépositionnel. S'il existe à partir de 1883, le

modèle du titre informationnel à partir d'une phrase verbale complète, lui, ne se stabilise quantitativement qu'à partir de 1896, date à laquelle l'emploi des titres index tend à s'amoinrir. Alors que l'accroissement des dépêches est généralement envisagé comme le symptôme du passage d'une presse d'opinion à une presse d'information, comme une rupture avec la presse de la première moitié du XIX^e siècle, l'appareil titulaire exploité de 1883 à 1903, pour sa part, n'a pas pour priorité de dire l'actualité, mais bien de l'organiser.

Si, *a priori* les dépêches privilégient le factuel, on ne manquera pas de souligner l'utilisation importante de titres-commentaires qui colorent l'information brute. Ces derniers, reflets de l'inadéquation énonciative entre la brève et son péritexte, sont le lieu d'une reprise de possession de la parole. En exprimant son opinion sur le contenu de la brève, *Le Petit Comtois* opère un contrôle de la voix hétérogène se donnant à lire dans l'article. Il la dénonce même, en pointant les non-dits et les formes discursives neutralisantes qui se déploient dans les brèves. Traces de la subjectivité du locuteur et de sa forte prise en charge de l'énoncé, les titres-commentaires se distinguent par leurs constructions syntaxiques, qui reposent de façon très significative sur des phrases verbales complètes et des modalités exclamatives, impératives ou interrogatives, mais aussi sur un style familier. En simulant l'oralité, en employant des expressions familières, la rédaction instaure par ce modèle de titre un dialogue virtuel avec son lecteur et entretient une accointance avec lui. Bien plus, il affirme son ethos de quotidien régional proche de son lectorat. Il n'est pas indifférent, dans cette perspective, que la majorité des titres à dimension ludique ait une visée de commentaire : l'allusion et les jeux de mots pour dire l'autre mettent en exergue les connaissances partagées de la rédaction et de son lectorat : à travers cette parole joueuse, sont construites et réaffirmées la connivence et l'appartenance à une même communauté socio-politique. Aux affectations du discours officiel présenté dans les brèves, *Le Petit Comtois* répond par le franc-parler. De la même manière, le quotidien, par le recours au proverbe, se met en scène comme la voix du « bon sens » tout en

exploitant l'autorité de la « sagesse des nations » pour asseoir ses opinions. Plus encore, prendre pour intitulé un proverbe, c'est conférer au titre la mission d'inscrire l'actuel dans l'intemporel et d'atténuer la singularité de l'actualité, dont le statut d'événement est estompé. Catégorisant l'information, la typifiant, atténuant sa singularité plus que l'exposant, les titres des dépêches du *Petit Comtois* ne reflètent ainsi aucune rupture radicale avec les pratiques discursives journalistiques de la première moitié du XIX^e siècle.

Conclusions sur la mise en forme de l'information dans *Le Petit Comtois*

Au cours de cette analyse, nous avons mis en évidence certains traits caractéristiques de la mise en forme de l'information dans *Le Petit Comtois*.

Nous avons ainsi pu constater que le terme de « mise en page » ne pouvait être utilisé, en toute rigueur, pour référer aux pratiques de mise en forme de l'information dans le quotidien sur la période étudiée. Rudimentaire, sa présentation repose presque exclusivement sur la seule dimension verticale de la page. L'aire scripturale n'est en l'occurrence pas aménagée pour faire cohabiter plusieurs surfaces distinctes mais uniquement pour les faire se succéder. L'austérité est de mise dans les trois premières pages du journal. Ces pages, que nous qualifierons de rédactionnelles, ne témoignent ainsi que très exceptionnellement d'une tentative de mise en relief esthétique ou même d'illustration : l'utilisation de caractères typographiques diversifiés et sophistiqués est l'exclusivité du placard publicitaire. Agrémentant très fréquemment les encarts publicitaires, les illustrations apparaissent pour leur part en 1893 dans les pages « rédactionnelles » du quotidien et sont utilisées avec parcimonie, sans que ne se dégage une réelle stabilité concernant leur emploi en diachronie.

Dans cet espace textuel dense, le rubriquage constitue un premier niveau de mise en forme. Organisant les nouvelles dans des espaces pré-

ordonnancés, il structure la masse informationnelle hétérogène. Élément stable d'un numéro à l'autre, le rubriquage ritualise la lecture du quotidien. Il se révèle être un élément essentiel tant pour la production que pour la réception. Pour l'analyste, le rubriquage constitue un lieu privilégié d'accès à l'identité du quotidien, en cela qu'il porte trace de la valeur attribuée par la rédaction à une information et qu'il reflète les choix idéologiques sous-jacents au découpage du monde ainsi effectué. Au gré d'une analyse quantitative des rubriques, nous avons dépeint une première esquisse du *Petit Comtois*. Si la place primordiale accordée aux événements locaux présentés dans la rubrique *Chronique régionale* confirme son statut de presse régionale, il reste que *Le Petit Comtois* est loin d'être centré sur les événements de proximité. Le rubriquage du quotidien fait en outre la part belle à la politique, perceptible à travers la récurrence des comptes rendus du Sénat et de la Chambre des députés, rubriques à part entière jusqu'en 1886, date à laquelle elles deviennent des sous-rubriques des *Dépêches de nuit*. Traitant de l'actualité nationale et internationale, cette dernière rubrique constitue l'autre bloc essentiel du quotidien, reflétant, outre l'ouverture de celui-ci aux faits politiques, l'importance accordée au factuel. L'emplacement topographique stratégique qui lui est accordé de 1883 à 1903 confirme la mise en valeur de l'importance capitale attribuée à l'information par le *Petit Comtois* : en les plaçant en Une, le quotidien expose à la vue de chacun ses dépêches, et par là même, se présente comme un quotidien d'information. Cependant, si le quotidien valorise certes sa mission d'information, il n'en demeure pas moins un quotidien d'opinion. Le factuel brut des dépêches est en effet encadré en amont par la rubrique du *Bulletin*, au travers de laquelle la rédaction livre sa lecture des événements politiques du jour et par des articles non-rubriqués dont les caractéristiques formelles et discursives sont proches de la forme générique de l'éditorial qui nous est contemporain. Suivent des articles caractérisés par une grande prise en charge énonciative, eux-aussi non-rubriqués. Encadrée par des articles d'opinion, l'information brute des dépêches est subordonnée à la voix

du quotidien qui exploite le factuel pour mieux confirmer et véhiculer sa lecture du monde. L'analyse du rubriquage permet, dans un premier mouvement, de définir *Le Petit Comtois* comme un quotidien d'opinion où l'information occupe une place privilégiée en vertu de la modernité de la technique qu'elle symbolise.

Le critère d'ordre temporel sous-jacent à un grand nombre de rubriques comme *Dépêches de nuit*, *Dernière Heure*, *Au jour le jour* met en exergue la profonde modification du rapport à la temporalité, permise par les technologies télégraphiques. Par ces intitulés mettant en avant un traitement rapide de l'actualité, la rédaction cherche à simuler la coïncidence de la temporalité du monde avec celle du journal. Par là-même, le quotidien met en scène sa modernité. Il convient cependant de noter le contraste entre la modernité que cherche à inspirer le quotidien avec des traits caractéristiques de la presse du début du XIX^e siècle : rareté des rubriques thématiques d'une part, recours à des intitulés de rubriques reposant sur la généralité d'autre part. Des intitulés tels que *Faits Divers*, *Variétés* ou *Causerie* dénotent la persistance des liens entre la matrice littéraire et la presse du XIX^e siècle : *Le Petit Comtois* découpe ainsi son propre discours à partir de catégories empruntées à la littérature. En instaurant ces cases, le quotidien ne réfère pas à autre chose qu'à des genres et témoigne de son incapacité à découper le monde en thématiques. Partant, le statut même de *rubrique* de ces cases peut être discuté.

À l'image du rubriquage, l'appareil titulaire du *Petit Comtois* demeure rudimentaire. Le péri-texte de l'article est rare bien qu'il tende à se développer en diachronie ; les fonctions du péri-texte du journal et du péri-texte de l'article sont parfois confuses. Afin d'analyser l'évolution de la forme et des fonctions des titres du *Petit Comtois*, nous avons constitué un corpus de titres observés dans une des rares rubriques disposant d'un appareil titulaire permanent : les *Dépêches de nuit*. Nous avons voulu observer les formes et fonctions de l'appareil titulaire de cette rubrique phare du quotidien, par laquelle *Le Petit Comtois* montre sa capacité à informer son lectorat des événements venant d'arriver n'importe où dans le monde.

Partant de l'hypothèse théorique qu'un titre peut constituer une unité autonome, mais que ses fonctions discursives et pragmatiques restent étroitement liées à l'unité textuelle qu'il surplombe, nous avons distingué quatre visées principales du titre : condenser l'information principale de l'article (*titre informationnel*), classer l'information (*titre index*), commenter l'information (*titre commentaire*) et divertir le lecteur (*titre ludique*).

Cette classification, loin de montrer une correspondance entre la modernité des dépêches et celle de son appareil titulaire, a pointé que les titres de cette rubrique, d'un calibrage très court (environ 3,5 mots par titre), ont majoritairement pour fonction de catégoriser l'information et non pas de la condenser. L'appareil titulaire, de 1883 à 1903, n'est pas tant voué à prédire l'information traitée dans l'article qu'à la rattacher à un horizon géographique, à un processus historique ou encore à une catégorie événementielle. Si le pourcentage de titres-index tend à s'amoinrir de 1896 à 1903, il est d'ailleurs stimulant d'observer que ce n'est pas au profit de titres condensant l'information principale de l'article. Aucune évolution quantitative des titres-informationnels ne se donne d'ailleurs à lire dans notre corpus sur la période 1883-1903. Ce sont les titres exprimant un commentaire sur l'article et l'information qu'il délivre qui sont quantitativement la deuxième macro-catégorie prégnante. Exprimant le point de vue du quotidien sur l'information, les titres-commentaires colorent l'information d'une sensibilité politique, et par là, influencent la réception du lecteur. Mettant en lumière la non-coïncidence énonciative entre l'article et son péri-texte, les titres-commentaires reprennent ainsi le contrôle sur l'information telle qu'elle est présentée dans les brèves. Enfin, au discours officiel, *Le Petit Comtois* oppose le franc-parler et une parole railleuse et ludique. Simulant l'oralité, se mettant en scène comme la voix du « bon sens », la rédaction instaure au gré de ces titres un dialogue virtuel avec son lecteur et entretient ainsi une connivence avec lui. Cette macro-catégorie de titres a donc pour rôle d'établir une connivence entre le journal et son lecteur bien plus que d'informer.

De l'analyse de la mise en page, en passant par l'analyse du péri-texte du journal à celui de l'article, notre corpus ne donne à lire aucune rupture radicale en diachronie. Il livre cependant de précieux renseignements sur l'ethos du *Petit Comtois* et sur l'état des pratiques journalistiques. *Le Petit Comtois*, tout en s'ornant des marques de la modernité et en exhibant sa mission d'information, demeure marqué par les traces persistantes des pratiques journalistiques du début du XIX^e siècle. Journal d'opinion faisant la part belle à l'actualité pour mieux conforter sa lecture du monde, *Le Petit Comtois* est avant tout un quotidien régional républicain, ayant pour ambition d'éclairer ses lecteurs sur les événements de la vie politique de la Troisième République.

Table des matières

[TOME 1]

INTRODUCTION GÉNÉRALE	1
CHAPITRE I.	
EXPLORER LA PRESSE RÉGIONALE DE LA TROISIÈME	
RÉPUBLIQUE : ÉTAT DE L'ART	9
1.1. Approches de la presse régionale de la Troisième République	
par les sciences humaines.....	10
1.1.1. Un champ de recherche peu investi par les sciences de	
l'information et de la communication	
ainsi que par l'analyse du discours	10
1.1.1.1. En analyse du discours.....	10
1.1.1.2. En sciences de l'information et de la communication	12
1.1.2. Une presse essentiellement étudiée par les sciences	
historiques	13
1.1.2.1. Bref panorama des recherches en sciences historiques	
sur et/ou par la presse régionale de la Troisième République	14
1.1.2.2. Les facteurs qui accentuent le retard de développement	
de l'étude de la presse régionale de la Troisième République	19
1.2. Le discours et sa dimension matérielle.....	24
1.2.1. Regards sur le <i>linguistic turn</i>	24
1.2.1.1. Une remise en cause des moyens	
de la connaissance historique.....	25
1.2.1.2. La controverse sur l'objectivité et la vérité.....	28
1.2.1.3. <i>L'histoire des concepts</i>	30
1.2.2. <i>L'analyse du discours du côté de l'histoire</i>	34

1.2.2.1. Retour sur la démarche inaugurale de l'analyse du discours du côté de l'histoire	35
1.2.2.2. La démarche configurationnelle	39
1.2.2.3. Vers l'histoire linguistique des concepts.....	43
1.2.3. L'analyse textuelle du discours	47
1.2.4. La dynamisation de notre champ de recherche est conditionnée par la constitution de vastes ensembles de données textuelles d'intérêt historique	51
1.2.4.1. Peu d'expériences inaugurales prennent la PQR3R pour objet	51
1.2.4.2. Retour sur des entreprises pionnières de constitution de base textuelle de presse ancienne	53
1.3. Construire une base textuelle de presse ancienne : aspects méthodologico-techniques	56
1.3.1. Constitution de données textuelles	57
1.3.1.1. Renforcer le statut scientifique de la presse ancienne comme ressource	57
1.3.1.2. Assurer portabilité et pérennité aux données textuelles	60
1.3.1.3. Le XML-TEI	62
1.3.2. La numérisation de la presse ancienne.....	66
1.3.2.1. La numérisation en mode image.....	67
1.3.2.2. La numérisation en mode texte	71
Conclusion.....	82
CHAPITRE II.	
CONSTITUTION DU DISPOSITIF D'EXPLORATION DU <i>PETIT COMTOIS</i>.....	85
2.1. Du support papier à la base textuelle.....	86
2.1.1. Brève présentation du projet de base textuelle PQR3R.....	86
2.1.2. Acquisition des données image.....	88
2.1.3. Techniques d'aide à l'OCR	96
2.1.4. Océrisation et correction automatique, semi-automatique et manuelle	100
2.1.4.1. Ouverture et traitement des images.....	101
2.1.4.2. Enregistrement du texte reconnu	102
2.1.4.3. Options de reconnaissance optique des caractères	102
2.1.4.4. Qualité de l'océrisation effectuée	102
2.1.4.5. Vérification et correction orthographique.....	105

2.1.4.6. La correction manuelle.....	108
2.1.5. Documenter les données	108
2.1.6. Établissement et normalisation des données textuelles : vers une philologie numérique	113
2.1.6.1. XML-TEI et philologie numérique.....	114
2.1.6.2. Encodage des données du <i>Petit Comtois</i> en XML-TEI.....	117
2.1.7. Bilan et perspectives	124
2.2. De la base textuelle au corpus	130
2.2.1. Retour sur notre projet de recherche	131
2.2.1.1. Bref historique des débuts de la Troisième République (1883-1903)	132
2.2.1.2. Un contexte de mutation de la pratique journalistique	140
2.2.2. Présentation du <i>Petit Comtois</i>	145
2.2.2.1. Contexte général de parution	145
2.2.2.2. <i>Le Petit Comtois</i> : un exemple typique de journal régional sous la Troisième République ?.....	148
2.2.3. Description du corpus	159
Conclusions partielles.....	173
 CHAPITRE III.	
LA MISE EN FORME DE L'INFORMATION	
DANS LE PETIT COMTOIS	175
 3.1. Description matérielle de l'objet imprimé	176
3.2. Le rubriquage du <i>Petit Comtois</i> (1883-1903).....	185
3.2.1. Approche générale du rubriquage	185
3.2.1.1. Retour sur l'étymologie du mot « rubrique ».....	185
3.2.1.2. Le point de vue des théoriciens sur la rubrique	187
3.2.2. Approche quantitative des rubriques du <i>Petit Comtois</i>	189
3.2.2.1. Méthodologie d'analyse quantitative des rubriques	189
3.2.2.2. La « cuisine » du <i>Petit Comtois</i>	196
3.2.3. Des critères sous-jacents au système de rubriquage qui se révèlent composites	211
3.2.4. Remarques conclusives sur le rubriquage du <i>Petit Comtois</i>	214
3.3. Formes et fonctions des titres.....	218
3.3.1. Approche générale du titre	218
3.3.1.1. Retour sur les fonctions du titre	219
3.3.1.2. Lectures du titre	220
3.3.2. Les titres dans <i>Le Petit Comtois</i>	223

3.3.3. Analyse des titres de la rubrique des « <i>Dépêches de nuit</i> »	229
3.3.3.1. Méthode de classification des titres.....	230
3.3.3.2. Aperçu sur les caractéristiques formelles générales des titres en diachronie.....	245
3.3.3.3. Répartition des fonctions des titres dans le sous-corpus.....	247
3.3.3.4. Remarques conclusives sur l'appareil titulaire des <i>Dépêches de nuit</i>	297
Conclusions sur la mise en forme de l'information dans <i>Le Petit Comtois</i>	299

[TOME 2]

CHAPITRE IV. EXPLORATIONS DU VOCABULAIRE DU <i>PETIT COMTOIS</i>	304
4.1. Retour sur la pratique d'analyse statistique textuelle.....	305
4.1.1. Lexicométrie, statistique linguistique, analyse des données textuelles.....	305
4.1.1.1. Une pratique interprétative objectivée du discours.....	306
4.1.1.2. Esquisse d'une praxis lexicométrique	308
4.1.2. Vers une herméneutique numérique	313
4.1.2.1. Une approche statistique des unités linguistiques plurielles du texte	314
4.1.2.2. Une approche statistique syntagmatique	317
4.1.2.3. Une statistique textuelle aux prises avec le sémantique	318
4.1.3. Les outils d'analyse textuelle assistée par informatique	320
4.1.3.1. Les outils disponibles	320
4.1.3.2. Astartex.....	324
4.2. Manipulations pré-exploratoires sur le vocabulaire du <i>Petit Comtois</i>	333
4.2.1. Consultation de l'index hiérarchique	334
4.2.2. Examen de la distribution du vocabulaire du <i>Petit Comtois</i> en diachronie.....	337
4.2.2.1. Analyse de la distribution du vocabulaire du <i>Petit Comtois</i> en diachronie.....	337
4.2.2.2. Examen du vocabulaire propre des séquences chronologiques.....	345

4.2.3. Remarques conclusives sur l'examen de la ventilation du vocabulaire en diachronie.....	378
4.3. Analyse de la micro-distribution du vocabulaire du <i>Petit Comtois</i>	379
4.3.1. Description de la configuration fine du vocabulaire du corpus	380
4.3.2. Synthèse sur la configuration fine du vocabulaire	403
4.4. Analyse de la structure fine du vocabulaire en diachronie.....	404
4.4.1. Premières expérimentations.....	404
4.4.2. Comparaison simultanée	410
4.4.3. Remarques conclusives provisoires	428
4.5. Le rubriquage, facteur structurant de la configuration fine du vocabulaire.....	429
4.5.1. Méthodologie d'analyse des dominantes lexico-thématiques des rubriques.....	429
4.5.2. Validation de la corrélation entre le système de rubriquage et la configuration fine du vocabulaire	432
Conclusion.....	444

CHAPITRE V.

APPROCHES DU NIVEAU MORPHOSYNTAXIQUE SUR LE CORPUS *PETIT COMTOIS*.

ESSAI DE TYPOLOGISATION TEXTUELLE..... 448

5.1. Bref cadrage de la pratique de typologisation textuelle à partir de variables morphosyntaxiques..... 450

5.1.1. Méthodes inductives et analyses statistiques des données 450

5.1.2. Genres et typologies textuelles..... 453

5.2. Méthodologie générale 454

5.2.1. Recueil et sélection des variables 455

5.2.2. Remarques sur l'étiquetage effectué par *Cordial*..... 457

5.2.3. Méthodologie 472

5.3. Analyse de la distribution des catégories grammaticales dans les rubriques du *Petit Comtois*

475

5.4. Analyse de la distribution d'un panel large de variables morphosyntaxiques dans les rubriques du *Petit Comtois*

482

5.4.1. Analyse de la distribution de 52 variables morphosyntaxiques dans les rubriques du <i>Petit Comtois</i>	483
5.4.2. Description des facteurs fortement discriminants	486
5.4.3. Synthèse sur les facteurs discriminants	511
5.4.4. Caractérisation morphosyntaxique des rubriques	513
5.4.4.1. Les dépêches	513
5.4.4.2. Les chroniques régionales.....	520
5.4.4.3. Les rubriques <i>Petite chronique, Chronique, Chronique de la semaine, Lettres hebdomadaires</i>	528
5.4.4.4. Les rubriques <i>Faits Divers, Tribunaux, Variétés, Causerie</i>	532
5.4.4.5. <i>Courrier de la mode, Chronique parisienne, Revue agricole, Agriculture, Bibliographie, Bulletin</i>	537
Conclusion	543
CONCLUSION GENERALE	545
BIBLIOGRAPHIE	I
INDEX	TOME 3
ANNEXES	TOME 3
TABLE DES FIGURES	TOME 3

UNIVERSITÉ DE FRANCHE-COMTÉ

ÉCOLE DOCTORALE « LANGAGES, ESPACES, TEMPS, SOCIÉTÉ »

Thèse en vue de l'obtention du titre de docteur en

SCIENCES DU LANGAGE

EXPLORATION TEXTUELLE DU DISCOURS D'UN
QUOTIDIEN RÉGIONAL AU CARREFOUR DU XIX^e ET DU
XX^e SIÈCLES : *LE PETIT COMTOIS* (1883-1903)

Présentée et soutenue publiquement par

Virginie LETHIER

Le 3 décembre 2009

Sous la direction de MM. Les Professeurs Jean-Marie VIPREY et Philippe
SCHEPENS

Membres du Jury :

Sonia BRANCA-ROSOFF, Professeure à l'Université Paris III, Sorbonne nouvelle.
Rapporteur.

Jacques GUILHAUMOU, Directeur de recherche au CNRS, UMR « Triangles »,
ENS-LSH, Lyon. Rapporteur.

Pierre FIALA, Maître de conférences à l'Université Paris 12.

Michael PALMER, Professeur à l'Université Paris 3.

Alain RABATEL, Professeur à l'Université Claude-Bernard, Lyon 1.

Philippe SCHEPENS, Professeur à l'Université de Franche-Comté.

Jean-Marie VIPREY, Professeur à l'Université de Franche-Comté.

Tome 2

Chapitre IV.

Explorations du vocabulaire du *Petit Comtois*

Le présent chapitre sera consacré à une série d'explorations dans le vocabulaire du *Petit Comtois*. Ces propositions, d'abord méthodologiques, s'inscriront dans le cadre d'une exploration textuelle du discours assistée par informatique, où les outils statistiques seront convoqués pour leur valeur heuristique. Il s'agira pour nous de dessiner une pratique de lecture du discours objectivée par le texte lui-même, dont les axes d'analyse seront déterminés au fil des saillances pointées par les manipulations sur la structure du vocabulaire du corpus.

Après avoir procédé à un bref retour sur l'état de l'art des pratiques d'analyse textuelle assistée par les outils informatiques et statistiques, visant à rappeler les enjeux et les chantiers de l'analyse textuelle des données, nous observerons l'index hiérarchique du vocabulaire de l'ensemble du corpus avant de nous tourner vers un examen de la ventilation du vocabulaire en diachronie.

Dans un second mouvement, notre approche se focalisera sur l'organisation fine du vocabulaire, à partir d'une analyse micro-distributionnelle. Guidée par les résultats issus des manipulations précédentes, nous observerons le comportement co-occurentiel des formes et des réseaux de formes dans une optique diachronique, pour enfin envisager dans quelle mesure le rubriquage participe à la configuration fine du vocabulaire observée.

Insistons dès à présent sur le statut « exploratoire » de ces procédures. Nous en attendons la possibilité de construire un trajet de lecture objectivé, où

émergent du corpus ses propres saillances linguistiques, textuelles et discursives. Celles-ci, dans l'espace de ce chapitre, ne pourront cependant pas donner lieu à une analyse exhaustive. Il s'agit pour nous de mettre en relief l'apport heuristique d'une démarche s'appuyant sur les outils de la statistique textuelle.

4.1. Retour sur la pratique d'analyse statistique textuelle

Pour la clarté de l'exposition, nous distinguerons deux grandes pratiques de quantification des faits langagiers, telles que déployées d'une part dans le cadre de la lexicométrie, et d'autre part, telles qu'elles se réalisent dans le cadre d'une nouvelle herméneutique numérique. Bien que l'évolution de la méthodologie de quantification des faits langagiers s'inscrive immanquablement dans la temporalité, ce *distinguo* repose moins sur un découpage chronologique entre ce qui pourrait être entendu comme une lexicométrie de première génération vs une lexicométrie de deuxième génération, qu'il ne vise à insister sur un processus d'adaptation des outils méthodologiques confrontés au texte par un effort de théorisation de cet objet.

4.1.1. Lexicométrie, statistique linguistique, analyse des données textuelles

Les méthodes de quantification des données textuelles ont été initiées en France dans les années 1950-60 sous des dénominations variées, traduisant des nuances importantes, notamment quant à leurs objets. Pionnières, la statistique linguistique (Guiraud, 1954) et la statistique lexicale (Muller, 1964) s'inscrivent dans le champ de la stylistique. Dans le champ de l'analyse du discours se côtoient alors l'Analyse Automatique du Discours (Pêcheux, 1969), l'analyse statistique des données textuelles (Benzécri, 1973) et la lexicométrie (Tournier, 1975), qui ont fait plus récemment place à la « textométrie » (Salem,

1987) et à la « logométrie » (Mayaffre, 2000). Malgré des divergences d'objets, d'entrées, de gestion du contexte dans la praxis d'analyse et d'algorithmes, le développement de ces méthodes de quantifications des données révèle un besoin fondamental des sciences humaines confrontées au texte : celui d'une modification profonde du geste de lecture traditionnel.

4.1.1.1. Une pratique interprétative objectivée du discours

Avant même l'argument du volume des données textuelles à traiter, c'est la nécessité d'une approche empirique et objectivée du discours qui est à l'origine des développements des méthodes lexicométriques.

Des travaux de Charles Muller à ceux de Maurice Tournier s'affirme en effet une volonté de briser le carcan d'un processus de lecture traditionnelle dominée par la subjectivité et des jugements de savoir *a priori* à partir desquels est reconstruit le sens.

Renouveler l'approche herméneutique du texte en l'abordant dans une perspective différentielle à partir des contrastes lexicaux observables en corpus : telle est la proposition sous-jacente à la mise en œuvre des outils statistiques et informatiques. La pierre angulaire d'un tel dépassement de l'acte traditionnel de lecture est l'observation fréquentielle des formes employées dans les différents textes d'un corpus. Le postulat sous-tendant le projet lexicométrique est qu'à travers l'observation des formes linguistiques mises en avant ou en retrait se joue l'accès à la thématique d'un texte, à sa structure stylistique ainsi qu'aux stratégies énonciatives et idéologiques qui le structurent :

Un texte est un tout solidaire en système d'oppositions avec d'autres, l'emploi d'un terme ne fonctionne pas isolément, mais à l'intérieur d'un vocabulaire, dans un univers de rapports aux autres emplois et aux autres termes. Dans la mesure de ces rapports, le dénombré devient significatif (Bonnafous & Tournier, 1995 : 69-70).

Une lecture objectivée doit pouvoir se targuer de reposer sur des observations quantitatives dont la qualité minimale est l'exhaustivité. Par leur rigueur, par leur

procédure systématique, les outils informatiques satisfont à cette nécessité. Ils garantissent ainsi au chercheur des relevés d'informations sûrs dont la précision et l'exhaustivité sont difficilement accessibles par le sujet humain, voire impossibles à constituer en fonction du volume de données à traiter. Auxiliaires du geste de lecture, les outils informatiques et statistiques ne procèdent pas d'une *lecture* du texte en ceci que cet acte implique un mouvement de compréhension et d'interprétation. Les outils lexicométriques procèdent à ce qui pourrait être entendu comme une *traversée* du texte dans l'épaisseur de sa structure paradigmatique. Les outils informatiques balayent le texte, l'indexent et procèdent à des relevés, que l'analyste se chargera d'abord d'observer et de décrire préalablement à tout geste interprétatif. On trouve notamment trace de l'importance accordée aux phases d'observation et de description dans la Revue *Mots*¹, lieu privilégié de réflexion sur le traitement lexicométrique. Dans le premier numéro de cette revue, sous la plume de Maurice Tournier (1980), sont ainsi édictées les consignes d'un nouveau geste de lecture. En réaction aux méthodes subjectivistes, M. Tournier prescrit une étape de description portant aussi bien sur les observations résultant des décomptes effectués que sur les outils méthodologiques utilisés, à l'égard desquels est lancé un appel à précaution :

Dernière règle, non la moindre : décrire avant de juger, en suspectant même les outils de la description et les modèles interprétatifs qui les accompagnent. Utiliser les saisies les plus neutres possible, les armes les mieux refroidies ou les reculs les plus arbitraires ne suffit pas [...] Aucune innocence n'est plus calculatrice que celle d'un ordinateur (Tournier, 1980b : 7).

L'ambition de Maurice Tournier est claire et radicale: substituer une pratique de lecture contrôlée et objectivée à une pratique de lecture reposant sur un entassement instable de connaissances *a priori* et intuitives, étayées par des arguments d'autorité. La rigueur et l'exhaustivité garanties par les outils informatiques ne sauraient suffire à accéder à une lecture contrôlée. Pour qu'elle soit objectivée, la lecture doit également être réflexive à l'égard des

¹ Acronyme signifiant dans les années 1980 « Mots-Ordinateurs-Textes-Sociétés »

outils et des algorithmes employés, qui, si formalisés qu'ils soient, n'en sont pas moins surdéterminés par des options épistémologiques, méthodologiques et herméneutiques. Il s'agit donc de procéder à une description exhaustive tant des résultats issus du traitement lexicométrique que du traitement lexicométrique lui-même. À cette étape succède une lecture fondamentalement syntagmatique et contextualisante dans un mouvement de retour au texte, qui s'attache à éclairer les quantifications relevées.

Dans le cadre d'une démarche hypothético-déductive, l'analyse assistée par informatique, convoquée pour confirmer ou infirmer des hypothèses préétablies, se révèle un puissant outil probatoire. Cependant, selon nous, l'outil informatique n'acquiert sa pleine dimension que lorsqu'il est sollicité pour son pouvoir suggestif. Comme le souligne J. Peytard, l'apport essentiel de l'outil informatique relève du fait qu'il élargit « les capacités heuristiques, sans que le chercheur ait à se déprendre de sa maîtrise de l'analyse » (Peytard, 1992 : 179). Lorsque le texte est abordé sans hypothèses préconstruites, sans questionnements prédéfinis, les outils informatiques et statistiques dessinent en effet des parcours de lecture objectivés par le texte lui-même. Les observations quantifiées, loin d'être abordées comme des résultats à proprement parler, sont alors destinées à devenir des pistes pour la poursuite de l'analyse.

4.1.1.2. Esquisse d'une praxis lexicométrique

Si, jusqu'à présent, nous nous sommes concentrée sur le statut des outils méthodologiques, nous souhaitons procéder à leur description de manière aussi brève que possible, dans la perspective plus large de proposer un regard sur la pratique qui les mobilise. Aussi ne s'agira-t-il pas de faire état ici des principes statistiques lexicométriques ni des débats qui s'y rattachent², mais de les mettre en perspective avec une certaine praxis. S'il est abusif de considérer la lexicométrie comme une méthodologie unifiée, un certain nombre

² Une abondante et riche littérature est déjà disponible. Parmi les ouvrages de référence : Benzécri (1981), Lebart & Salem (1994).

de principes régulateurs en sont néanmoins constitutifs. Dans une perspective volontairement schématique, on peut ainsi considérer que la pratique lexicométrique s'articule autour des étapes suivantes, identifiées par un chercheur qui s'est attaché, pour la première fois, à une comparaison et mise en perspectives des principaux logiciels du domaine :

Une démarche classique en lexicométrie consisterait à examiner les caractéristiques des différentes partitions, afin de prendre connaissance du corpus et de vérifier sa pertinence statistique, l'index hiérarchique (sur l'ensemble puis pour chaque partie) puis à observer les grandes oppositions au moyen d'une analyse factorielle ou de tout autre outil de statistique multidimensionnelle. Le second mouvement consisterait à examiner les spécificités par partie, c'est-à-dire à s'interroger en termes de vocabulaire commun (partagé), vocabulaire spécifique, vocabulaire original. L'ordre de ces étapes étant susceptible de variations selon les points de vue que l'on désire apporter. Une variante de cette démarche pourrait inclure des préoccupations relatives à la richesse, à l'accroissement du vocabulaire soit dans une perspective stylométrique, soit comme chez Labbé et Monière pour mettre en évidence des ruptures, une modification de vocabulaire pouvant révéler des revirements thématiques (Leblanc, 2005 : 56).

Une fois le corpus constitué, ce dernier est soumis à une procédure visant à segmenter la chaîne textuelle en occurrences et à regrouper ces données en un tableau raisonné d'occurrences identiques. La consultation de cet index du corpus et de ses partitions esquisse une première prise de vue sur le texte et permet la formulation de premières hypothèses. On renverra par exemple à l'étude de Jean-Marie Cotteret et René Moreau (1969) sur le vocabulaire du général De Gaulle : à partir du décompte des déictiques dans les discours étudiés, l'étude aboutit à une typologie entre discours-appels et discours-bilans. Dans un second mouvement, l'analyste se tourne vers un ensemble d'outils statistiques afin d'obtenir des informations sur la ventilation des formes graphiques brutes dans les structures externes du corpus (diachronie, source, chapitre, etc.). Outil de mesure d'ordre descriptif, l'Analyse factorielle des correspondances (AFC) est alors convoquée pour sa puissance de représentation graphique des phénomènes statistiques complexes :

L'analyse factorielle traite les tableaux de nombres et elle remplace un tableau difficile à lire par un tableau plus simple à lire qui soit une bonne approximation de celui-ci (Cibois, 1983 : 5).

Plus précisément, l'analyse factorielle des correspondances pratiquée dans le cadre d'une analyse du discours traditionnelle offre la visualisation d'une collection de renseignements numériques relatifs à des unités lexicales, collection contenue dans des tableaux à double entrée ou *tableaux de contingence*. Dans ces tableaux, peuvent se croiser des vocables en ligne et des modalités (variables) de partitionnement externe en colonnes (année, auteur, titre, rubrique, etc). L'Analyse factorielle des correspondances (AFC) offre ainsi une représentation synthétique de la distribution du vocabulaire où chaque entaille dans l'espace graphique signale une zone « où ça signifie » (Peytard, 1999 : 357). En ce sens, elle constitue, comme le rappelle Jean-Pierre Massonie (1990), un outil profondément heuristique :

[...] la statistique devient donc l'art et la manière de poser des questions nouvelles à l'endroit qu'il convient dans le texte...Non de résoudre des problèmes, mais de les faire naître par une lecture nouvelle, en indiquant où on doit relire (Massonie, 1990a : 102).

Science de l'écart par excellence, la statistique sert ainsi une lecture des lieux de variance du texte et de ses ruptures. En cela, la lexicométrie peut être perçue comme une lecture *différentielle* du texte (Peytard, 1992). Dans le cadre de cette démarche, l'examen des spécificités³ permet ensuite de prendre une connaissance complémentaire et plus fine vocabulaire spécifique de chacune des partitions du corpus. On distingue classiquement deux modes de calcul, l'un privilégié par Saint-Cloud (les spécificités), l'autre par le courant des linguistes comme P. Guiraud, C. Muller (l'écart-réduit). Les premières ont pour référence la distribution aléatoire, le second s'établit par référence à l'équidistribution. Leurs résultats convergent fortement. De par l'emploi du logiciel *Astartex*, nous avons opté pour l'écart-réduit. À l'aune de la norme endogène du vocabulaire du corpus, les sur-emplois et les sous-emplois d'une

³ L'analyse des spécificités est une méthode conçue par les linguistes de Saint-Cloud (notamment : Lafon, 1980).

forme dans une partie du corpus sont les outils de mesure de la structure du vocabulaire.

Les apports des outils lexicométriques sont riches : le vocabulaire hiérarchique permet d'avoir une vue d'ensemble des thématiques d'un corpus et de ses principales caractéristiques énonciatives, voire stylistiques ; l'AFC esquisse les oppositions principales exercées sur le vocabulaire du corpus ; le calcul des spécificités permet quant à lui d'expliquer cette configuration du vocabulaire par la mise en lumière du profil distributionnel des formes qui la structurent. La « boîte à outils » lexicométrique permet donc de pointer les formes et les partitions du corpus en tant que saillances significatives, mais n'intègre pas l'accès au sens. Celui-ci est relégué au terme du traitement lexicométrique :

On ne peut, au départ d'une étude lexicométrique, que cumuler, confronter, réunir ou opposer des occurrences de formes textuelles. Les questions de sens sont à poser ultérieurement, par un travail sur le dépouillement ou les résultats, qui sera spécifique mais dont la statistique ne garantira plus forcément la valeur (Bonnafous & Tournier, 1995 : 69).

En d'autres termes, la lexicométrie se caractérise par une distanciation méthodologique temporaire entre les formes lexicales et le sens. L'accès au sens s'effectue en effet par un retour au contexte (manuel ou assisté), qui échappe aux premières étapes de l'analyse linguistique. La consultation de l'index hiérarchique se résume à une prise de vue sur une liste de formes atomisées, dont le sens résiste par essence dès lors qu'elles sont extirpées de leur co(n)texte. L'examen de la distribution du vocabulaire dans ses structures exogènes, s'il intègre la macro-structure d'un texte ou d'un recueil de textes, se prête à la même critique. Tous ces outils fournissent uniquement des observations quantifiées, qui, une fois décrites exhaustivement par l'analyste, exigent d'être mises en perspective prioritairement avec le contexte linguistique, puis avec le contexte, en tant qu'environnement extralinguistique. Le retour au (con)texte s'impose ainsi en garde-fou méthodologique obviant toute interprétation hâtive :

Le travail sur contexte doit être systématique : on ne connaît jamais son corpus au point de ne pas risquer de gauchir les faits à la lumière d'une interprétation qui semble ingénieuse (Habert, 1985 : 143).

Le geste final d'interprétation impliquera cependant de recourir à un contexte non enclos dans le corpus, n'ayant pas fait l'objet de la même objectivation que les données jusqu'au préalable convoquées. C'est tout le paradoxe du geste lexicométrique, au terme duquel le chercheur « abandonne au moment de l'interprétation le parti initial de l'exhaustivité des données et des traitements » (Fiala, 1994 : 120).

C'est cette rupture méthodologique sur laquelle nous voulons mettre l'accent par le schéma présenté ci-dessous. Si l'analyse lexicométrique permet d'objectiver le *pluralisme des interprétations possibles* (Duchastel & Daoust, 2007 : 257-268), elle ne peut faire l'économie d'une pluralité de choix herméneutiques, s'imposant depuis la constitution du corpus et le prétraitement des données jusqu'aux étapes de l'analyse lexicométrique proprement dite. D'où une représentation du corpus et de l'analyse lexicométrique cerclée de vert, visant à insister sur leur surdétermination par un ensemble de choix heuristiques. Si nous avons déjà largement évoqué les choix présidant à la constitution d'un corpus, il nous reste à constater, à la suite de B. Habert (1985) et de P. Fiala (1994), que le traitement lexicométrique est également le lieu de choix herméneutiques. Depuis le choix de la segmentation opérée et de la partition convoquée jusqu'à ceux des formes lexicales prises en considération, des seuils retenus et de la méthode sélectionnée pour évaluer la surreprésentation des formes dans une partition, l'analyse lexicométrique est caractérisée par une série d'options elles-mêmes surdéterminées par des considérations épistémologiques, méthodologiques et théoriques.

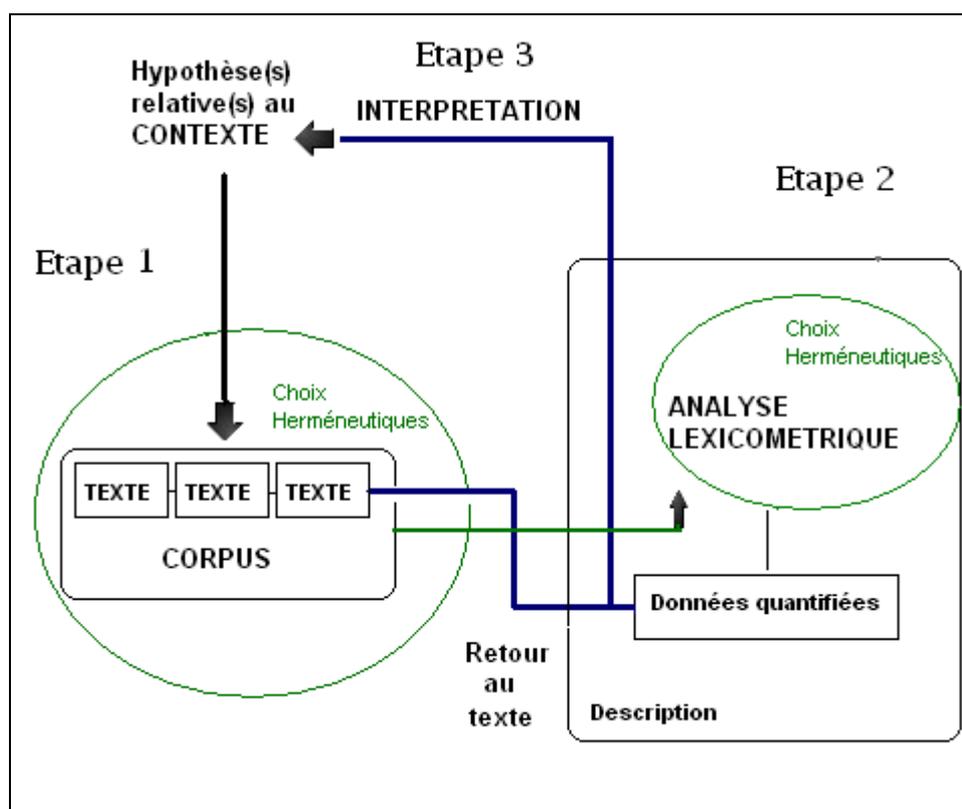


Figure IV-1 : Le geste d'analyse lexicométrique.

4.1.2. Vers une herméneutique numérique

Dès la fin des années 1980, on rencontre dans la littérature consacrée à l'analyse des données textuelles les pistes d'un changement de paradigme du champ, qui ne se réalise néanmoins pleinement qu'à la fin des années 1990. Alimentée par les progrès de capacité des logiciels et son corollaire, la rapidité des traitements de grands ensembles textuels, l'analyse textuelle des données en tant que « logométrie » (Mayaffre, 2000) ou « textométrie » est portée par une ouverture aux questions herméneutiques et à un effort de définition de l'objet *texte*.

Nous distinguerons ici trois principaux chantiers, étroitement liés, autour desquels se joue ce déplacement vers une démarche *herméneutique et-ou philologique*. Il conviendra ainsi tout d'abord de faire état de la prise en compte

de nouvelles unités pour une approche statistique libérée de l'entrée de l'unité graphique ; ensuite, nous traiterons du dépassement de la seule dimension paradigmatisée du texte par une approche statistique désormais syntagmatique. Enfin, nous envisagerons le tournant herméneutique à l'œuvre *via* l'initiation d'une statistique de la co-occurrence.

4.1.2.1. Une approche statistique des unités linguistiques plurielles du texte

Si l'approche statistique du texte était traditionnellement une analyse de données lexicales brutes, elle s'est aujourd'hui enrichie d'une approche des unités linguistiques, sémantiques, grammaticales et syntaxiques tissant le texte. Cette évolution, résultant principalement des progrès technologiques réalisés dans la dernière décennie du XX^e siècle, marque l'évolution d'une lexicométrie à une *logométrie* :

Tant et si bien que le monde étroit et contestable de la lexicométrie de la première génération semble dépassé pour s'ouvrir sur l'univers nouveau d'une logo-matique pleine et entière qui embrasse le texte, qualitativement et quantitativement, dans toutes ses dimensions linguistiques, lexicale, grammaticale, syntaxique, sémantique, rhétorique (Mayaffre, 2002b).

Une évolution des plus importantes de l'analyse textuelle des données est tout d'abord relative au dépassement de la forme brute comme unique entrée du texte. Prenant pour unité le « mot » en tant que segment graphique isolé par des blancs selon une conception issue de la tradition typographique, la lexicométrie s'est longtemps vu reprocher de ne procéder qu'à une description superficielle du texte, fondée sur sa matérialité graphique. L'infirmité de la pertinence d'une analyse lexicométrique menée sur des données lexicales brutes reposait sur une série d'objections relatives tout d'abord à l'ambiguïté constitutive de la langue. Alors que le taux de formes ambiguës dans un discours en langue française est estimé à 15%, un décompte des formes graphiques présente le risque potentiel de réunir sous la même entrée deux unités de sens distinctes. Le cas des lexies composées, décomptées en unités

graphiques séparées ne renvoyant individuellement à aucune unité sémantique, confortait la position selon laquelle une analyse sur des données lexicales brutes négligeait le poids sémantique du discours. La fiabilité des relevés lexicométriques était de surcroît questionnée en raison de la dispersion graphique variable des unités de langue : la déclinaison d'un verbe en un éventail plus large de formes qu'un nom commun était ainsi mise en avant pour mettre en cause le socle méthodologique de la lexicométrie. On retrouvera ces arguments plus détaillés ainsi que les réponses des tenants de la surface graphique dans une abondante littérature⁴, alimentant le débat entre « formalistes », attachés à l'étude de données brutes, et partisans de la lemmatisation, militant pour la substitution d'une unité linguistique pertinente à l'unité graphique traditionnelle. La lemmatisation peut être définie comme une opération visant à regrouper les unités graphiques correspondant aux différentes flexions d'une même unité de lexique, sous une « forme canonique ». Dans une phase de regroupement paradigmatique, on procède à la réunion des formes hétérographes et à la séparation des formes homographes. Les formes verbales sont ainsi ramenées à l'infinitif, les substantifs au singulier, les formes élidées transformées en formes sans élision, les pronoms et les adjectifs pronominaux présentant une flexion en genre et en nombre regroupés. Le développement d'une analyse quantitative de données textuelles lemmatisées a cependant été lent, moins en raison des arguments des formalistes relatifs à la perte sémantique inhérente à la pratique de lemmatisation que par la nature chronophage des opérations manuelles induites et l'absence de lemmatiseurs automatiques opérants et fiables sur le plan linguistique.

Depuis le début des années 2000, on constate le développement de logiciels automatiques d'annotation perfectionnés, assurant la lemmatisation mais aussi un certain degré de reconnaissance linguistique, tels que *Cordia*⁵

⁴ Notamment Muller, C. (1993) et Lafon, P. (1984).

⁵ Société Synapse-Développement.

ou *TreeTagger*⁶, et d'autre part, l'implémentation de traitement de données annotées dans certains outils lexicométriques. Un logiciel comme *Hyperbase* traite ainsi directement les sorties des logiciels *Cordial* et *TreeTagger*, permettant un traitement statistique des diverses unités linguistiques du texte (forme graphique, lemme, codes grammaticaux). Ces logiciels de lemmatisation et d'annotation grammaticale garantissent la rapidité du prétraitement tandis que leurs résultats sont jugés fiables, comme en témoigne D. Mayaffre :

Aujourd'hui les lemmatiseurs-étiqueteurs affichent, dans leurs fonctions de base (désambiguïsation des homographes, regroupement flexionnel des noms ou des verbes, reconnaissance des catégories grammaticales élémentaires), un taux de réussite proche de 100 %: pour ces fonctions de bases quelques erreurs subsistent mais aucune susceptible de fausser le traitement statistique que nous opérons par la suite; les vérifications manuelles restent utiles mais non plus obligatoires (Mayaffre, 2005a).

Nous émettons cependant de grandes réserves par rapport à la fiabilité d'une lemmatisation automatique et il nous faut nuancer le jugement de D. Mayaffre. Si les logiciels automatiques d'annotation garantissent la systématisme du prétraitement des données, l'utilisation d'un système d'étiquettes cohérent, ils sont intrinsèquement *d'imparfaits instruments* (Habert, 2005 : 37) qui nécessitent immanquablement une procédure de vérification manuelle. Remarquons que d'après B. Habert (2005 : 40), le taux de réussite de ces logiciels oscille entre 95 et 97% de mots correctement annotés pour le français écrit standard. Nous concédons que ces 3% d'erreurs ne remettent pas en cause la validité des résultats d'une analyse quantitative de première intention, puisque la méthodologie statistique sur laquelle elle repose vise à dégager les grandes tendances d'une série de textes, d'où une certaine tolérance vis-à-vis de ce pourcentage d'erreur. Il nous semble intéressant de mettre en perspective la fiabilité imparfaite de la lemmatisation opérée avec les constats issus de certaines études comparatives : celles-ci ont démontré que les résultats d'une analyse statistique basée sur un schéma d'urne, sur des

⁶ Développé au sein du projet TC (Textcorpora and Erschliessungswerkzeuge) à l'institut de linguistique computationnelle de l'université de Stuttgart : <http://www.ims.uni-stuttgart.de/projekte/tc/>.

données brutes et lemmatisées ne présentaient aucune différence substantielle. É. Brunet (2000) montre en effet l'insensibilité des résultats de classification de première intention (évolutions diachroniques massives, etc.), qu'elle que soit l'unité (graphique ou lemme) convoquée, ce que constate également Mayaffre (2004). Si l'intérêt de disposer de données textuelles lemmatisées pour une analyse statistique traditionnelle n'excède pas, au niveau des résultats, celui des données brutes, le statut même de la lemmatisation exige d'être radicalement revu dans une perspective de philologie numérique, où la lemmatisation est entendue comme un véritable état du texte. De fait, toute lemmatisation ne saurait être réalisée de façon exclusivement automatique : dans la continuité des efforts d'établissement du texte, notamment appelé par une analyse textuelle du discours, la lemmatisation doit être une opération *contrôlée*, transparente et *justifiée* (Viprey & Lethier, 2008a). À cette seule condition, la lemmatisation ne constitue plus une substitution destructive à la surface graphique, mais bien une *annotation* intégrée à l'enrichissement des données dans le cadre des protocoles XML-TEI (Viprey, 2006b: 175-178).

4.1.2.2. Une approche statistique syntagmatique

Dans la perspective lexicométrique classique, le texte est assimilé à une population, à un « sac de mots » (Rastier, 2001) par l'analyse des données textuelles. D'où la méfiance persistante dans les disciplines du texte à l'égard de cette méthodologie. Tandis qu'elles s'expliquent par une première adaptation des méthodes de statistique générale, les insuffisances de cette approche pointent le déficit de définition de l'objet *texte* caractéristique de l'analyse quantitative des données textuelles, elle-même inscrite dans une analyse du discours sujette à la même faiblesse (Sarfati, 2003 : 432).

L'ère du numérique a très certainement contribué à ce que nos disciplines se recentrent sur le texte et à ouvrir le débat entre la linguistique textuelle et l'Analyse des données textuelles (désormais ADT). La conférence

plénière inaugurale aux Journées d'Analyse de Données Textuelles 2006 tenue par Jean-Michel Adam (2006b) prend acte de l'ouverture du dialogue entre les deux communautés. Ce dernier souligne que, bien loin de s'opposer, linguistique textuelle et ADT sont aujourd'hui destinées à travailler conjointement. Si la définition privilégiée par la linguistique textuelle mettait essentiellement l'accent sur la dimension linéaire du texte, elle s'est aujourd'hui enrichie d'une ouverture sur sa *structure non-séquentielle* et *réticulaire* (Viprey, 2005b) pointée par les récents développements de l'ADT. De son côté, l'approche paradigmatique traditionnellement privilégiée par la statistique textuelle est désormais enrichie d'une prise en compte de la dimension linéaire du texte. Les travaux de C. Lamalle et A. Salem (2002) marquent à ce titre un tournant de l'analyse des données textuelles, que viennent confirmer les chantiers de « topologie textuelle » menés par X. Luong, D. Longrée et S. Mellet (2004, 2007, 2008). Au-delà de différences significatives, cette classe de recherches a pour but de dépasser le cadre d'une pratique d'analyse contrastive reposant sur un macro-découpage statistique du corpus, en l'enrichissant d'une approche de la séquentialité du texte. La distribution d'une forme (lexicale ou syntaxique) n'est dès lors plus envisagée indépendamment de la progression du texte, mais au contraire, au fil de ses occurrences dans le texte. Parce que cette méthodologie rend compte des enchaînements et des ruptures d'emploi d'une forme, elle porte en germe la redéfinition de nouvelles coupes dans le corpus par le chercheur, et par la même de la dynamisation du dispositif d'observation en fonction des objectifs des pistes de recherche pointées par des manipulations pré-exploratoires.

4.1.2.3. Une statistique textuelle aux prises avec le sémantique

Traditionnellement, l'analyse lexicométrique s'attache à mesurer le profil distributionnel d'une forme lexicale dans les différentes parties du corpus au travers d'analyses comparatives. Les observations ainsi objectivées sur les

occurrences du corpus sont précieuses mais décontextualisées, et donc déchargées de leur sens. Dans la pratique lexicométrique, le retour au texte devient alors ce temps essentiel de l'analyse mais disjoint du traitement statistique, au cours duquel la consultation du contexte permet à l'analyse de toucher au sens. Cette statistique de l'occurrence fait aujourd'hui place à une statistique de la co-occurrence. Cette dernière, bien que portée en germes dans les travaux pionniers de M. Demonet *et al.* (1975), M. Tournier (1980), de P. Lafon (1984), ne s'est concrétisée que depuis une dizaine d'années, s'illustrant notamment à travers les contributions récentes de J.-M. Viprey (1997), S. Heiden (1998 et 2004), W. Martinez (2003) ou E. Brunet (2006). La plupart des logiciels sont aujourd'hui implémentés de fonctions permettant une mesure des faits co-occurentiels, se distinguant néanmoins par leurs méthodes de calculs et de représentation⁷.

La co-occurrence est définie comme la rencontre de deux unités linguistiques au sein d'un contexte linguistique délimité. Elle donne lieu à une mesure de la co-présence des unités. À l'aune du rapport fréquentiel entre deux items co-présents dans le corpus au sein d'une fenêtre co-textuelle délimitée, émergent les phénomènes dynamiques d'associations et de répulsions lexicales tissant le texte. Or, c'est par ses voisinages et ses antagonismes lexicaux que se construit le sens d'un mot, comme le précise M. Tournier :

Le mot : cet acteur de sens que seuls d'autres mots peuvent actionner dans un sens. Tout, dans l'énonciation, est séquence, réseau, co-occurrences, équilibre ou fils formes, scansion d'emplois affrontées (Tournier, 1980b : 7).

Atomisé, isolé, esseulé, un mot ne fait pas *sens*. Le sens ne se fait qu'en contexte. Or, qu'est-ce que le contexte accessible pour l'analyste, si ce n'est le co(n)texte en tant que contexte linguistique, réalisé par les unités lexicales entourant l'occurrence interrogée ? La co-occurrence, en cela qu'elle interroge

⁷ Nous reviendrons plus en détails sur ces caractéristiques. À titre d'exemple concernant les modes de visualisation, *Weblex* propose notamment une représentation ordonnée de lexicogrammes simples, et une représentation des réseaux d'affinités par les lexicogrammes récursifs tandis qu'*Hyperbase* privilégie une représentation synthétique.

l'environnement d'une unité linguistique dans une fenêtre de texte donnée, constitue donc un premier palier d'accès au contexte, et par suite, un premier palier d'accès pour l'analyse statistique à la sémantique. Par conséquent, la co-occurrence peut être envisagée, selon l'expression de D. Mayaffre, comme *l'unité textuelle élémentaire* à partir de laquelle est permise l'interprétation :

Nous considérerons en effet, en corpus, que la forme minimale du contexte d'un terme, nécessaire à sa compréhension-interprétation, n'est pas le syntagme ou la phrase mais la co-occurrence ; ou, dit autrement encore, nous définirons ici la co-occurrence comme la forme minimale du contexte qui présente l'avantage de se trouver accessible de manière systématique, étant entendu que nous ne saurions considérer, même avec un concordancier, un par un, tous les mots dans toutes leurs chaînes (Mayaffre, 2008a : 95).

Dans un mouvement d'intégration du co(n)texte, l'analyse textuelle des données s'ouvre à la sémantique et met à mal le *distinguo* établi entre analyse quantitative et analyse qualitative. L'analyse textuelle des données intégrant quantitativement le contexte pour accéder à la sémantique nécessairement qualitative ouvre, on peut en tout cas le souhaiter, le passage *d'une linguistique logico-grammaticale* à une *logique réthorico-herméneutique* (Rastier, 2001). Mais plus encore, l'ADT peut pleinement revendiquer son statut de méthodologie porteuse d'un renouveau des études lexicologiques. À l'heure du numérique et des bases textuelles, celle-ci devient ce dispositif d'observation des sens d'une forme lexicale à l'aune de la diachronie ou des formations discursives.

4.1.3. Les outils d'analyse textuelle assistée par informatique

4.1.3.1. Les outils disponibles

Différents logiciels dédiés à l'analyse informatisée de corpus de textes sont aujourd'hui disponibles pour la communauté scientifique. Chacun d'entre eux repose sur des choix de programmation nécessairement soumis à l'évolution des questionnements relatifs à l'objet d'étude de l'analyse informatisée de corpus textuels. Ils renvoient, plus largement, à l'inscription

théorique et institutionnelle des concepteurs des programmes. Au-delà de la variation des méthodes statistiques et des algorithmes convoqués ainsi que des formats d'entrée, deux grandes tendances de logiciels ont pu être dégagées. Proposant une description des outils statistiques, le site *Textopol*⁸, qui est géré par P. Fiala et J.-M. Leblanc (Ceditec, Paris XII), classe ces outils selon qu'ils relèvent de la classe des logiciels lexicométriques ou de la classe des co-occurrenceurs. Cette classification n'est pas sans faire écho au *distinguo* entre lexicométrie et herméneutique numérique. Par logiciel de type lexicométrique, sont désignés les programmes qui présentent la particularité, après une segmentation initiale du texte d'après la forme graphique, d'offrir au chercheur un éventail d'outils qui permettent d'analyser un corpus donné au gré de ses partitions préalablement définies. Communément regroupé sous cette première catégorie de logiciels, *Lexico 3*⁹ créé par C. Lamalle, W. Martinez, S. Fleury, A. Salem (Syled-Cla2T, EA 2290, Université de Paris 3) offre un ensemble de fonctionnalités aptes à prendre connaissance de la distribution du vocabulaire dans les différentes partitions du corpus. De l'index hiérarchique aux calculs de spécificités par partition (mais aussi recherche de groupes de formes, graphiques de ventilation en fréquences absolues ou relatives), en passant par le repérage des segments répétés, le logiciel se présente ainsi comme une boîte à outils traditionnelle de la lexicométrie. Aussi est-il également pourvu de fonctions documentaires permettant la consultation des contextes d'occurrences d'une forme interrogée par la convocation de concordanciers. Si le geste de retour au texte est donc prévu par le programme, les capacités de navigation dans le corpus sont néanmoins limitées. Concernant les fonctions d'analyses factorielles, *Lexico 3* permet d'observer les grandes oppositions d'un corpus relatives à la distribution des unités lexicales. Récemment, de nouvelles fonctions ont été implémentées dans *Lexico 3*, qui conduisent à reconsidérer l'appartenance de ce logiciel à la classe des outils lexicométriques au sens

⁸ <http://textopol.free.fr/> (Ressources informatisées pour l'analyse du discours politique)

⁹ <http://www.cavi.univ-paris3.fr/ilpga/ilpga/tal/lexicoWWW/>

restreint du terme. En effet, ce logiciel permet désormais de procéder à des analyses de co-occurrences. De surcroît, depuis 2001, le logiciel est pourvu d'une fonction de topographie textuelle, permettant d'explorer le profil fréquentiel d'une forme¹⁰ à la lumière de la chaîne syntagmatique. Le corpus d'étude est ainsi représenté dans l'interface par autant de carrés qu'il compte de paragraphes. Ces carrés se colorent selon la présence ou l'absence de la forme dans le paragraphe qu'ils représentent, mais aussi selon sa fréquence d'utilisation dans l'empan considéré.

Hyperbase, développé par E. Brunet au sein de l'UMR 6039 « Bases, corpus, langages » (Université de Sophia-Antipolis, Nice), témoigne également d'un enrichissement de ses fonctionnalités qui illustre le changement de paradigme de l'analyse textuelle des données. Les premiers développements d'*Hyperbase* s'articulent autour des outils lexicométriques de base, comme la recherche de formes lexicales et de lemmes, le calcul des spécificités par rapport à la norme du corpus (spécificités endogènes) ou par rapport à un corpus de référence (spécificités exogènes), l'analyse factorielle et arborée du vocabulaire. Ce logiciel offre également la possibilité de travailler sur des données catégorisées par l'étiqueteur *Cordial* et de dépasser l'entrée de la forme graphique. Dernièrement, des fonctions relatives aux phénomènes de co-occurrences ont été implémentées dans *Hyperbase* : celui-ci permet désormais le calcul et la construction de graphe de co-occurrences, une analyse arborée et la construction de représentations topologiques. Le logiciel propose également la construction de graphiques de co-occurrence généralisée semblables à ceux de micro-distribution d'*Astartex*, sur lesquels nous reviendrons ultérieurement. Outil très complet permettant la mise en œuvre d'une analyse statistique à la fois paradigmatique et syntagmatique, *Hyperbase* constitue un des logiciels les plus aptes à servir une herméneutique numérique, et cela en raison de son hypertextualité. En effet, à la différence de la grande majorité des logiciels d'analyse textuelle informatisée, *Hyperbase* intègre dans

¹⁰ Ou de deux formes.

son ergonomie le mouvement de retour au texte. Les liens hypertextes concrétisent le postulat de la nécessaire contextualisation des unités lexicales. Il suffit ainsi d'un clic sur une forme présentée dans une liste ou sous une quelconque représentation graphique (histogramme, AFC, etc.) pour faire apparaître tous ses emplois dans le texte considéré.

Comptant parmi les outils qui allient fonctions lexicométriques et d'analyse cooccurrentielle, *Weblex*¹¹, conçu par Serge Heiden (ICAR, CNRS/ENS-LSH) est un outil lexicométrique accessible en ligne. Doté des fonctions documentaires et statistiques traditionnellement usitées en lexicométrie, l'outil est également implémenté de fonctionnalités de calcul des co-occurrences, à travers l'établissement de lexicogrammes simples et récursifs.

Élément saillant de la classe des co-occurrenceurs, *Coocs*¹², développé par William Martinez (Syled-Cla2T, EA 2290, Université de Paris 3), est entièrement consacré à l'analyse des phénomènes de co-occurrence. Ce logiciel est pensé pour fonctionner en aval de *Lexico 3* : *Coocs* ne segmente pas de bases textuelles, mais prend en charge celles qui ont été segmentées en amont par *Lexico 3*. Ne pouvant de fait être pensé comme un logiciel « autonome », *Coocs* est composé d'une série de modules pour l'analyse des co-occurrences lexicales. Il permet de détecter les co-occurrences lexicales autour d'une forme pôle, de repérer les co-occurrences dans une section d'un corpus partitionné, de détecter des réseaux de co-occurrence virtuelle, d'extraire des contextes de co-occurrence et enfin d'analyser la ventilation des cooccurrents autour d'un pôle. Nous insistons néanmoins sur le fait que *Coocs*, même s'il relève d'une classe d'analyse contextualisante, ne prévoit pas dans sa conception le retour au plein texte.

Pour finir ce panorama des outils statistiques disponibles, nous citerons également le logiciel *Alceste* (Analyses des Lexèmes Cooccurrents des

¹¹ <http://weblex.ens-lsh.fr/wlx/>

¹² <http://williammartinez.fr/coocs/page.php>

Énoncés Simples d'un Texte) conçu par M. Reinert, commercialisé par la société *Image*¹³. Les postulats théoriques et méthodologiques sous-tendant la conception de ce logiciel diffèrent sensiblement des autres logiciels. Sur le plan méthodologique, contrairement aux autres logiciels précédemment évoqués, *Alceste* ne procède pas à un partitionnement du texte, mais à la constitution automatisée de classes d'énoncés indépendamment des partitions du corpus. La structure externe du texte n'est donc pas prise en compte par ce programme. Le logiciel effectue un repérage des unités lexicales, identifiées au moyen d'un dictionnaire, puis procède à une lemmatisation-racinisation. Les énoncés sont alors triés en fonction de la présence ou de l'absence des formes qui les composent, puis subissent la méthode de classification descendante hiérarchique. *Alceste* obtient ainsi des classes de mots, triés selon leur coefficient d'association à la classe par la méthode du khi2, ainsi jugés les plus représentatifs des énoncés, et des classes d'énoncés. L'analyse statistique distributionnelle est convoquée pour mettre en évidence les grandes articulations du corpus dans une logique de fait inductive, selon la mise en lumière de « mondes lexicaux » (Reinert, 1993).

4.1.3.2. *Astartex*

Notre exploration du corpus *Petit Comtois* s'appuiera sur le programme *Astartex*, développé par J.-M. Viprey dans le cadre du pôle *Archive, Base, Corpus* de la MSHE Claude-Nicolas Ledoux). Plus qu'une présentation de ce programme¹⁴, il s'agit essentiellement ici de justifier notre choix de son utilisation, en insistant sur la cohérence des fonctionnalités de cet outil avec les besoins de notre projet de recherche.

Dédié à l'exploration de corpus assistée par la statistique textuelle, *Astartex* est une interface qui offre une palette de fonctionnalités parentes d'une partie de

¹³ http://www.image-zafar.com/index_alceste.htm

¹⁴ Nous renvoyons le lecteur à la page web :http://laseldi.univ-fcomte.fr/document/viprey/page_JMV.htm

celles proposées par *Hyperbase* : l'exploitation documentaire et la requête lexicale simple ou multiple en vue de la constitution de concordances et de contextes, le calcul statistique en temps réel sur les sur- et sous-emplois lexicaux des contextes convoqués et sur la distribution, l'emploi de graphes d'Analyse Factorielle des Correspondances et la constitution de dictionnaires de corpus. Il partage également avec *Hyperbase* la caractéristique d'être un dispositif hypertextuel, où sont mis en relations les divers états du texte par des liens hypertextuels. L'aspect hypertextuel du logiciel est à relier avec une certaine conception du processus d'exploration. L'ergonomie de l'interface est pensée pour que le parcours d'analyse de l'utilisateur soit en permanence un mouvement de va-et-vient entre les observations quantifiées et la démarche documentaire. L'amélioration des capacités de navigation exigées par le retour au texte est un des chantiers d'expérimentations mis en œuvre dans le cadre du développement de ce logiciel. Aussi, tous les résultats et ressources configurant le texte (tableaux, listes, nuages de points, dictionnaires) proposés par *Astartex* ont la particularité d'être organisés en hypertexte expert, dans un souci de dynamisation du dispositif d'exploration. Selon son concepteur :

Il s'agit, dans un environnement d'analyse textuelle fondé sur le primat de la lecture non-linéaire des corpus, de disposer en sus des listes, concordances et champs d'affichage plein-texte, de graphes à nuages de points où ceux-ci sont cliquables en vue de diverses fonctions de poursuite, optimisant ainsi le caractère dynamique de l'appareillage (Viprey, 2006c : 990).

D'un clic sur un item, le chercheur peut ensuite avoir accès au plein-texte, aux concordances et aux listes où apparaît l'item interrogé. Illustrons dans la mesure du possible notre propos par une série de prises de vues sur l'interface *Astartex*, bien que la présentation fige l'aspect dynamique que nous visons à décrire. Présentons tout d'abord le dictionnaire du corpus *Petit Comtois* tel qu'accessible dans *Astartex*. Les items sont classés par ordre de fréquence décroissante, par défaut :

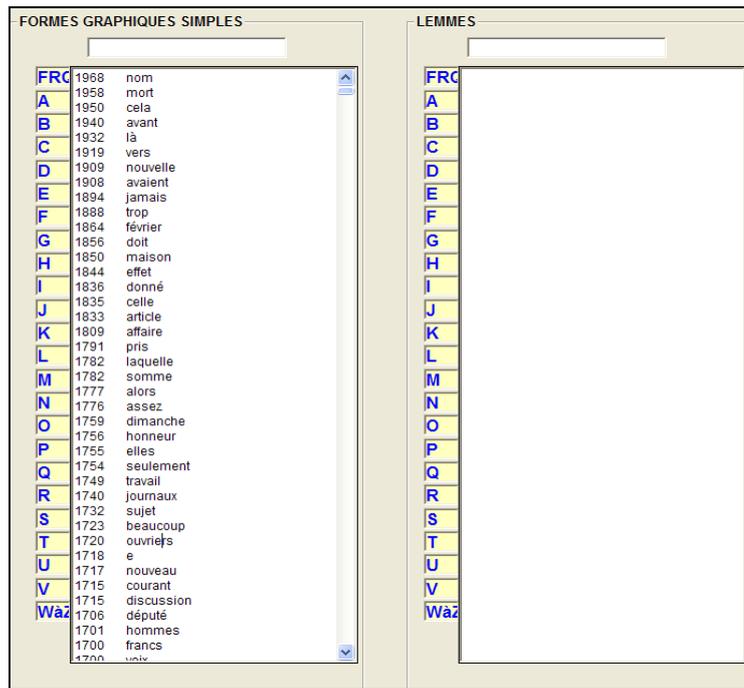


Figure IV-2 : Dictionnaire du corpus *Petit Comtois* (curseur sur « ouvriers »).

Un simple clic sur l’item de notre choix, en l’occurrence « ouvriers », permet de mettre en mémoire la sélection et de poursuivre l’analyse sans saisir au clavier l’item désiré :

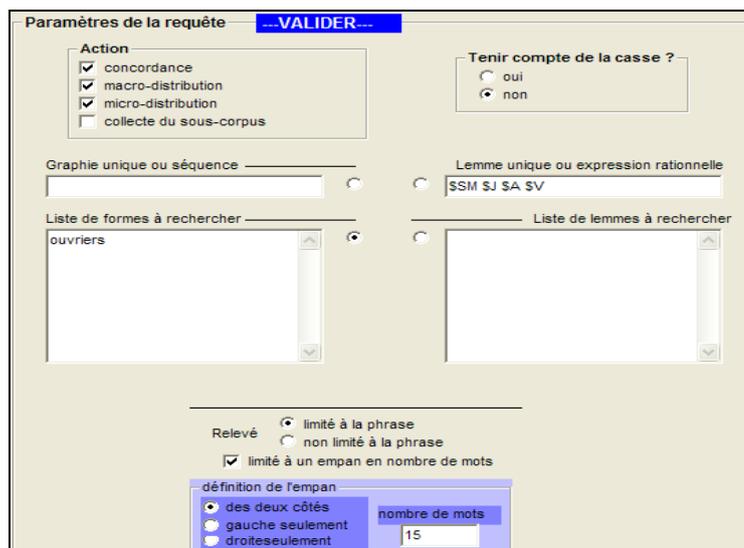


Figure IV-3 : Lancement de l'analyse.

Toujours d'un simple clic, nous obtenons une concordance, triable à gauche et à droite, permettant de prendre une première connaissance globale des contextes :

	ouvriers	, qui permettrait d' envoyer à la Chambre des députés sortant
	ouvriers	français arrivés à soixante-cinq ans et n' ayant pas de reven
	ouvriers	, et une allocation égale versée par les patrons . La retenue
	ouvriers	.
	ouvriers	en moyenne sur cinq cents travaillent dans chaque fosse .
	ouvriers	, les paysans , le peuple , qui , depuis bien des années , a
	ouvriers	sont victimes dans l' exercice de leur travail . M . Georges
	ouvriers	. Enfin , il termine en rappelant le mot d' un de ses amis dé
	ouvriers	connaissent les sentiments républicains et patriotiques de M
	ouvriers	grévistés , sans trop de rixes et sans trop de bourrades . No
	ouvriers	dans toutes les jeunes . couches de la société : action d' au
	ouvriers	mineurs , à décidé de réincorporer dans la loi les articles q
	d' ouvriers	ont pris part , avant-hier , à une réunion provoquée par les
	d' ouvriers	circulent en ordre et très calmes .
	d' ouvriers	Français . L' orateur dit qu' il
ux industriels employant au moins 70 pour 100	d' ouvriers	travaillant sur les chantiers du tramway . L' auteur du vol e
a été commis ces jours derniers au préjudice	d' ouvriers	travaillaient à la construction du tunnel , lorsqu' une grand
Une bande	d' ouvriers	italiens . Deux moines ont " été grièvement blessés . Le parq
res de Narbonne , a été attaqué par une bande	d' ouvriers	. Vous ne voyez dans les rues , pendant la semaine , que des
ation étant à peu près exclusivement composée	d' ouvriers	d' artillerie de M . le capitaine Marcy qui , en personne , a
, comprenait une bonne partie de la compagnie	d' ouvriers	, qui travaillait dans les ateliers du plateau des Chèvres ,
, un artilleur appartenant à la 4 e compagnie	d' ouvriers	militaires et cherchait à contourner le fort Griffon , quand
it échappé à un artilleur de la 4 e compagnie	d' ouvriers	donnait déjà les premiers soins . Cet homme avait tenté de se
individu auquel un soldat de la 4 e compagnie	d' ouvriers	du génie , et pendant ce temps le commerce de Saint-Claude pa
bler le vide si on n' amène pas une compagnie	d' ouvriers	mineur venu en hâte de la mine pour écouter et applaudir l'
L' assistance , composée en grande partie	d' ouvriers	charbonniers .
après-midi étaient composées en grande partie	d' ouvriers	. M . Crozet-Fourneyron . - Elle est la négation de l' admiss
vation du travail pour un nombre considérable	d' ouvriers	dans la situation la plus précaire . Sur plusieurs autres poi
Anvers , qui maintient un nombre considérable	d' ouvriers	réclamant une augmentation de salaire ni diminution d' heures
La grève , en effet , n' est pas celle	d' ouvriers	socialistes , sorte d' académie d' application ou de conféren
projet de création , à Berlin , d' une école	d' ouvriers	. Mais comme le Ilot a grossi rapidement . Nous avons vu dans
sure l' existence et le bien-être à une foule	d' ouvriers	se dirigeaient vers la zone minière pour travailler , mais en
Une vingtaine	d' ouvriers	employés dans une souffrière se rendaient , hier matin , vers
Sicile . - - Une vingtaine	d' ouvriers	italiens environ . Toutefois les habitants ont organisé des p
iroilles , où il reste encore une cinquantaine	d' ouvriers	se trouvent forcément empêchés .
jours l' exécution des travaux . Une centaine	d' ouvriers	, tant français - qu' italiens . Le patron de l' établissemen
dans un café où étaient attablés une centaine	d' ouvriers	travaillent au hangar des messageries maritimes .
in sur les quelques chantiers où une centaine	d' ouvriers	français en sortait , furent invectives et menacés par ceux -
ant le café Nicole , au moment où une dizaine	d' ouvriers	. Le travail est extrêmement divisé et , par suite , le bon o
ans de petits ateliers contenant une douzaine	d' ouvriers	

Figure IV-4 : Concordancier.

À partir de ce concordancier, un retour au texte est possible pour chaque occurrence par un clic. La répartition de l'item interrogé est renseignée par un module statistique (paramétré – partie inférieure de la fig. IV-3), présentant notamment un histogramme basé sur le calcul de l'écart-réduit à l'équidistribution (ici, par années), ainsi que la liste des co-occurents de l'item donné :

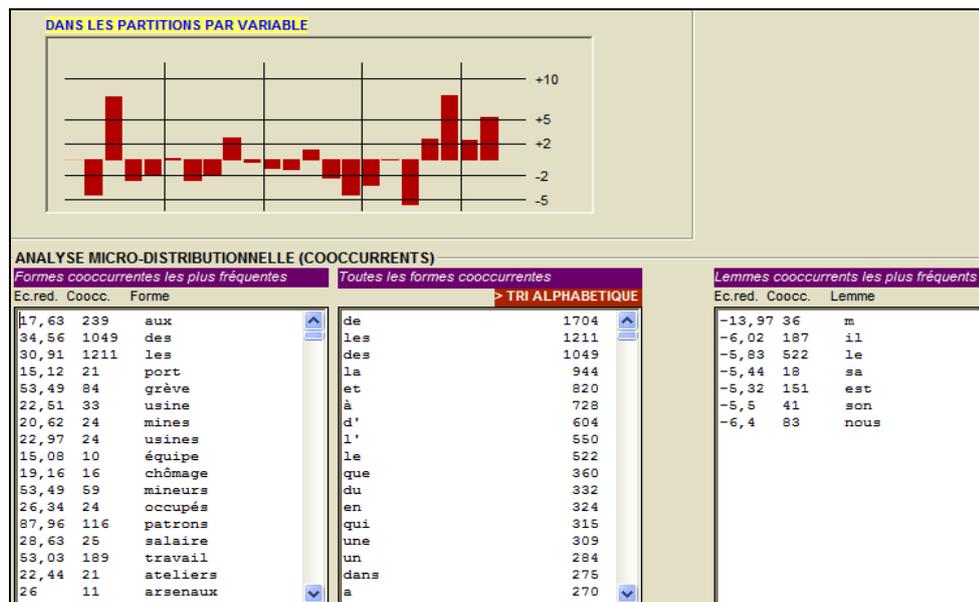


Figure IV-5 : Fenêtre de distribution du vocabulaire.

Parallèlement à l'élaboration de l'hypertexte d'exploration que nous venons de présenter, *Astartex* est le terrain d'expérimentations relatives à l'ergonomisation des sorties de graphes d'AFC. Bien que convoquées pour leur capacité à dresser une vue synthétique d'un ensemble de données, les visualisations AFC se révèlent paradoxalement difficiles à lire. Tout chercheur rompu à l'analyse statistique multidimensionnelle concédera bien volontiers que les sorties visuelles classiques d'AFC sur le plan de deux facteurs s'avèrent très souvent illisibles. Et ce, plus spécifiquement en leur centre, au croisement des points d'origine, parasité par un nuage dense d'items, dont ni les positions ni les relations de proximité ne sont significatives. Seuls quelques items en raison de leurs co-occurrences très fortes, notamment relatives au phénomène de lexicalisation (ou de semi-lexicalisation) des syntagmes, se distinguent de cette masse d'information non pertinente. De fait, les phénomènes mis à jour par les méthodes d'AFC apparaissent déformés. Dans le dessein de faire apparaître les informations amenées par l'exploitation du troisième facteur, *Astartex* propose depuis 2006 un mode de projection *sphérique* ou *géodésique*.

La projection géodésique se présente sous la forme d'un planisphère où les parties significatives du nuage de points se distinguent de façon très marquée (et ce, même sans recours à une solution d'écrêtage des données). La distance à l'origine dans les trois dimensions est signalée par le jeu des niveaux de gris et tailles de police : plus un point se trouve proche du centre, moins celui-ci sera visible, tout en restant repérable. Ainsi le parasitage dû aux faits de lexicalisation de cooccurrences est en bonne partie résolu. Comparons ainsi une sortie traditionnelle d'AFC sur deux facteurs, et la même projection sous mode géodésique :

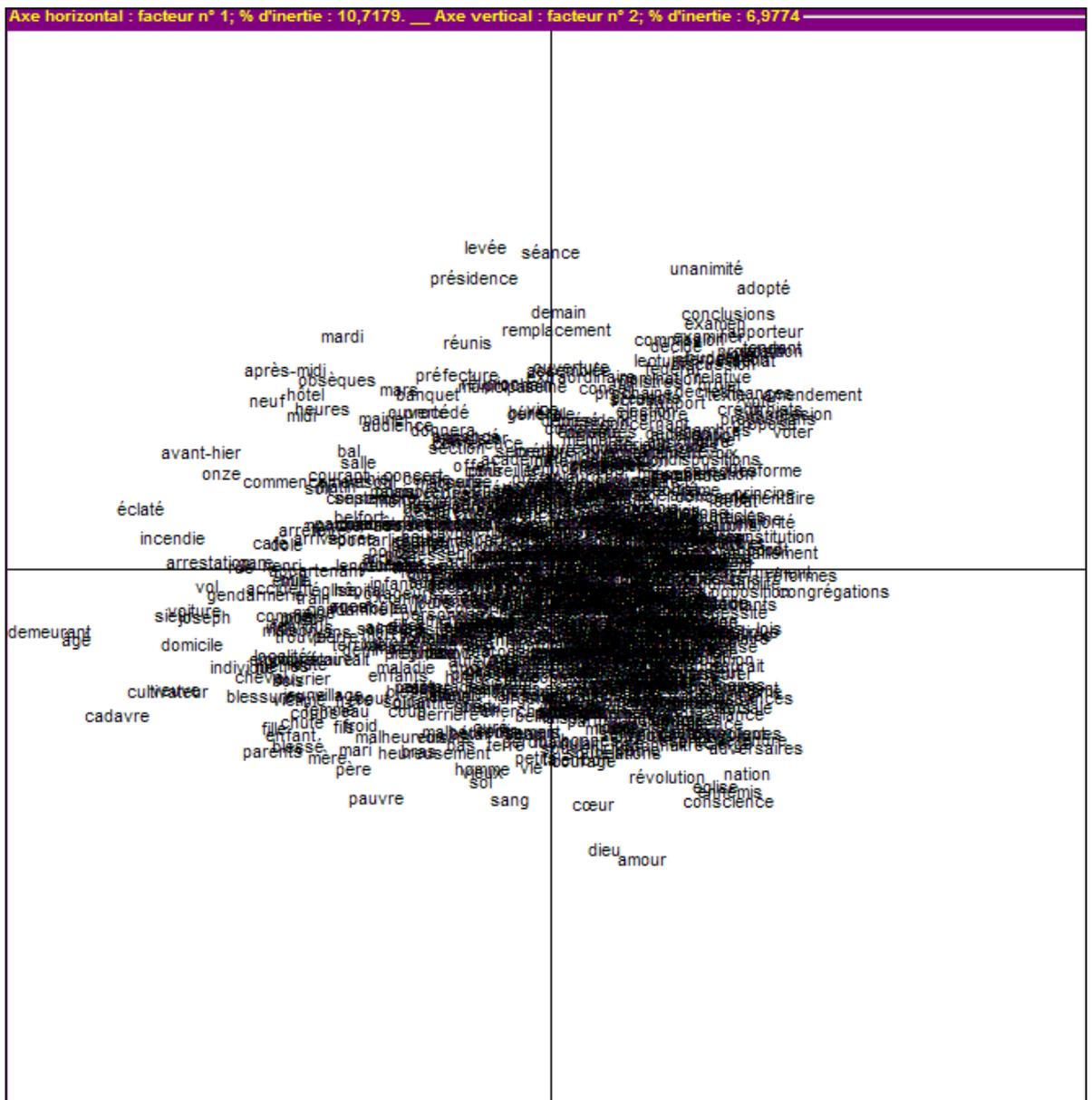


Figure IV-6 : AFC traditionnelle de la micro-distribution des 1000 premiers items du corpus *Petit Comtois*.

D'un clic, les zones d'ensembles distributionnels s'offrent à la lecture « zoomée », permettant un accès facilité à l'étude des configurations de la structure du vocabulaire. Remarquons ainsi le nuage de points assez dense sur le graphe d'AFC ci-dessus, encadré par nos soins d'un rectangle. D'un clic, *Astartex* permet d'accéder à une visualisation focalisée sur ce nuage avec une lisibilité très supérieure :

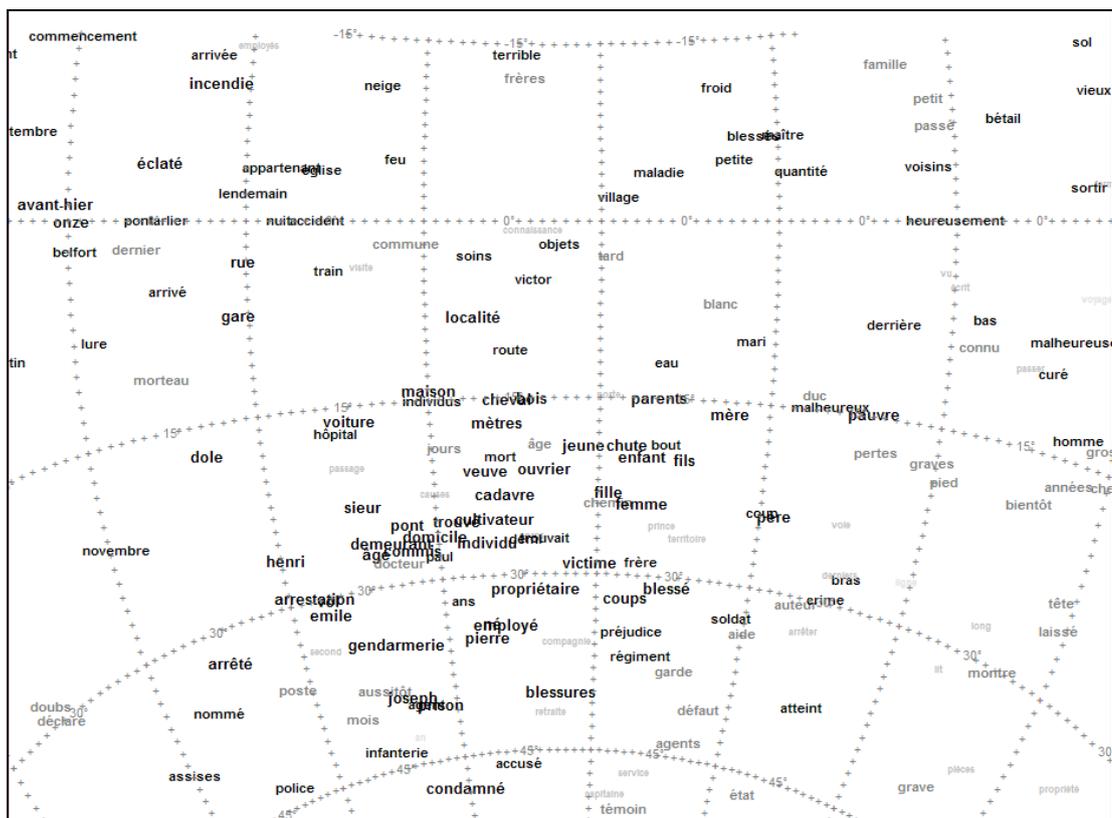


Figure IV-8 : Zoom sur région du graphe géodésique de micro-distribution des 1000 premiers items du corpus *Petit Comtois*.

Le développement de techniques visant à optimiser le potentiel de visualisation de l'AFC est d'autant plus important que l'ensemble de données textuelles analysées est conséquent. Dans le cadre d'une lecture assistée informatiquement de l'archive et de ses millions d'occurrences, le besoin d'obtenir des méthodes de visualisations qui restent lisibles même lorsqu'elles

synthétisent la distribution d'un nombre très important de formes lexicales s'impose comme enjeu incontournable.

Module complémentaire d'*Astartex*, *DiaTag* est un environnement d'établissement du texte numérisé (modules de normalisation graphique, d'étiquetage lexical et flexionnel, de balisage), alternant des phases automatiques et des phases ouvertes interactives et conviviales en vue d'un étiquetage fin et contrôlé. La lemmatisation experte offerte par *DiaTag* répond ainsi aux besoins d'une analyse *textuelle du discours* (ATD) soucieuse de respecter la matérialité textuelle du texte.

Face au caractère expérimental de notre projet d'étude, c'est très logiquement que nous avons fait le choix d'analyser notre corpus à l'aide d'un laboratoire d'expérimentations relatives aux besoins d'une analyse textuelle du discours. Le fait qu'*Astartex* tolère une taille de corpus supérieure à celle nécessitée par le *Petit Comtois* selon un nombre de partitions dépassant très largement les capacités des autres logiciels d'ADT a donc été l'une des nombreuses raisons ayant orienté notre choix vers ce logiciel.

4.2. Manipulations pré-exploratoires sur le vocabulaire du *Petit Comtois*

4.2.1. Consultation de l'index hiérarchique

de	231727	5,32%	il	44881	1,03%	ne	21516	0,49%	ses	9555	0,22%	était	7837	0,18%
la	141825	3,26%	une	40270	0,92%	se	20849	0,48%	lui	9363	0,21%	leur	7776	0,18%
le	104539	2,40%	m.	40069	0,92%	n'	20497	0,47%	sa	9336	0,21%	tout	7659	0,18%
à	100186	2,30%	qui	39448	0,91%	sur	19575	0,45%	vous	9289	0,21%	ils	7626	0,18%
l'	89512	2,06%	est	36014	0,83%	s'	18913	0,43%	fait	9160	0,21%	comme	7597	0,17%
et	87003	2,00%	dans	35082	0,81%	été	16943	0,39%	y	9115	0,21%	être	7399	0,17%
les	79738	1,83%	pour	30662	0,70%	cette	15293	0,35%	je	9098	0,21%	faire	7197	0,17%
d'	60070	1,38%	au	30019	0,69%	plus	14926	0,34%	bien	8886	0,20%	ou	7093	0,16%
des	58827	1,35%	qu'	29090	0,67%	son	14615	0,34%	c'est	8825	0,20%	tous	6907	0,16%
a	52779	1,21%	par	28629	0,66%	ont	14527	0,33%	elle	8518	0,20%	dont	6470	0,15%
que	51995	1,19%	nous	25601	0,59%	aux	12577	0,29%	si	8498	0,20%	notre	6400	0,15%
du	51234	1,18%	pas	23108	0,53%	sont	12294	0,28%	ces	8453	0,19%	même	6233	0,14%
un	46389	1,07%	on	22821	0,52%	avec	11714	0,27%	avait	8307	0,19%	sans	6192	0,14%
en	46277	1,06%	ce	22296	0,51%	mais	10942	0,25%	deux	7938	0,18%	où	6010	0,14%

Tableau IV-1 : Index des 70 formes les plus fréquentes du corpus *Petit Comtois*.

En tête de liste des mots les plus utilisés, on observe l'habituelle prédominance des mots-outils, remarquables dans tous les corpus de langue française. L'ordre de fréquence des mots-outils est tout à fait commun : ainsi, « de » est la forme la plus fréquente de notre corpus, avec 231 727 occurrences. Outre la prédominance inversée du « la » sur « le », il est néanmoins intéressant de noter que la forme « m. », correspondant à l'abréviation de *monsieur*, figure au 17^{ème} rang des formes les plus utilisées. *Nous* est le premier pronom personnel dans l'ordre hiérarchique : son rang dans l'index offre un indice fort de l'appartenance du *Petit Comtois* à une presse d'opinion. Le pronom de première personne du singulier est fréquemment utilisé, indiquant une prise en charge du discours cohérente avec notre précédente remarque. Le pronom d'adresse « vous » est utilisé dans la même proportion. En raison des caractéristiques de la langue française mais aussi du volume de notre corpus, les mots pleins ne figurent qu'à des rangs très nettement inférieurs dans la hiérarchie des emplois.

forme	occ.	forme	occ.	forme	occ.	forme	occ.
heures	5827	partie	1993	bonne	1506	applaudissements	1161
lieu	4766	membres	1989	commune	1502	trouvé	1157
conseil	4423	mois	1989	républicains	1496	suivante	1156
paris	4411	nom	1968	ministère	1489	enquête	1154
fr	4391	mort	1958	jeune	1485	but	1145
général	4310	nouvelle	1909	ensuite	1463	administration	1138
gouvernement	4223	jamais	1894	gauche	1454	députés	1133
chambre	4125	maison	1850	monde	1448	municipal	1123
ministre	4016	donné	1836	commerce	1437	seul	1120
france	3955	article	1833	sénat	1424	samedi	1116
président	3916	affaire	1809	prix	1412	cours	1111
ans	3877	somme	1782	amis	1410	paraît	1108
soir	3604	dimanche	1759	salle	1400	nombreux	1101
aujourd'hui	3600	honneur	1756	coup	1383	demie	1100
temps	3575	seulement	1749	instruction	1378	suisse	1100
hier	3488	travail	1749	public	1367	tête	1098
commission	3404	journaux	1740	républicain	1357	intérêts	1097
jour	3345	sujet	1732	comité	1350	œuvre	1090
loi	3337	ouvriers	1720	justice	1346	rapport	1086
matin	3262	e	1718	cour	1339	compagnie	1080
question	3245	nouveau	1717	dernière	1334	suivant	1079
ville	3204	courant	1715	famille	1323	française	1069
république	3202	discussion	1715	sait	1307	âgé	1065
mm	3107	hommes	1701	affaires	1303	paix	1060
grand	3100	francs	1700	discours	1298	argent	1059
jours	3023	voix	1697	femme	1297	jeunes	1051
besançon	2959	député	1694	travaux	1291	allemagne	1046
société	2758	personnes	1692	dépêche	1285	déclaré	1042
dernier	2755	état	1675	école	1285	militaire	1040
journal	2735	français	1671	succès	1278	veut	1040
séance	2641	budget	1668	proposition	1273	canton	1030
droit	2583	mme	1666	vue	1272	savoir	1026
projet	2567	premier	1666	occasion	1265	note	1024
politique	2549	chef	1630	enfants	1260	presse	1020
demande	2513	heure	1630	nuit	1252	vie	1017
pays	2499	première	1610	raison	1251	intérêt	1016
nommé	2378	bon	1607	arrêté	1250	produit	1016
guerre	2339	police	1607	écrit	1249	générale	1015
toujours	2333	bureau	1603	porte	1243	donne	1011
grande	2327	service	1594	reste	1235	fer	1011
nombre	2325	armée	1591	ancien	1225	vu	1011
maire	2282	parti	1584	prochain	1216	avis	1010
rue	2253	réunion	1581	millions	1204	doubs	1008
lettre	2238	situation	1569	demain	1203	électeurs	987
ordre	2209	voir	1562	liberté	1188	directeur	985
suite	2172	année	1546	annonce	1186	main	985
petit	2135	reçu	1542	présidence	1180	décidé	983
donner	2090	corps	1537	devoir	1171	feu	982
homme	2046	fort	1530	concours	1166	incendie	979
etat	2039	publique	1529	sieur	1166	nouvelles	979

Tableau IV-2 : Les 200 premières formes (mots-pleins) du corpus *Petit Comtois*.

Le tableau ci-dessus indique les formes apparaissant hiérarchiquement parmi les plus fréquentes du corpus, mots-outils comptés à part. Ce sont tout d'abord les caractéristiques d'une production discursive médiatique du XIX^e siècle qui se laissent apercevoir dans cet index, comme en témoigne l'abondance des marqueurs de temps, déictiques (*aujourd'hui, hier*) ou non (*soir, matin, janvier, mars*). Culminant en tête de l'index hiérarchique avec 5 827 occurrences, la forme « heures » souligne le rôle primordial de la temporalité dans le discours journalistique, qui insère en tête de ses nouvelles leurs dates et horaires de provenance, pour exhiber l'actualité. On notera également l'utilisation massive de formes référant au champ journalistique : *journal, journaux, article, dépêche et presse*.

Parmi les toponymes figurant dans cet index, *paris* arrive en tête de liste, avec 4 411 occurrences, très loin devant *besançon* (2 959 occurrences). Cette indication pourrait sembler paradoxale alors que le corpus traité relève de la presse régionale, supposée privilégier les informations de proximité. Cette fréquence très importante du toponyme *paris* traduit d'emblée l'ouverture du quotidien sur les événements nationaux, et pointe le centralisme politique parisien de la Troisième République : si la capitale française affleure aussi nettement à la surface du discours du *Petit Comtois*, c'est parce qu'elle est le lieu par excellence du politique, centre d'intérêt primordial de la presse au carrefour du XIX^e siècle et du XX^e siècle. En témoignent les items *gouvernement, ministre, président, politique, député, parti, état, électeurs* qui culminent en bonne place de l'index hiérarchique. À la politique en tant que champ lexico-thématique, il faut également rattacher les formes *république, justice, liberté, honneur et succès*, qui peuvent être considérées comme autant de notions martelées par le quotidien républicain, comme l'indique leur nombre important d'occurrences. Nous faisons ici le choix de ne pas aller plus en avant dans l'interprétation de ces premiers résultats, du fait des limites inhérentes à la pratique de consultation d'un index hiérarchique. Les informations fournies par ce type d'inventaires basé sur la comptabilité des graphies présentent en effet

le risque de donner lieu à une interprétation par essence hasardeuse de formes, en cela que leur sens n'est accessible qu'en contexte. Un retour au texte montre ainsi très clairement qu'un item tel que *justice* ne correspond en contexte pas exclusivement à un nom commun, mais qu'il est également employé comme un nom propre, qui désigne un titre de presse nationale abondamment cité par *Le Petit Comtois*.

Ch.IV (01), « La démission du ministre de la guerre », Dépêches de nuit, (1885-01-03)

Tous les journaux commentent la démission du ministre de la guerre, qui est certaine, affirme-t-on.

Parmi les généraux indiqués pour la succession à ce portefeuille, figure le général Wolf. Cependant, le choix du gouvernement paraît s'être fixé sur le général Lewal. Déjà les feuilles radicales intransigeantes disent que le général Lewal est orléaniste.

Le nouveau ministre n'a pas encore pris possession de son portefeuille qu'il est déjà attaqué par la *Justice*, organe de M. Clémenceau.

Les premières caractéristiques du vocabulaire du corpus ayant été esquissées par l'examen de l'index hiérarchique, nous nous proposons à présent d'interroger la ventilation du vocabulaire en diachronie. Il s'agit précisément de procéder à une analyse contrastive du vocabulaire menée à partir de la norme endogène de notre corpus pour ensuite observer les régularités et les ruptures du discours du *Petit Comtois* qui émergent ainsi.

4.2.2. Examen de la distribution du vocabulaire du *Petit Comtois* en diachronie

4.2.2.1. Analyse de la distribution du vocabulaire du *Petit Comtois* en diachronie

Pour étudier l'évolution du vocabulaire du *Petit Comtois* en diachronie, nous avons soumis les 1000 formes les plus fréquentes du corpus aux procédures de l'Analyse Factorielle des Correspondances (Lebart & Salem,

1994). Nous avons retranché de cette liste les mots-outils, les noms de mois¹⁵, ainsi que certaines formes parasites¹⁶, ce qui porte à 988 le nombre définitif d'items analysés¹⁷. Bien que parasitée en son centre par un nuage dense d'items, la représentation bidimensionnelle (Figure IV-9) issue de cette analyse permet d'obtenir une vue synthétique intéressante de la répartition du vocabulaire du corpus par année de parution. S'il s'agit d'une représentation classique d'AFC, où les axes 1 et 2 expliquent respectivement 14,2% et 9,6% de la variance totale, ce graphique implique de respecter certaines règles d'interprétation particulières.

¹⁵ Bien que méthodologiquement, le choix de conserver les mois de référence des échantillons annuels ne semble pas biaiser les données.

¹⁶ Ces formes parasites résultent majoritairement de scories de correction (Exemple : « Sic », « ii », etc.).

¹⁷ Les mots-outils ont ici été expurgés de l'analyse afin de ne pas perturber la lisibilité du graphique d'une part, et d'autre part, en raison de la faible pertinence de l'information apportée.

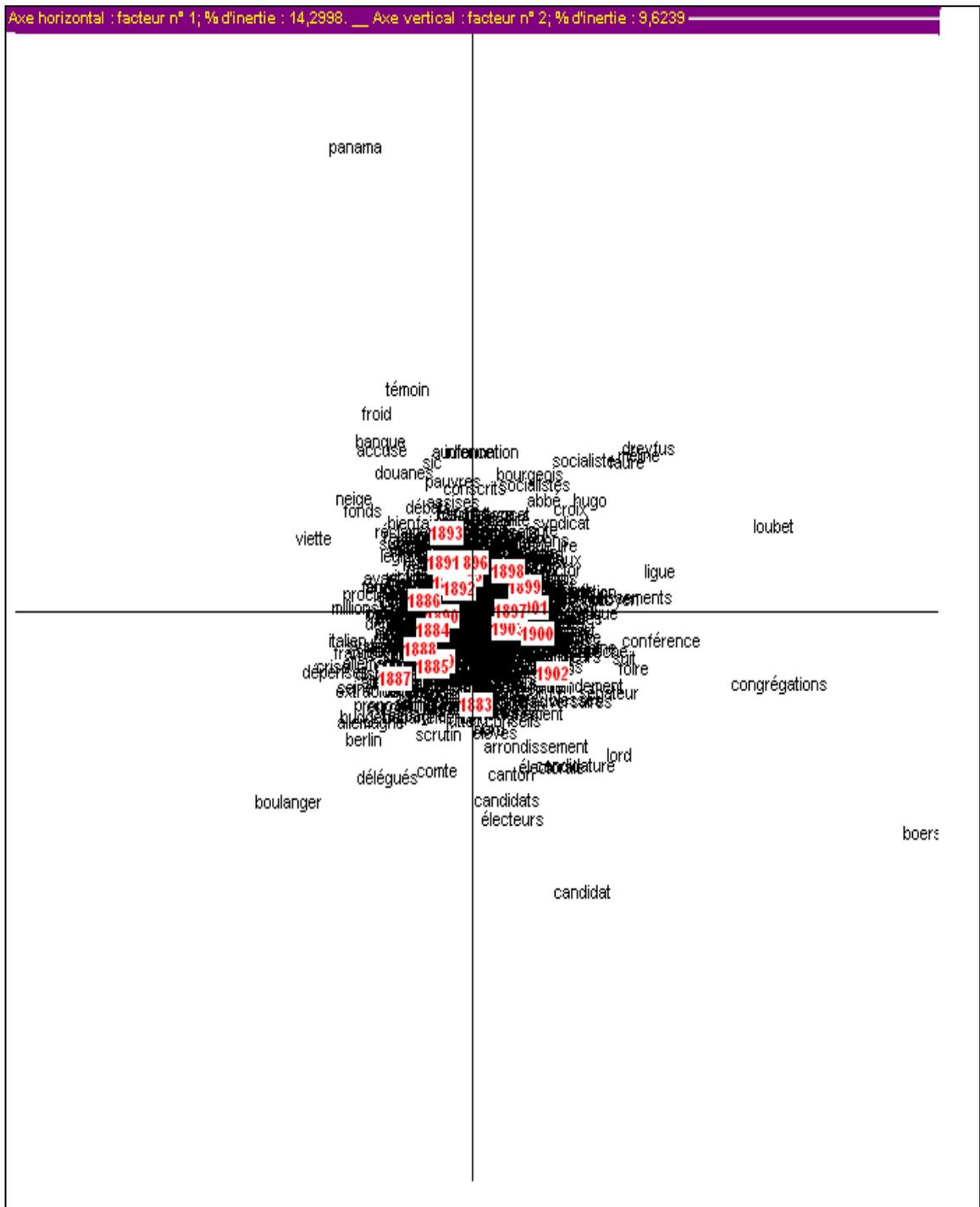


Figure IV-9 : Analyse macro-lexicale des 1000 items les plus fréquents du corpus.

On doit à A. Salem (1994) d'avoir posé que la lecture d'une sortie AFC d'une *série textuelle chronologique*, en tant que corpus de textes produits dans des

conditions similaires par un locuteur collectif, implique d'autres garde-fous méthodologiques que ceux présidant traditionnellement à l'interprétation d'un graphe AFC représentant l'analyse d'un corpus contrastif. Ce rappel, D. Mayaffre l'explique en ces termes :

Tout au contraire, les AFC portant sur des corpus constituant des séquences textuelles chronologiques doivent être regardés dans une vision d'ensemble par la considération de l'harmonie des points d'un alignement incurvé autour du centre de gravité du graphique (Mayaffre, 2000d :10).

L'étude de séries textuelles chronologiques pointe généralement une organisation particulière du vocabulaire en diachronie, reposant sur des convergences lexicales très fortes entre les périodes consécutives et des dissemblances entre des discours séparés par un intervalle de temps plus long. Ce phénomène, nommé *temps lexical* (Lebart & Salem, 1994), (Salem, 1998), est imputable au fait que :

Tout émetteur produisant des textes sur une période assez longue utilise sans cesse de nouvelles formes de vocabulaire qui viennent supplanter, du point de vue fréquentiel, d'autres formes dont l'usage se raréfie. Il s'ensuit que les vocabulaires des parties correspondant à des périodes consécutives dans le temps présentent en général plus de similitudes entre eux que ceux qui correspondent à des périodes séparées par un intervalle de temps plus long (Salem, 1994 : 221).

Sur le plan des deux premiers facteurs, cette caractéristique se traduit par une suite régulière et chronologique des parties du corpus et donc, par un ordonnancement parabolique lié à certaines règles mathématiques connues dans la littérature spécialisée en tant qu'*effet Guttman* (Benzécri, 1969). Cette remarque est néanmoins d'autant plus valable que la diachronie envisagée est longue. C'est pourquoi on remarquera d'emblée sur la figure IV-9 une disposition moins régulière (que pour une série diachronique plus étendue) des périodes sur le premier axe factoriel, et la quasi absence d'un ordonnancement parabolique. Très classiquement, les rares items aux profils très saillants s'avèrent pour la plupart des noms propres et des toponymes correspondant à des événements fortement circonscrits dans le temps (*boers, panama, dreyfus, faure, loubet, boulanger, congrégations, boers*). Au-delà de cette première

remarque, on notera sur le facteur 1 la position neutre de l'année 1883 (la première de la période, en rupture donc avec tout ordonnancement temporel), mais surtout la très nette opposition entre, d'une part, la première tranche diachronique du corpus (1884-1896), qui s'inscrit dans la partie gauche du graphique, et d'autre part, une seconde tranche diachronique, à droite, correspondant à la période 1897-1903.

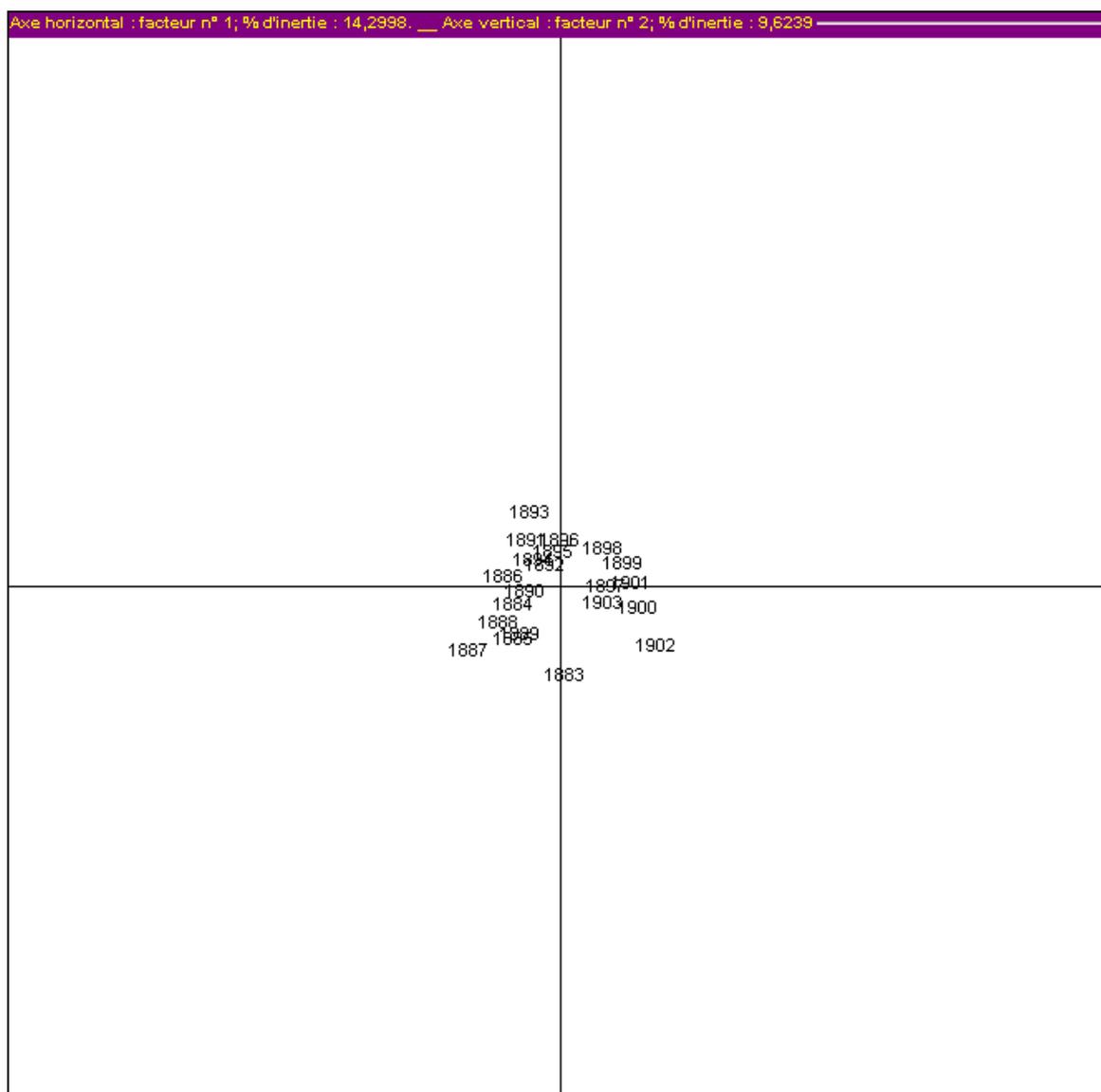


Figure IV-10 : Esquisse de l'analyse de la macro-distribution du corpus.

On ne pourra s'empêcher de signaler au passage et à nouveau la mauvaise lisibilité des représentations AFC en deux dimensions de l'analyse de séries chronologiques, qui malgré une réduction de la projection aux seuls noms des variables en colonnes (figure IV-10), demeure un obstacle fort dans la recherche d'indications plus fines. On peut certes observer que les figures IV-9 et IV-10 se distinguent d'un net schéma parabolique classique à *effet Gutman* et que certaines années se croisent selon un ordre chronologique non régulier et discontinu. Il s'avère néanmoins délicat d'accéder à toute observation plus fine. Aussi procéderons-nous ici à une projection géodésique. Le recours à une telle sortie AFC en trois dimensions permet, d'une part, de préciser les continuités et ruptures des séries chronologiques par une visibilité accrue, et d'autre part, de livrer de précieuses informations sur les coordonnées du troisième facteur. La visualisation AFC suivante (cf. figure IV-11) a été construite à partir des résultats d'analyse de la même matrice utilisée pour l'établissement des deux précédentes figures. Elle recoupe en premier lieu les informations livrées par les figures IV-9 et IV-10 en indiquant une opposition sur le facteur 1 des années 1883-1896 d'une part, 1897-1903 de l'autre. Elle met également en exergue la très singulière position de l'année 1883, neutre sur l'axe 1, saillante à la fois sur l'axe 2 et sur l'axe 3, qui se démarque de façon très visible de la série 1884-1890. À partir de ce constat, nous pourrions émettre l'hypothèse selon laquelle la saillance de l'année 1883 pourrait être l'indice d'une inflexion du discours du *Petit Comtois* après les débuts immédiats de parution du quotidien.

La figure IV-11 nous permet parallèlement d'affiner notre observation de la succession des années du corpus. Si l'on ne pouvait précédemment que deviner, faute de lisibilité, une opposition sur le deuxième facteur entre la série 1884-1890, la sortie géodésique permet de repérer des groupements d'années,

que nous envisagerons comme des *séquences chronologiques*¹⁸. On distinguera ainsi que la période 1884-1890 s'oppose nettement à la tranche 1891-1896 du fait des contributions du facteur 2. Située dans la moitié droite du planisphère, la série 1897-1903 forme une zone plus diffuse que les séries précédentes. Sur le facteur 2, les années 1900-1902-1903 tendent à s'inscrire dans le quart-ouest de la moitié gauche du plan et les années 1897-1898-1899-1901 dans son quart-est. Nous distinguerons de fait deux sous-ensembles imbriqués, 1897-1898-1899-1901 et 1900-1902-1903, pour rendre compte des spécificités de cette tranche diachronique.

¹⁸ Pour éviter toute confusion avec la notion de *série textuelle chronologique* (Salem, 1998), nous utiliserons ici le terme de « séquence chronologique » pour désigner les regroupements de variables pointées par l'AFC.

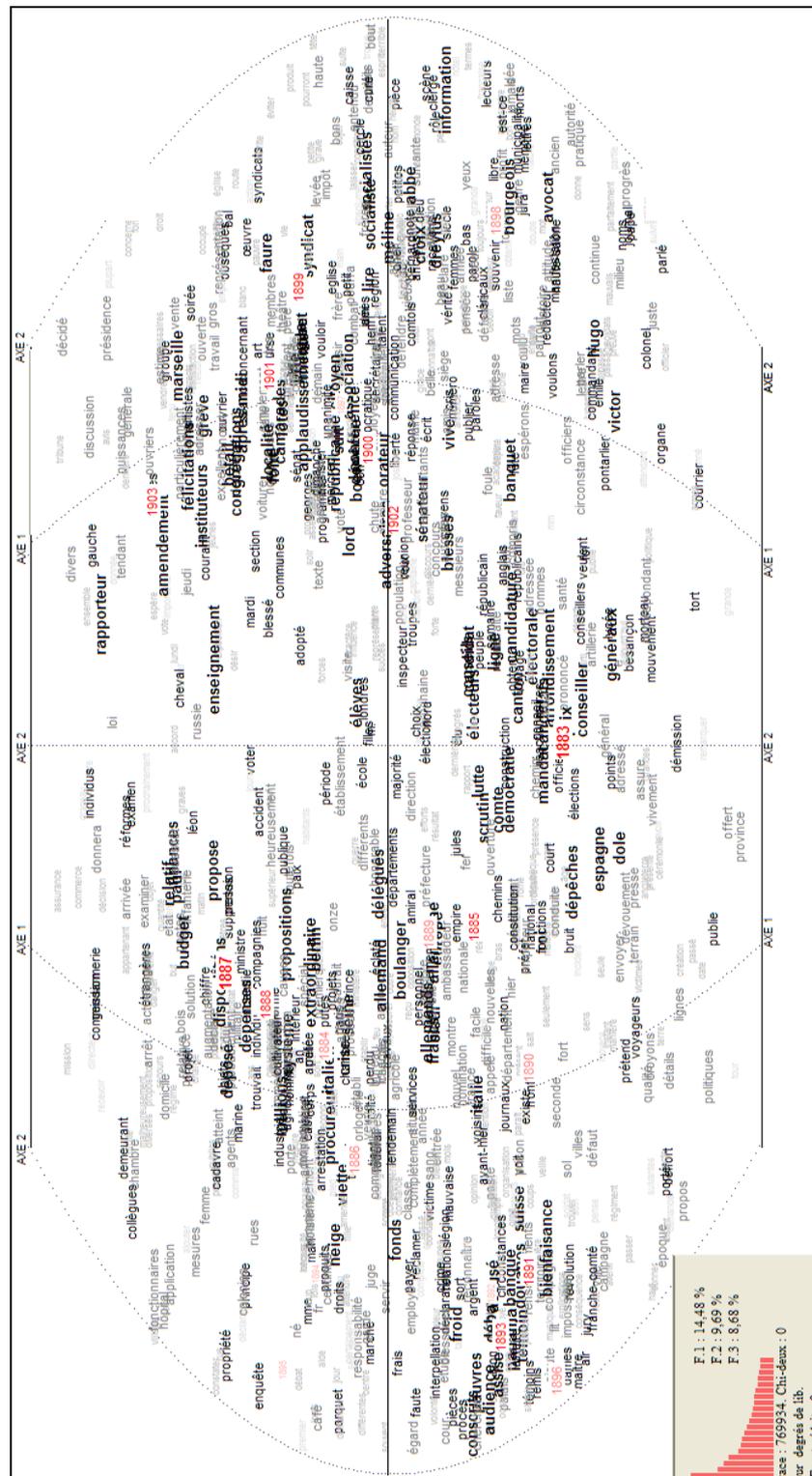


Figure IV-11 : Projection AFC de la configuration fine du vocabulaire du *Petit Comtois*.

En somme, le premier examen de l'analyse de la distribution du vocabulaire pointe, au-delà de quelques ruptures dans les bornes chronologiques extrêmes du corpus, une dichotomie diachronique importante. Il convient dès lors de prendre connaissance du vocabulaire propre de chacune des séquences chronologiques mises au jour par l'AFC. Pour ce faire, nous nous orienterons vers l'examen détaillé des zooms correspondants.

4.2.2.2. Examen du vocabulaire propre des séquences chronologiques

L'année 1883

En premier lieu, nous nous proposons de prendre connaissance des items spécifiques de l'année 1883.

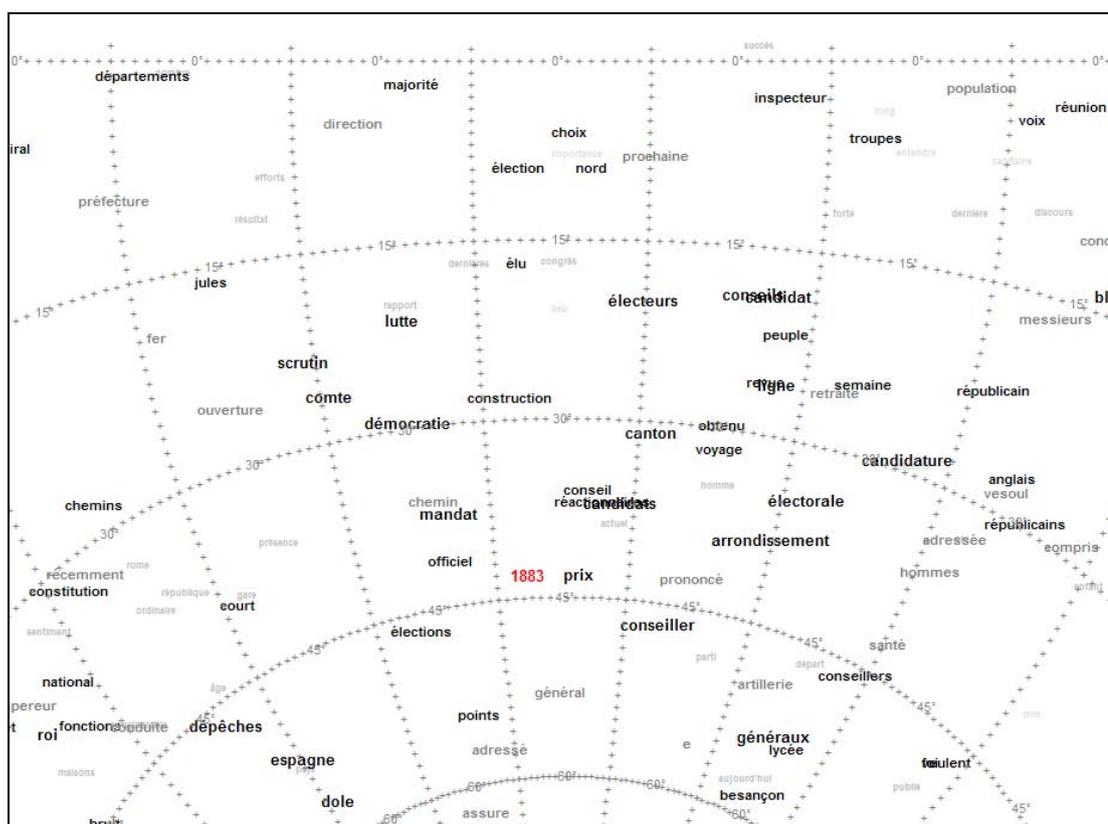


Figure IV-12 : Zoom sur la région 1883.

Nous avons émis précédemment l'hypothèse selon laquelle le profil saillant de cette année sur le planisphère de la distribution du vocabulaire en diachronie pouvait être relatif aux débuts de parution du quotidien *Le Petit Comtois*. On remarquera sur ce zoom centré de l'année 1883 que le vocabulaire propre de cette année est majoritairement constitué d'items relatifs au champ électoral (*scrutin, candidats, élection, électorale, mandat, candidature*). Le retour au texte nous indique que le mois d'août 1883 est une période d'élections des conseillers généraux : des traces de cet événement du calendrier électoral sont aisément repérables sur la figure IV-12 par le biais des formes lexicales *arrondissement, conseiller, généraux, canton*. On remarquera également que le vocabulaire propre de l'année 1883 s'articule autour des items *lutte, démocratie, réactionnaires, républicains, républicain* manifestant la confrontation des deux camps politiques en présence lors de cette élection. Les spécificités de l'année 1883 semblent ainsi principalement liées au calendrier politique de cette période, qui se révèle un facteur fortement structurant de ces productions journalistiques, dont les articles ci-dessous sont représentatifs :

Ch.IV (02), Chronique régionale Jura, « Elections départementales », (03-08-1883)

Bletterans. M. Bellissime, pharmacien, maire de Bletterans, a été choisi par les électeurs du canton, pour succéder à M. Molard, conseiller d'arrondissement, récemment décédé. Il n'y a pas de compétiteurs.

Arinthod. M. Labordère, ancien préfet, pose sa candidature au conseil général. Son élection est assurée. M. Labordère, ancien avocat au conseil d'Etat, a été successivement préfet de la Haute-Loire, des Landes et des Côtes-du-Nord ; il est actuellement en disponibilité.

Ch.IV (03), Chronique régionale Haute-Saône, « Vesoul », (17-08-1883)

Les esprits sont encore à l'élection de dimanche dernier.

L'*Avenir de la Haute-Saône* constate la multiplicité des candidatures républicaines dans les quatre cantons d'Autrey, Frosne-Saint-Mamès, Faucogney et Lure. L'*Avenir* interprète cette pluralité de candidats de même nuance comme un signe de force pour le parti républicain, qui ne se diviserait pas s'il se trouvait en présence de candidats réactionnaires.

Nous ne demandons pas mieux d'être de l'avis de notre excellent confrère, mais nous ne cachons pas que nous préférierions une entente absolue avant tous les scrutins. En somme, le résultat est excellent dans la Haute-Saône. Les républicains gagnent un siège, celui de Combeaufontaine, où l'honorable

M. Parguez a été élu, malgré ou peut-être à cause du patronage donné à son concurrent réactionnaire par le conseiller sortant, ancien membre des commissions mixtes. Les républicains gagnent de plus quatre sièges au conseil d'arrondissement.

D'autres éléments du vocabulaire propre de cette période nous indiquent cependant que le calendrier électoral n'est pas le seul facteur expliquant la position tout à fait saillante de l'année 1883 et semblent, au contraire, converger vers l'hypothèse de spécificités lexicales propres aux débuts de parution du quotidien. Il est ainsi intéressant de remarquer sur ce zoom la forme *dépêches* et d'examiner son profil distributionnel à partir de l'histogramme suivant, construit à partir de l'écart-réduit de l'effectif réel et de l'effectif théorique de l'item dans les différentes sections chronologiques du corpus. On observe très nettement un sur-emploi de la forme dans la section 1883, tandis que son profil distributionnel est très souvent non-significatif dans les autres sections, voire déficitaire dans les dernières années du corpus.

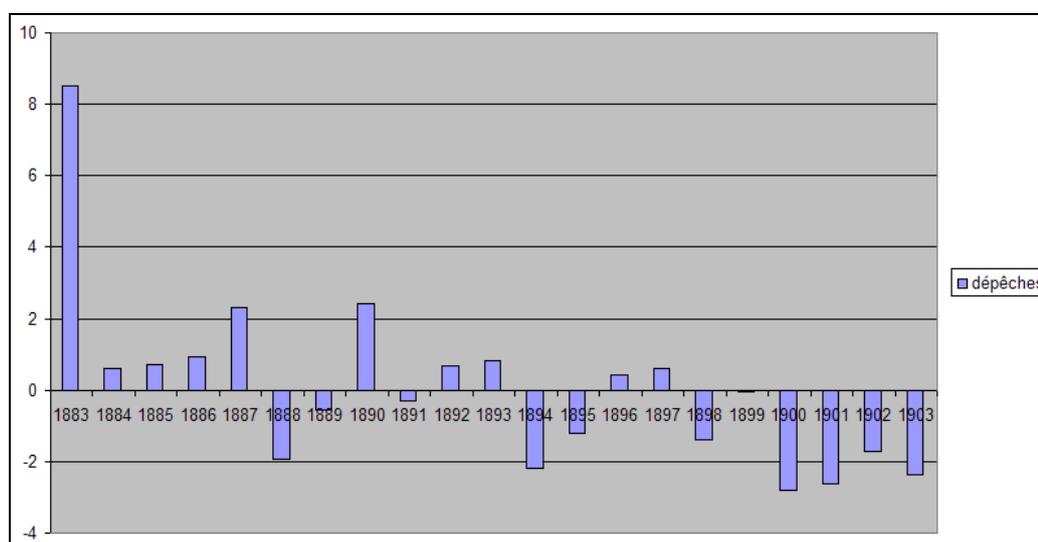


Figure IV-13 : Histogramme de la distribution de la forme *dépêches* (en écarts-réduits).

Le sur-emploi de la forme (*dépêches*) qui désigne un matériau essentiel du discours journalistique dans la première année de parution du quotidien, nous semble légitimement pouvoir être interprété comme l'indice d'une pratique

journalistique en construction. Durant la première année de parution du quotidien, la circulation des discours, l'origine des informations, sont tout particulièrement mis en relief par la rédaction du *Petit Comtois*, comme nous le montre le retour au texte. Les nombreuses occurrences de l'item pointent tout d'abord une activité de reformulation vis-à-vis du matériau brut que sont les dépêches.

Ch.IV (04), Dépêches de nuit, « Marche décroissante du choléra », (1883-08-11)

Toutes les dépêches reçues s'accordent à annoncer que le nombre des décès cholériques diminue. La maladie entre sérieusement dans une période de décroissance.

Il est très intéressant de noter que la plupart des occurrences de *dépêches* s'accompagnent parallèlement d'une prise de distance avec l'information qu'elles rapportent. Ainsi, dans les exemples suivants, on remarquera l'insistance du quotidien à mettre en garde le lecteur sur la véracité des dépêches relatives à une insurrection en Espagne, que la rédaction juge « obscures » et « contradictoires » :

Ch.IV (05), Dépêches de nuit, « Révolution en Espagne », (1883-08-07)

Paris le 6 août, 8 h. 42 soir.

Une insurrection en faveur de la république a éclaté à Badajoz, dans l'Estramadure, pendant la nuit de dimanche à lundi. Le mouvement paraît avoir été concerté entre civils et militaires. [...] Des dépêches obscures et contradictoires disent que les conjurés sont au nombre de onze cents, tant civils que militaires, et que l'insurrection a été réprimée par des troupes venues de Madrid. Elles ajoutent que le gouvernement réprimera ce mouvement avec la dernière énergie. La vérité est que la situation n'est pas encore exactement connue, que les conjurés ont proclamé la république, et qu'ils ont acclamé M. Zorilla comme chef de leur gouvernement. [...] Il est prudent de se méfier des dépêches qui représentent l'insurrection comme réprimée, avant que les troupes aient eu le temps de se transporter et de transporter le matériel de guerre sur le théâtre des événements.

Ch.IV (06), Dépêches de nuit, « Affaires d'Espagne », (1883-08-11)

Les dépêches d'Espagne continuent à être contradictoires. Une dépêche du conseil des ministres à Madrid dit que les insurgés de Barcelone n'ont pas touché aux six millions renfermés dans la banque. Donc les insurgés ont été maîtres un

instant de Barcelone. Cependant certaines dépêches prétendent que l'insurrection a été réprimée sans peine et que les révoltés sont en fuite. [...]

On peut dès lors s'interroger sur les raisons présidant à cet état de fait. Les exemples ci-dessus suggèrent une relation entre l'emploi insistant observé de la forme plurielle *dépêches* et la volonté de la rédaction de confronter plusieurs sources d'information en temps de conflits. Ce procédé rédactionnel n'est cependant pas l'apanage des périodes de conflits ou de guerre, bien que ce climat semble la favoriser. Pendant la première année de parution, les brèves ne cessent de faire circuler des informations, dont l'origine énonciative est effacée au profit de la mise en lumière de leur mode de transmission. Il ne s'agit donc pas tant pour la rédaction de « recouper » des informations pour en confirmer la valeur et la véracité. Envisager cette pratique comme les traces d'une déontologie journalistique nous semblerait abusif, puisque la déontologie journalistique est en effet une notion anachronique au XIX^e siècle. La substitution de la voix au profit de la technologie nous semble devoir être au contraire éclairée par le statut très fortement valorisé de la dépêche au XIX^e siècle. Parce qu'elle est le fruit d'une technologie fraîchement élaborée, la dépêche est mise en scène en tant qu'argument commercial, ce qu'on peut observer dans tous les quotidiens à la fin du XIX^e. La représentation de fils télégraphiques dans la manchette du quotidien national *Le Matin*, le premier article paru du *Petit Comtois* sont tout autant de formes de monstration de la technologie maîtrisée par le journal, qui est mise en scène par de multiples biais. Il fait sens que, dans son premier article où le quotidien se présente, *Le Petit Comtois* se targue d'accéder aux fils télégraphiques officiels et à la totalité des sources d'informations disponibles, mais aussi que la rubrique des *Dépêches de nuit* occupe une place stratégique dans l'aire scripturale du quotidien. Toutes ces stratégies de mise en scène relèvent d'une forme d'idéologie visant à s'emparer de l'image valorisée du progrès technique et de la modernité pour asseoir et valoriser l'ethos du journal. Le sur-emploi de l'item *dépêches* au pluriel est la trace discursive de cette stratégie de monstration de

la technologie : mentionner le flux de dépêches accessibles par le quotidien, c'est montrer sa modernité.

Par ailleurs, on notera que bon nombre d'occurrences de *dépêches* relèvent d'une rubrique figurant en Une du nom de *Bulletin*, qui est tout spécialement consacrée à un retour sur le contenu des dépêches publiées. Cette case rédactionnelle ne paraît dans le *Petit Comtois* que de 1883 à 1885. Ayant une fonction d'appel manifeste, cette rubrique vise à susciter l'attention du lecteur sur le contenu d'une dépêche jugée importante. L'importance qu'accorde la rédaction à cette dépêche témoigne d'un processus rédactionnel de hiérarchisation de l'information et participe également de la diffusion de la ligne éditoriale du quotidien.

Ch.IV (07), Bulletin, (1883-08-11)

Nous ne voyons à signaler dans nos dépêches que deux ou trois nouvelles ayant une signification de quelque valeur. [...]

Ch.IV (08), Bulletin, (1883-08-13)

Nous appelons l'attention du lecteur sur une de nos dépêches annonçant que la *Epoca* journal alphonsiste, accuse la France d'avoir fomenté l'insurrection espagnole.

Outre l'item *dépêches*, d'autres formes observables sur le zoom de l'année 1883 évoquent l'état des pratiques discursives journalistiques. Ainsi, l'item « court » s'avère tout spécialement spécifique de cette année où la forme réalise son seul écart-réduit significativement positif, comme on peut le constater sur l'histogramme suivant.

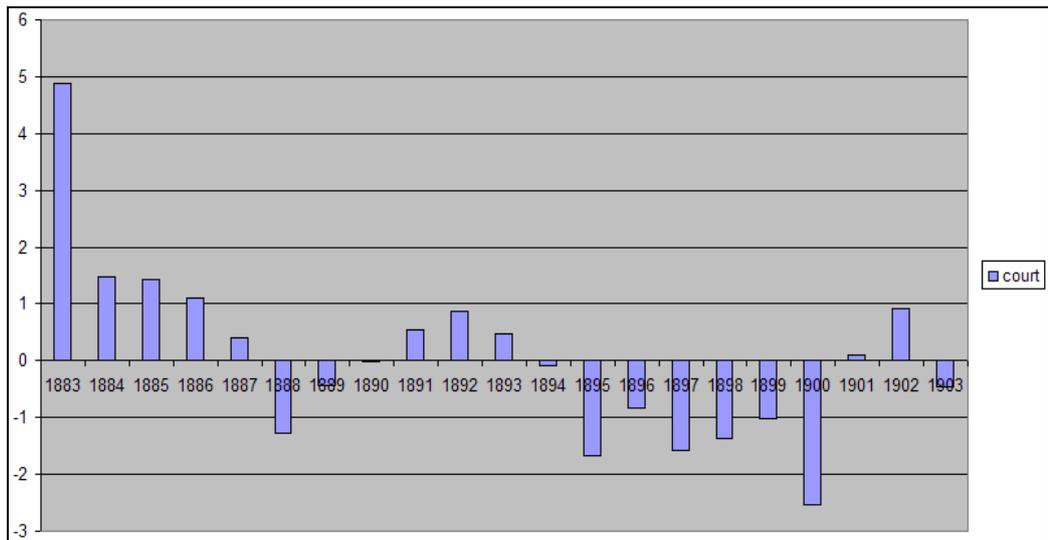


Figure IV-14 : Histogramme de la distribution de la forme *court* (en écarts-réduits).

Un retour au texte montre que cet item est majoritairement employé avec le substantif *bruit*, figurant également sur le zoom de l'année 1883. Visant à introduire un discours rapporté dont la prise en charge est nulle, l'expression « le bruit court », très fréquemment utilisée dans la première année de parution du quotidien, nous semble représenter un indice de l'état des pratiques journalistiques. Manifestement, *Le Petit Comtois*, dans sa première année de parution, colporte la rumeur, que nous définirons comme une information non vérifiée. La non-vérification de l'information délivrée est explicite dans les exemples (9) et (10). Alors que les rumeurs fleurissent dans ses colonnes, le quotidien met l'accent sur la multitude des voix qu'il rassemble, et ce, indépendamment des frontières. *Le Petit Comtois* n'est pas uniquement un compte-rendu des conversations qui se déroulent à l'échelle locale ; il rapporte également au lecteur les rumeurs ayant cours dans l'espace national mais aussi international. Or, à la rumeur, au XIX^e siècle, on attache une dimension séditieuse. On se réfèrera sur ce point aux travaux de l'historien F. Ploux (2003), qui a s'est interrogé sur le rôle de cette originale de communication dans la politisation des Français au début du XIX^e siècle. Comment interpréter la saillance de la rumeur dans la première année de parution du quotidien ?

Le TLFi (Trésor de la Langue Française informatisé) définit la « rumeur » comme une « nouvelle sans certitude qui se répand de bouche-à-oreille, bruit inquiétant qui court »¹⁹. Conformément à cette définition, la prégnance de la rumeur dans les chroniques régionales publiée en 1883 nous semble liée au fait que le réseau de correspondants spéciaux du *Petit Comtois* n'est pas encore suffisamment rodé à la temporalité introduite par le quotidien. La saillance du « bruit qui court » s'expliquerait ainsi par l'inaptitude des correspondants à vérifier l'information dans le laps de temps qui leur est imparti, mais aussi de la persévérance de formes de communication traditionnelle dans la construction du discours journalistique tenu par des non-professionnels. Les exemples ci-dessous illustrent ce cas de figure où l'information est présentée comme rumeur dès lors qu'elle n'a pu être vérifiée ou authentifiée par une source autorisée :

Ch.IV (09), Chronique régionale, (1883-08-25)

Le bruit court qu'à la suite des élections dernières, les maires de Morez et de Saint-Claude auraient donné leur démission.

La *Sentinelle du Jura* s'est même fait l'écho de ces bruits.

Nos correspondants spéciaux ne nous ont encore rien appris à ce sujet.

Ch.IV (10), Chronique régionale, Jura, Dole, (21-08-1883)

Un empoisonnement par les champignons vient de causer une pénible émotion dans la ville de Dole. [...] Le bruit court en ville que deux autres des victimes auraient succombé aujourd'hui. Le temps me manque pour vérifier cette nouvelle. Ce qui est malheureusement vrai, c'est que M. le docteur Gremaud a déclaré que deux des malades sont dans un état à peu près désespéré.

Ch.IV (11), Dépêches de nuit, Etranger, « Les rapports du gouvernement français avec le Vatican », (01-08-1883)

Le bruit court que l'évêque d'Angers M. Freppel, aurait pris l'initiative de certaines démarches auprès du pape Léon XIII, afin de connaître très exactement ses intentions, relativement au gouvernement français. Plusieurs évêques, parmi lesquels on compte M. Guibert, archevêque de Paris, feignent de croire que le pape serait tout à fait résolu à une rupture avec le gouvernement de la République.

Ces rumeurs sont riches d'information sur l'imaginaire social, les angoisses, les ressentiments et les aspirations des contemporains du *Petit Comtois*. En ce qui

¹⁹ TLFi en ligne, consulté le 20-07-2009.

concerne les rumeurs présentées dans les brèves sur l'actualité politique nationale et internationale, elles sont exploitées en tant qu'événements en devenir. Les brèves rapportent alors une information strictement officieuse plus que non-vérifiée : il s'agit de dévoiler aux lecteurs les arrières-plans du politique. La fréquentation de notre corpus nous invite à élargir la définition de la rumeur et à mettre en lumière la dimension idéologique de la rumeur. Mettre en scène une information comme rumeur, c'est avant tout disqualifier le discours de l'autre, comme le dit très justement le sociologue E. Taïeb :

[...] la « rumeur » c'est aussi l'idéologie de l'autre, que les médias entendent démasquer comme telle et ne pas transformer en information légitime. Qualifier une position politique de « rumeur », c'est à la fois ne jamais lui permettre d'accéder au statut d'information, donc d'être crédible, et marginaliser la position qui a tenté de faire passer une telle information (Taïeb, 2006 : 137).

La dépêche suivante illustre parfaitement notre propos. Le locuteur-citant disqualifie ici explicitement les informations communiquées par « la presse ministérielle » : par l'utilisation de l'expression *répandre le bruit que*, d'un conditionnel et d'un commentaire des plus catégoriques où le discours du locuteur-cité est relégué au statut de « racontars », substantif axiologique des plus négatifs. L'information n'est pas ici posée comme « bruit » parce qu'elle n'est pas vérifiée, mais parce que le quotidien la juge formellement erronée. Il ne s'agit pas ici de prendre ses distances avec une information incertaine mais de dénoncer les visées politiciennes de la source diffusant la nouvelle :

Ch.IV (12), « Calomnies de la presse officieuse », (1883-08-19)

La presse ministérielle répand le bruit qu'il existe une société, au capital de 750,000 francs, ayant pour but de provoquer des *pronunciamientos*.

Cette société se serait acquis le concours de M. Zorilla, moyennant finances.

Il va sans dire que tout est absolument faux dans ces racontars.

Ce refus de l'information de l'autre est explicitée par le cotexte de la dépêche : ainsi, le bulletin du 19-08-1883, surmontant la rubrique des *Dépêches de Nuit*, dont est extrait la brève ci-dessus, accuse, dès les premières lignes du numéro du quotidien, le gouvernement espagnol de mener une opération de désinformation :

La proximité sur le plan des années chronologiquement continues 1884-1885-1886-1887 est relative à un ensemble d'items connaissant des profils distributionnels très marqués sur cette période. Les items *allemands*, *allemand*, *Allemagne*, *Berlin* participent tout particulièrement à la constitution de cette séquence chronologique. Si l'on examine les profils distributionnels de ces items, on remarquera la concentration de leur plus forts écarts-réduits sur la période 1885-1889, tandis que leur profil tend à être déficitaire sur les tranches diachroniques suivantes.

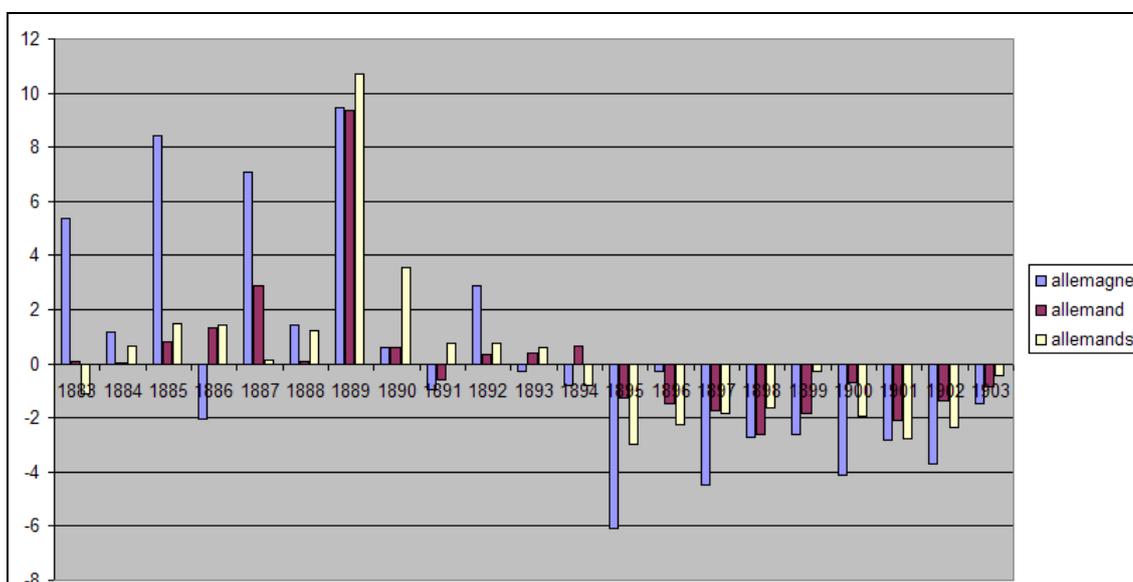


Figure IV-16 : Histogramme parallèle de la distribution des formes *Allemagne*, *allemand*, *allemands* (en écarts-réduits).

De 1885 à 1889, la question allemande est au centre des préoccupations du moment, comme le suggère la consultation des articles. Suite à un incident de frontière qui vient exalter le sentiment revanchard d'une France traumatisée et amputée par la défaite de 1871, la guerre semble imminente en 1887. L'expansion allemande, les déclarations de Bismarck, les rapports entre l'Allemagne et la Russie sont autant de sujets d'inquiétude et de polémique dont *Le Petit Comtois* se fait l'écho :

Ch.IV (14), Dépêches de nuit, « La situation extérieure », (25-01-1887)

Paris, 24 janvier. Il fallait s'y attendre. Les journaux anglais, qui répondent peut-être mieux encore que les journaux allemands aux désirs intimes de M. de Bismark, ont repris leur campagne de fausses nouvelles sur les relations de l'Allemagne avec la France. En arrivant aujourd'hui à la Chambre, nous avons appris que le *Daily News* publiait une information des plus alarmantes. D'après ce journal, il n'y aurait pas seulement des points noirs à l'horizon : l'orage serait tout formé et menacerait d'éclater à bref délai sur nos têtes. En style d'information, l'Allemagne s'apprêterait à demander des explications à la France sur les mouvements de troupes qui auraient lieu du côté de notre frontière de l'Est. On voit qu'il s'agirait purement et simplement d'une complication diplomatique pouvant aboutir à une déclaration de guerre. [...]

Ch.IV (15), « 1889 », (01-01-1889)

[...] En surexcitant sans scrupule le chauvinisme allemand, dans le but d'obtenir le vote des crédits nécessaires au développement exagéré de leur armement, les gouvernants de l'Allemagne ont certainement porté le coup le plus terrible à notre commerce et à notre industrie.

Ce contexte de tension a pour pendant le martèlement d'un certain nombre de formes lexicales, de fait présentes sur le zoom du vocabulaire spécifique de la séquence 1884-1890. Tandis que le conflit pressenti menace l'équilibre précaire de l'*Europe*, les appels à la *paix* se multiplient :

Ch.IV (16), Dépêches de nuit, « Au ministère de l'Intérieur », (01-01-1887)

[...] Nous voulons la paix, nous l'avons assez dit pour qu'il soit inutile de le répéter désormais, personne ne peut douter de notre sincérité et certainement personne n'en doute.

Ce qui est vrai, c'est que l'Europe tout entière vit sur le pied d'une sorte de paix armée et qu'un tel état de choses suffit à expliquer, à certains moments, les préoccupations de l'opinion publique. Cet état de chose existe; il ne dépend pas de nous de le faire cesser. Le gouvernement n'a aucune raison de croire que la guerre en doive sortir; il espère fermement qu'elle n'en sortira pas. Toute sa politique est dirigée dans cette vue. [...]

L'histogramme suivant témoigne des suremplois de la forme *paix* en 1887 (+6,63) et 1890 (+3,63).

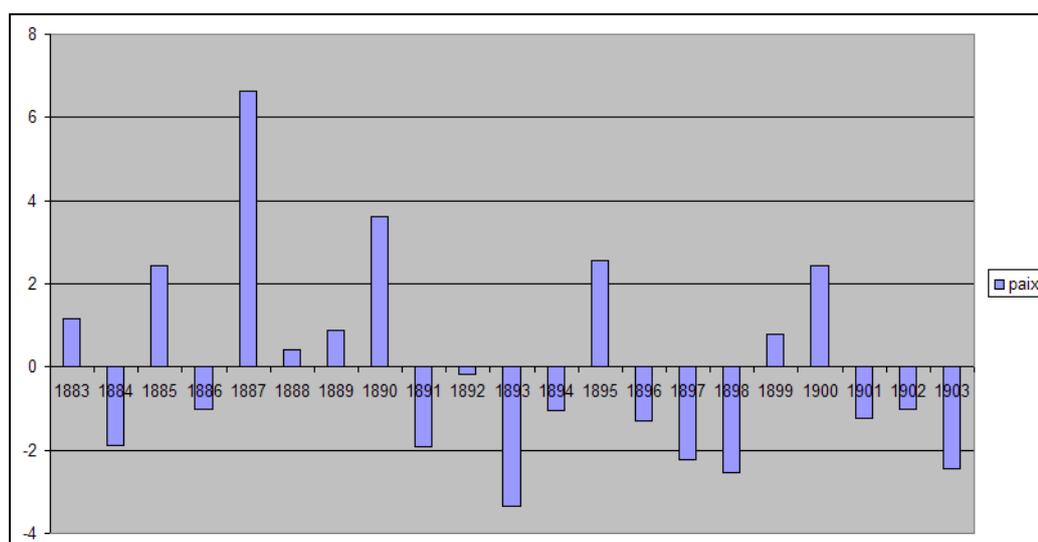


Figure IV-17 : Histogramme de la distribution de *paix* (en écarts-réduits).

Corollaire du climat de tension entre l'Allemagne et la France, la question du *budget du gouvernement*, et plus particulièrement de l'*intérieur*, s'inscrit comme une thématique centrale de cette période. On remarquera en effet les items se rapportant à la thématique des finances : *finances, budget, chiffre, millions, dépenses, francs*, qui constituent le cœur du vocabulaire spécifique de la séquence 1884-1890.

Ch.IV (17), Dépêches de nuit, « Le budget de 1887 », (1887-01-05)

Paris, 14 janvier. On avait paru assez préoccupé, jeudi soir, au sujet du vote de la Chambre sur le chapitre du budget de l'intérieur relatif aux dépenses secrètes de la sûreté générale. Les dispositions des députés radicaux semblaient peu favorables. On craignait des abstentions nombreuses qui auraient été de nature à réduire considérablement l'effectif des voix gouvernementales. Les droites, qui comptent 180 membres menaçaient, en s'adjoignant une quarantaine de députés d'extrême gauche, de constituer une majorité. Ces appréhensions n'étaient pas fondées. Les intentions des membres de l'extrême gauche et de la gauche radicale semblaient, au moins vingt-quatre heures plus tard, toutes différentes. La gauche radicale fera peut-être des réserves, mais elle votera, en définitive, les fonds secrets.

De 1884 à 1890, un des termes clés est également celui de *réforme*, item qui connaît des valeurs positives en 1884, en 1887, et son plus fort sur-emploi en 1887. L'heure est définitivement à la *réforme*, qu'elle ait pour objet les droits de

succession, le régime des boissons, mais surtout la réforme du budget de l'intérieur tandis qu'une guerre imminente est redoutée.

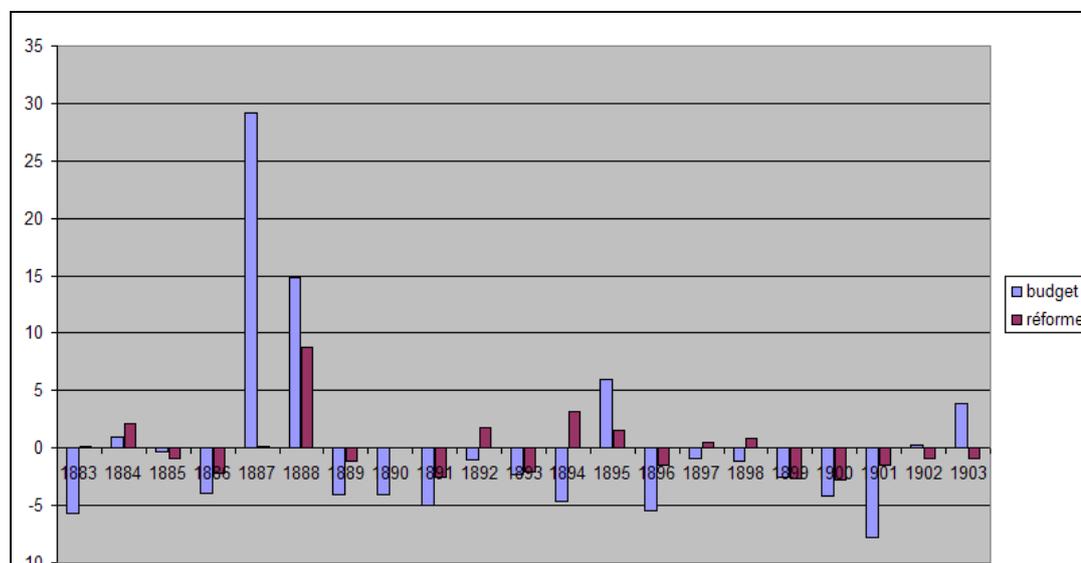


Figure IV-18 : Histogramme croisé des formes *budget* et *réforme* (en écarts-réduits).

Reflétant la période trouble que connaît alors la Troisième République, un autre item est au centre de la série chronologique 1886-1890 : celui de *boulangier*, nom propre qui renvoie au général du même nom. Sur l'histogramme suivant, un profil distributionnel très marqué sur cette période se donne à lire : on relève ainsi un emploi significativement positif successivement en 1886 (3,5), 1887 (4,03), jusqu'à culminer en 1889 à 45,08. Ces emplois donnent à lire la nomination du général Boulanger au poste de Ministre de la guerre en 1886, son implication dans le scandale des décorations en 1887, et ses succès d'élection au poste de député en 1889.

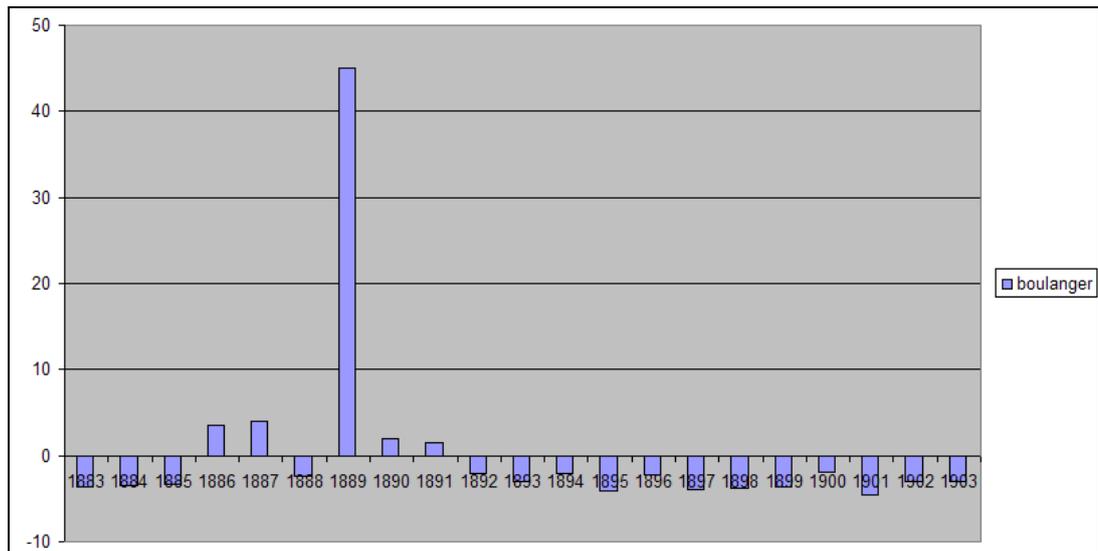


Figure IV-19 : Histogramme de la distribution de la forme *boulanger* (en écarts-réduits).

1891-1896

Les items relatifs au domaine juridique constituent la clé de voûte de la série 1891-1896. On repérera ainsi, rassemblés sur ce zoom, les items *procès*, *audience*, *témoins*, *témoin*, *parquet*, *interpellation*, *accusé*, *crime*, qui connaissent tous des écarts-réduits significativement positifs sur les années successives de cette série.

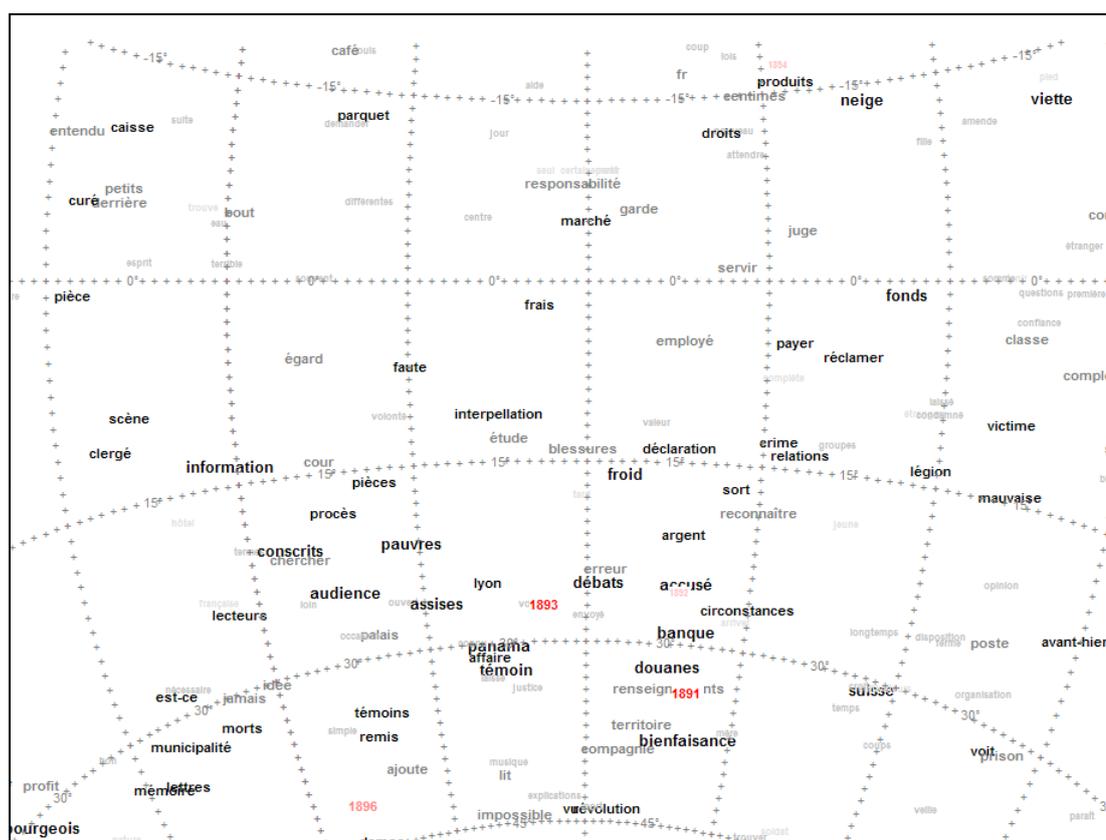


Figure IV-20 : Zoom sur la région 1891-1896.

De la plus célèbre, celle du Panama, à la plus anecdotique, en passant par l'affaire Fouroux et l'affaire Padlewski en 1891 ou encore l'affaire des 104 en 1896, l'« affaire » est à la Une de cette séquence chronologique :

Ch.IV (18), « L'affaire Fouroux », (1891-01-08)

Au moment où s'ouvrent, devant la cour d'assises du Var, à Draguignan, les débats du procès Fouroux, nous allons résumer en quelques lignes cette affaire qui eut, il y a quelques mois, à cause des personnages qui y figurent, un si grand retentissement. On se souvient que M. Fouroux, officier de marine démissionnaire et maire de Toulon, avait depuis longtemps pour maîtresse la femme d'un officier de marine, Mme de Jonquières ; leur intimité faisait même beaucoup jaser la ville.

Ch.IV (19), Dernière heure, « L'affaire Lebaudy », (1896-01-07)

M. Bernard, commissaire aux délégations judiciaires, chargé de l'instruction de l'affaire de chantage Lebaudy, a entendu MM. Jacques et Robert Lebaudy et M. de Certi. Après conférence avec M. Bertrand, procureur général, et M. Meyer, juge d'instruction, désigné pour examiner l'affaire, on a procédé à l'arrestation de

M. de Certi, qui a été conduit à Mazas. Dans la soirée de demain, d'autres personnes seront entendues.

Ch.IV (20), Chronique régionale, Jura, « Une grave affaire », (1891-01-30)

M. le procureur général près la cour d'appel de Besançon, après avoir donné l'ordre d'arrêter le sieur Verdène, garde forestier domanial à la Châtelaine, est arrivé inopinément à Arbois, accompagné d'un conseiller à la cour. Ces messieurs vont procéder eux-mêmes à l'instruction de cette affaire. On parle d'importants détournements.

On aura remarqué la proximité de l'item *panama* avec *affaire*, et sa correspondance avec l'année 1893. Comme l'indique l'histogramme suivant, la forme réalise un écart-réduit de plus de 50 en 1893, date à laquelle le scandale du Panama, rendu publique en 1892, est au centre de tous les débats. On notera d'ailleurs sur l'histogramme ci-dessous les sur-emplois de cette forme, majoritairement concentrés sur la période 1888-1896 (5,2 en 1888 ; 2,83 en 1891 ; 4,07 en 1893 ; 5,44 en 1896).

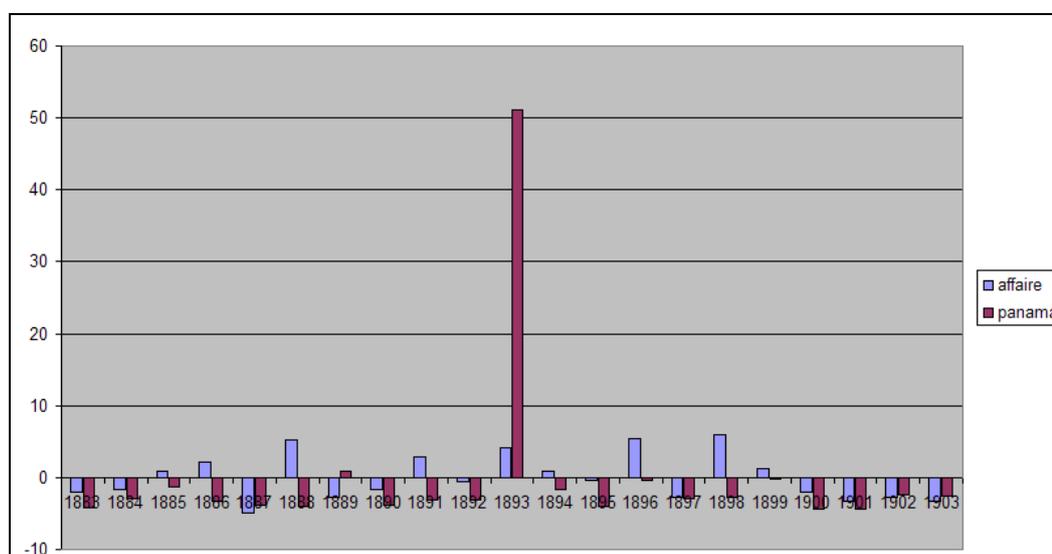


Figure IV-21 : Histogramme parallèle des formes *affaire* et *panama* (en écarts-réduits).

Les items *argent*, *fonds*, *banque*, ayant pour corolaire les items *payer* et *réclamer*, semblent autant d'indices des scandales financiers de cette période, à l'instar de l'affaire du Panama. On remarquera sur l'histogramme suivant les

profils distributionnels des items *argent* et *banque*, positivement marqués sur la période 1890-1893.

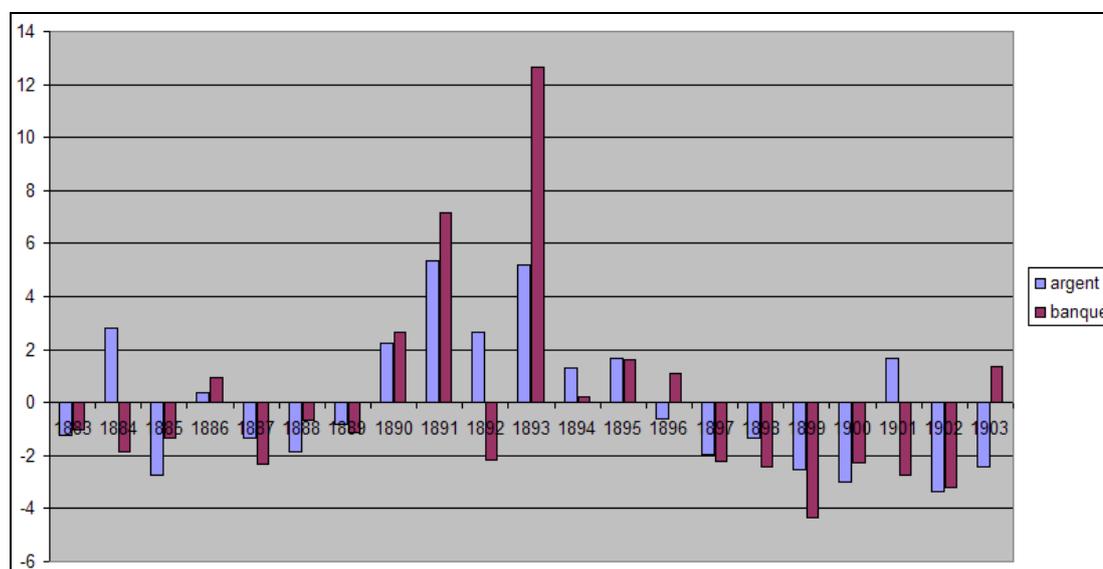


Figure IV-22 : Histogramme parallèle des formes *argent* et *banque* (en écarts-réduits).

Si un certain nombre d'occurrences de ces items s'inscrivent dans le cadre d'articles consacrés à l'affaire du Panama, ils relèvent, plus généralement, de divers scandales financiers, nationaux et internationaux :

Ch.IV (21), Dépêches de nuit, « Autour du Panama », (1893-01-17)

On lit dans le *Petit Lyonnais* : Il est si peu vrai, dit l'*Echo de Paris*, que l'instruction relative à M. Baihaut soit sur le point d'être close, que M. Franqueville a envoyé hier, aux parquets de Vesoul, de Gray et de Lure, des commissions rogatoires, à l'effet de rechercher, dans les différentes maisons de banque du département de la Haute-Saône, les opérations financières que M. Baihaut a pu faire dans la région de l'Est. Jusqu'à présent, M. Baihaut reconnaît avoir reçu de la compagnie de Panama, à divers titres, une somme de 375,000 fr. Or, M. Franqueville a eu en mains la preuve que le député de la Haute-Saône a touché plus d'un million. On retrouve les traces des 373,000 fr., avoués par M. Baihaut, au Crédit lyonnais et à l'ancien Comptoir d'escompte ; on recherche maintenant la différence entre la somme avouée et la somme réellement touchée. De là les commissions rogatoires adressées aux parquets de Gray, de Vesoul et de Lure.

Ch.IV (22), Dépêches de nuit, Extérieur, Italie, « Les scandales de la Banque romaine », (1896-01-24)

La chambre des mises en accusation, chargée d'examiner l'affaire de soustraction de documents dans le procès de la Banque romaine, a déclaré, conformément à

la demande du ministère public, qu'il n'y avait pas lieu de poursuivre MM. Felzani, ancien questeur [...]

Enfin, on notera, disséminés sur ce zoom, la présence des items *clergé* et *curé*. Si l'on considère l'histogramme des profils distributionnels de ces items, on remarquera que les emplois significativement positifs de la forme *clergé* se concentrent en 1890 (3,02), 1891 (2,04) et que la forme connaît son plus fort suremploi en 1892 (6,83). Quant à la forme *curé*, ses emplois significativement positifs sont à relever en 1892, 1895, 1896 et 1899.

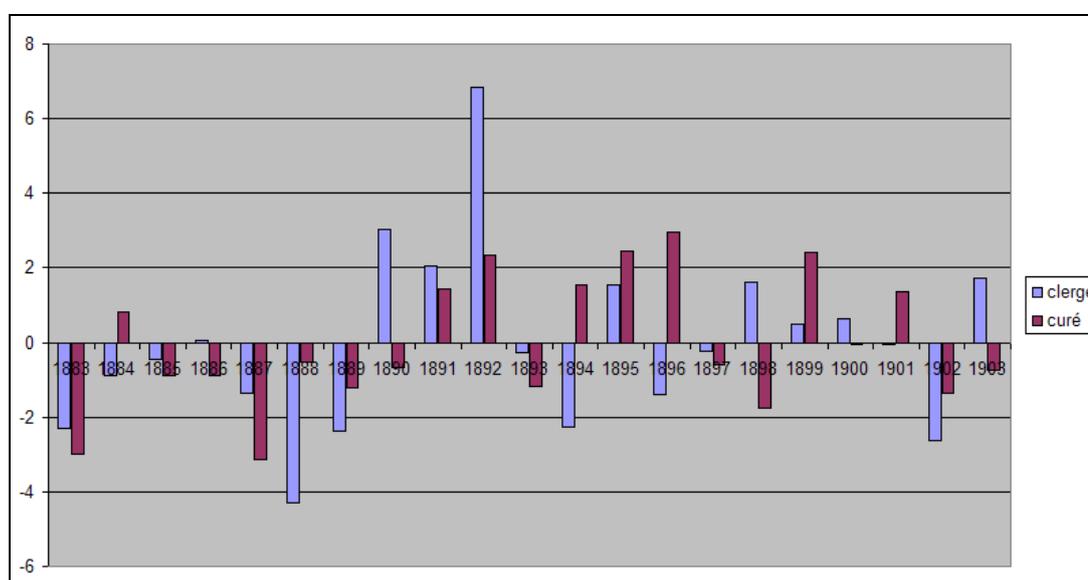


Figure IV-23 : Histogramme parallèle des formes *clergé* et *curé* (en écarts-réduits).

Le sur-emploi de ces formes pendant la période 1890-1896 pointe l'intensification d'un climat anticlérical, comme on peut le constater dans les exemples suivants :

Ch.IV (23), « Une nouvelle encyclique », (1890-01-18)

Léon XIII vient d'adresser au clergé catholique une encyclique dans laquelle il s'occupe, entre autres choses, des devoirs et des droits des catholiques comme des citoyens. Cette encyclique vient fort à propos, c'est-à-dire au moment même où certains membres du clergé français partent en guerre contre le gouvernement de la République et rééditent contre lui toutes les accusations idiotes et injustes qui traînent dans les colonnes des feuilles cléricales. Il n'est donc pas sans intérêt de connaître ce que le chef de l'Eglise pense de la conduite

que le clergé doit tenir vis-à-vis du pouvoir civil. Léon XIII commence - n'est-ce pas le cliché obligatoire de tout écrit clérical ?- par déplorer la situation du monde chrétien, les persécutions dont l'Eglise est l'objet. Tous ces maux, suivant lui, viennent de l'oubli des principes de la sagesse chrétienne et de la perte de la foi. On pourrait répondre au pape que si la foi s'éteint de plus en plus dans les âmes, ce sont les membres du clergé qui en sont la cause, car ils font tout ce qu'il faut pour amener un pareil résultat, quand ils se mêlent aux choses de la politique, qu'ils se jettent dans la mêlée des partis comme aux dernières élections et qu'ils prêchent la haine des pouvoirs établis que l'Évangile leur commande de respecter. [...]

Ch.IV (24), Chronique régionale, Jura, « Campagne maçonnique », (1890-01-20)

Ceux qui aiment la douce gaîté pourront s'en donner à cœur joie en lisant le *Courrier du Jura* de cette semaine. Lâchant d'un cran la polémique cléricale dont elle ne sait plus que dire, la feuille jésuitique s'occupe des scandales soulevés par le clergé à Pontoise, à Grenoble, et dans mille autres lieux, et elle appelle cela : « Campagne maçonnique. » Les faits sont dans toutes les mémoires :

Un curé refusant de recevoir le corps d'un maire, malgré les supplications d'une veuve, ardente cléricale, parce que ce maire était franc-maçon. Un évêque interdisant un prêtre parce qu'il avait approuvé les lois concernant l'instruction et le service militaire ; Un autre évêque défendant qu'on enterre religieusement un préfet, - encore malgré les prières de la famille, - sous le prétexte, d'ailleurs imaginaire, que les francs-maçons de Grenoble avaient résolu d'assister officiellement en corps et avec les insignes à cet enterrement d'un fonctionnaire républicain, etc., etc. Voilà ce que le *Courrier du Jura* appelle : une campagne maçonnique.

Ch.IV (25), Dépêches de nuit, Intérieur, (1891-01-06)

Les journaux conservateurs annoncent qu'aux obsèques du frère Iblasius, directeur des écoles chrétiennes de Bordeaux, assistaient, outre le maire de la ville, le préfet, les généraux, des colonels, des magistrats, des avocats, des officiers et dix mille personnes environ.

Plusieurs discours ont été prononcés. On voit bien que nous vivons dans un temps de persécution contre le clergé et les congrégations !

L'item *droits* connaît également un profil distributionnel marqué, en réalisant ses seules valeurs positives d'écart réduit en 1888, 1894 et 1898. Comme en témoignent les exemples suivants, la forme lexicale renvoie notamment aux *droits de succession* et aux *droits de douanes*, qu'il faut relier au débat sur le protectionnisme :

Ch.IV (26), « Le budget et les droits de succession », (1888-01-06)

Au point où nous en sommes, il ne faut pas s'attendre à voir encore introduire dans le budget, pour l'exercice 1888, de ces réformes fiscales qui comptent dans le système financier d'un pays.

Cependant, il n'y aurait rien d'étonnant à ce que la Chambre votât, dès cette année, quelque modification dans le régime appliqué aux boissons et peut-être aussi quelques combinaisons nouvelles dans le mode et le taux de perception des droits de mutation en matière de succession. La commission du budget a précisément adopté, en effet, une proposition d'un de ses membres, mort il y a quelques jours à peine. M. Duché, qui devait procurer des ressources non prévues au budget des recettes présenté par le gouvernement. Cette proposition avait pour principal objet de donner aux droits de succession un caractère progressif plus accentué que celui, établi par la loi de frimaire an VII ; de là une élévation sensible du tarif des droits ainsi que de leur rendement, comme de juste.

D'autre part, comme compensation offerte aux héritiers futurs appelés à tomber sous le coup de ces tarifs nouveaux, on déduirait les dettes du défunt du montant des valeurs de la succession soumises aux droits.

Ch.IV (27), Dépêches de nuit, Extérieur, Russie, « A MM. les protectionnistes », (1894-02-06)

Le *Nouveau Temps* publie un article sur la question des blés. Il dit que le prélèvement des droits des blés en France causera un sensible dommage aux intérêts de l'agriculture et du commerce russes. Il énumère ensuite les nombreux avantages que la France retire de la convention commerciale de 1893 avec la Russie, avantages qui seront encore accrus à la conclusion du traité russo-allemand, tandis que, pour la Russie, l'unique avantage relatif au pétrole perd sa signification quand les conditions de l'exportation de son principal produit, le blé, deviendront difficiles. En conséquence, le maintien de la convention de 1893 doit être considéré par la Russie comme désavantageux. [...]

Note du journal. — La fureur protectionniste, après nous avoir aliéné la Suisse, ira-t-elle jusqu'à tourner contre nous la Russie, notre alliée naturelle ? Ce serait un comble dont M. Méline et ses amis devraient peut-être compte au pays le jour où ils seraient parvenus à faire augmenter le prix du pain.

1897-1898-1899-1901

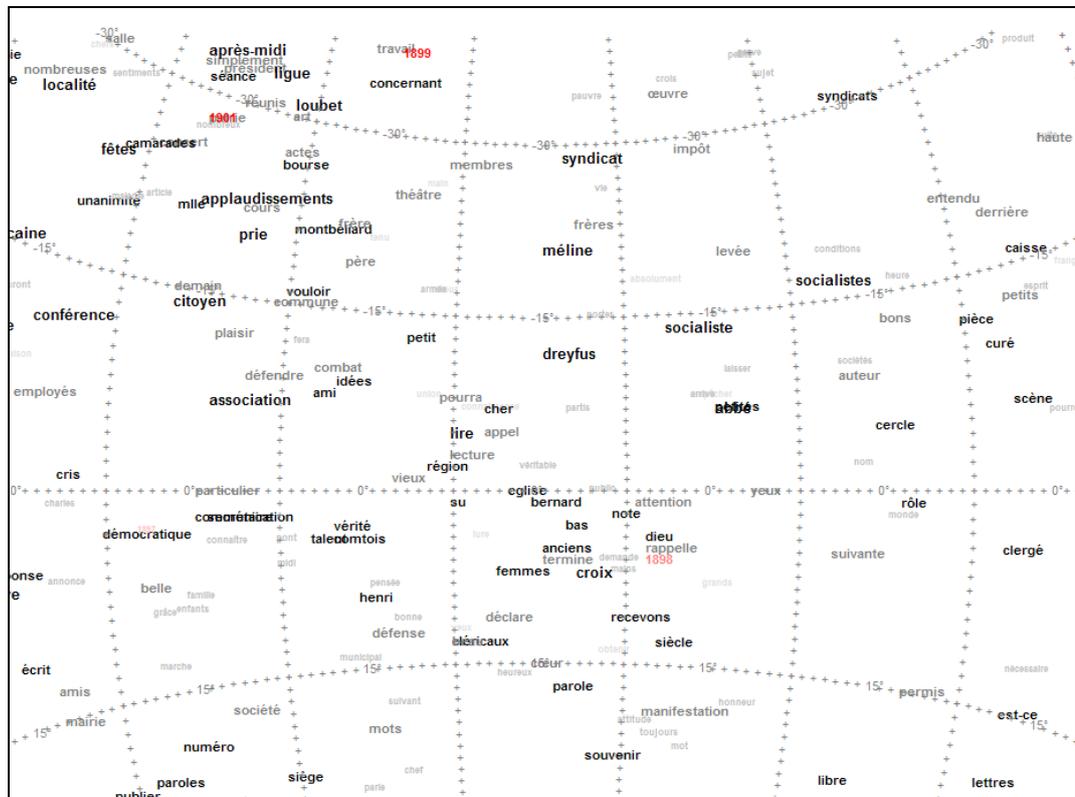


Figure IV-24 : Zoom sur la région 1897-1898-1899-1901.

Sur ce zoom des années 1898-1901-1903, on relèvera tout d'abord les noms des « hommes du jour » (Angenot, 1989 : 633) de cette séquence chronologique, à savoir Dreyfus, Méline, Loubet, et à travers eux, l'actualité de cette période. Cette actualité, c'est tout d'abord le dossier Dreyfus, qui devient une *affaire* en janvier 1898 à la suite de la parution de la lettre ouverte d'E. Zola au président de la République. C'est aussi la lutte électorale entre Jules Méline et Émile Loubet, l'élection de ce dernier au poste de Président de la République par le Congrès réuni à Versailles le 18 février 1899.

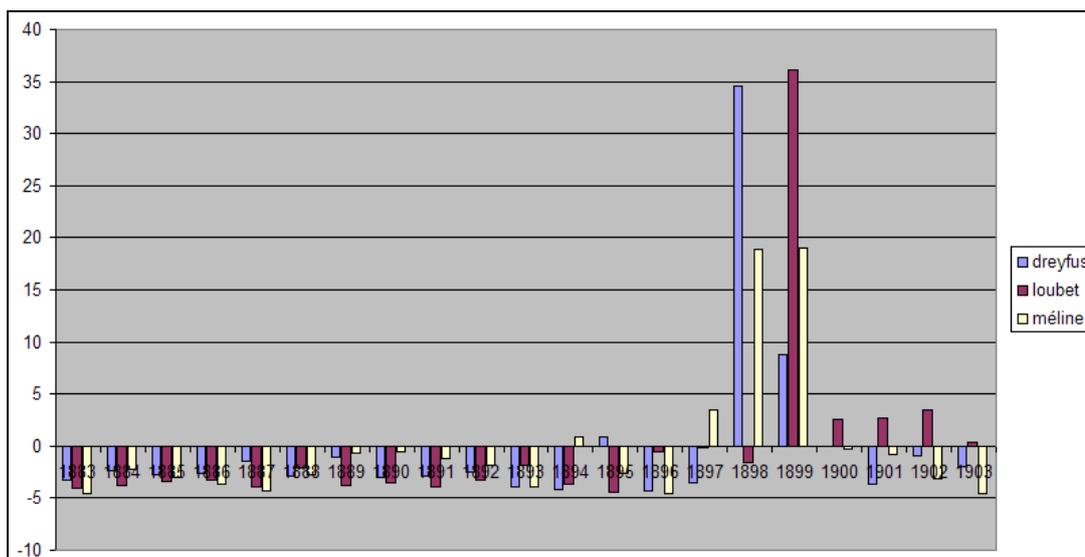


Figure IV-25 : Histogramme des profils distributionnels des formes *dreyfus*, *loubet*, *meline*.

On observera sur la figure IV-24 la présence de deux items d'une même forme déclinée au singulier et au pluriel : *socialiste* et *socialistes*. Si l'on veut rendre compte du fonctionnement de ces items grâce à l'histogramme suivant, on notera que ceux-ci connaissent sur la première moitié diachronique un profil déficitaire, mis à part en 1893 où ils sont employés de façon positive significative en 1893. Après une période d'accalmie, on assiste à une résurgence extensive de leur utilisation à partir de 1897. Eclairée par le contexte d'apparition de ces formes, l'interprétation de leur profil distributionnel est à relier à l'essor du mouvement socialiste, devenant une réelle force parlementaire lors de l'élection de 50 députés socialistes aux élections législatives en 1893, entraînant l'essor de groupes socialistes locaux en vue de la préparation des élections législatives de 1897 et 1899, sur lequel la presse républicaine ne manque pas de revenir.

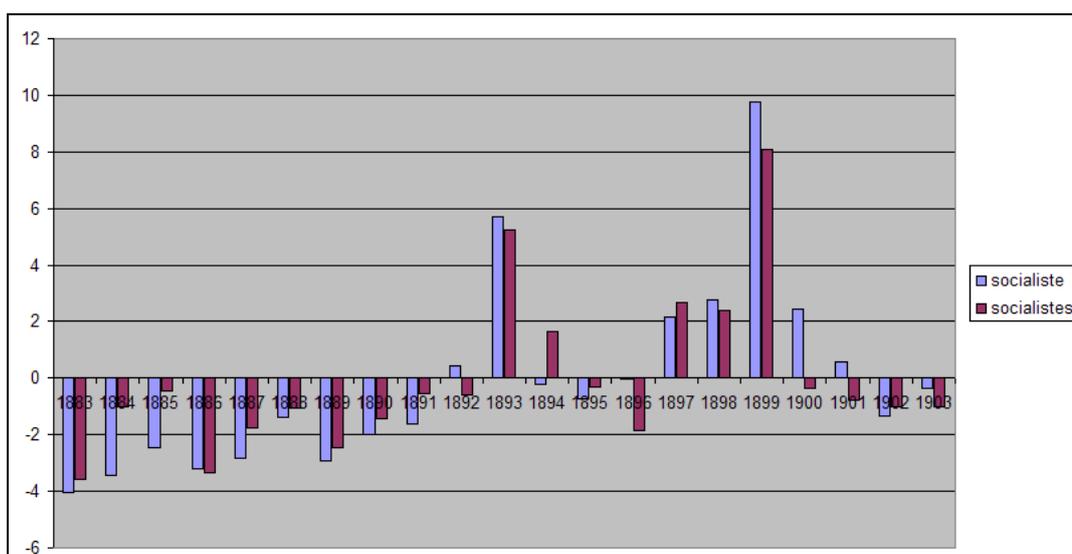


Figure IV-26 : Histogramme des formes *socialiste* et *socialistes* (en écarts-réduits).

De l'échelle nationale et internationale à l'échelle locale, des dépêches aux chroniques régionales, la progression des « socialistes » s'impose comme un événement majeur de la deuxième partie diachronique de notre corpus, comme on peut le constater à la lecture des exemples suivants :

Ch.IV (28), Dépêches de nuit, « Les socialistes révolutionnaires », (1893-01-04)

Le gouvernement s'est préoccupé des réunions tenues par les socialistes révolutionnaires. Aucune organisation révolutionnaire ne sera tolérée ; tous les attroupements sur la voie publique, le jour de la rentrée des Chambres, seront dispersés, tant à Paris que dans les départements.

Ch.IV (29), Dépêches de nuit, « Politique d'apaisement », (1899-02-06)

Marseille. - On télégraphie que la ligue des patriotes avait organisé, ce matin, une réunion dans laquelle devait se faire entendre MM. Marcel Habert et de Magallon, orateur des ligueurs marseillais, et Pourcin, délégué d'un groupe socialiste parisien. Environ 2.500 personnes étaient à l'intérieur du théâtre de l'Athambra, où la réunion avait lieu. Aux alentours, toute la police était sur pied. L'entrée s'est effectuée avec assez de calme ; mais à l'extérieur l'agitation était extrême, et dans les groupes, on discutait avec la dernière violence.

On notera que la question socialiste donne même lieu à la création d'une *Chronique socialiste* dans *Le Petit Comtois* à partir de 1893 :

Ch.IV (30), Chronique socialiste, (1893-01-30)

[...] Le jour où la bourgeoisie, - nous parlons de celle qui est sincère, - se sera fait ce raisonnement à elle-même, le procès des socialistes sera bien près d'être gagné, au moins pour partie. En attendant, notre bourgeoisie à nous s'effarouche volontiers des apostolats trop virulents, et c'est pourquoi, dans ces Chroniques socialistes, nous aimerions un peu plus d'exposition et moins de discussion. C'est là, d'ailleurs, une simple observation que nous soumettons à notre collaborateur, sans prétendre aucunement restreindre sa liberté de penser et d'écrire.

Un autre substantif décliné au singulier et au pluriel relève du vocabulaire propre de cette série chronologique : *syndicat* - *syndicats*. Si le pluriel du substantif connaît quelques emplois significativement positifs sur la première moitié diachronique du corpus (en 1886, 1893, 1894), le singulier du substantif ne s'avère quant à lui que plus tardivement utilisé dans des mesures significatives, puisque présentant un profil très nettement déficitaire dans la première moitié diachronique. La forme *syndicat* ne monte en effet qu'à partir de 1894, pour enfin être en suremplei dans les années 1898, 1901, 1903.

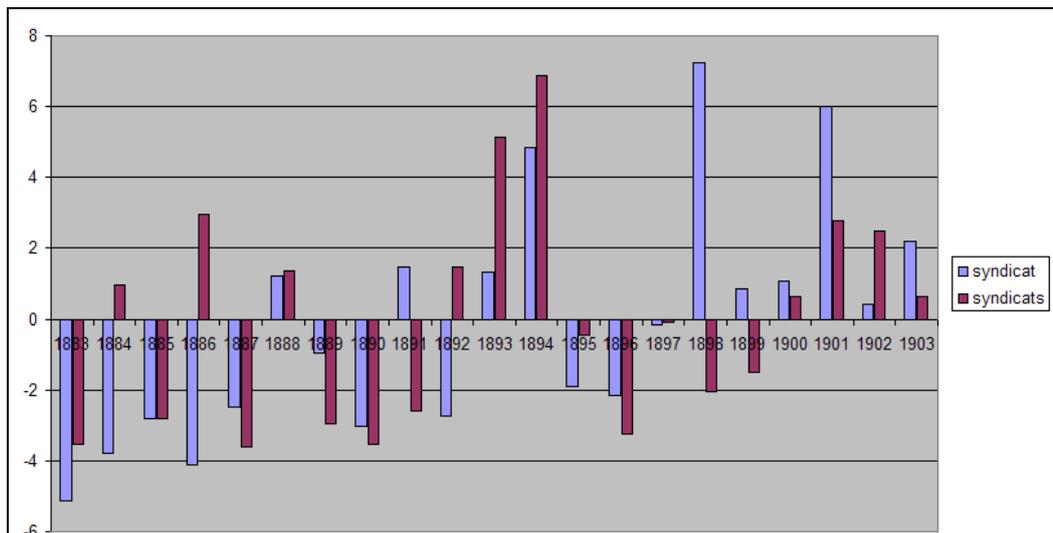


Figure IV-27 : Histogramme parallèle des formes *syndicat* et *syndicats*.

Les occurrences de la forme *syndicat* en 1898 s'inscrivent majoritairement dans le cadre d'articles visant à informer de la formation de ceux-ci, ainsi que l'atteste l'exemple suivant :

Ch.IV (31), Chronique régionale, Doubs, Montbéliard, « Le mouvement syndical », (1898-02-05)

Blamont. Pour faire suite aux adhésions que nous avons déjà fait connaître, ajoutons que les anciens horlogers sur cylindres de Blamont et des environs déclarent qu'ils enverront des délégués à la réunion du 6 février, à Maîche, en vue de la formation d'un syndicat.

Frambouhans. Il en est de même à Frambouhans, qui a formé un syndicat local, et qui envoie aussi des délégués à Maîche.

À proximité de l'année 1898, on relèvera la présence des items relatifs au domaine de la religion : *dieu, eglise, cléricaux*.

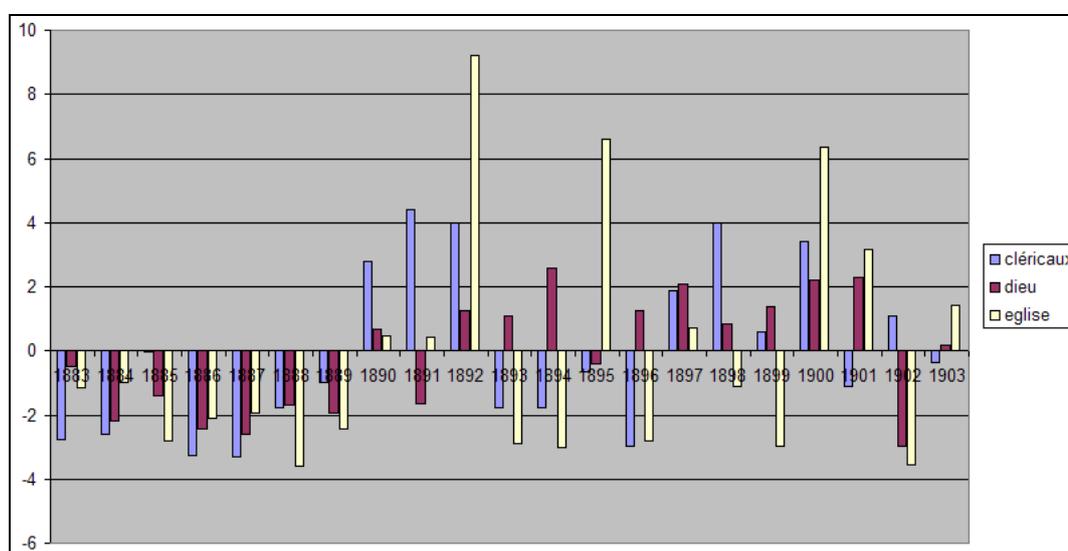


Figure IV-28 : Histogramme parallèle des profils distributionnels des formes *cléricaux*, *dieu*, *eglise* (en écarts-réduits).

Si l'on examine le profil distributionnel de ces items, on constatera que ceux-ci connaissent des profils déficitaires jusqu'en 1889. La forme *cléricaux* est significativement employée successivement en 1890, 1891 où elle réalise son plus fort écart-réduit, en 1892, 1898 et 1900. La forme *dieu* est quant à elle significativement employée en 1894, 1897, 1900 et 1901, tandis que la forme *eglise* est très massivement utilisée en 1892 (9,27), en 1895, 1900 et 1901.

On reviendra enfin sur l'item *association* présent sur le zoom centré du vocabulaire propre de la séquence considérée. Si cette forme apparaît

largement déficitaire durant la première partie diachronique du corpus, la deuxième période du corpus, à partir de 1891, s'annonce sous des auspices différents : on relève de forts sur-emplois en 1897, 1900, 1901 et 1903.

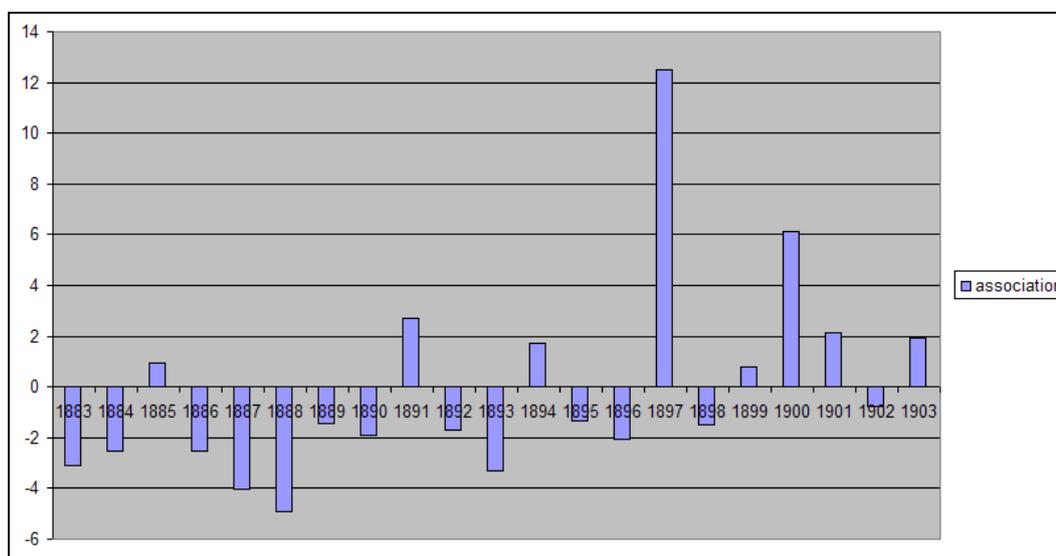


Figure IV-29 : Histogramme du profil distributionnel de la forme *association* (en écarts-réduits).

1900-1902-1903

Le vocabulaire propre des années 1900, 1902, 1903 sera observé ici à l'aide deux zooms centrés : l'un représentant la région 1900-1902, l'autre représentant la région 1903, située au nord de la zone précédente du fait des contributions du facteur 3.

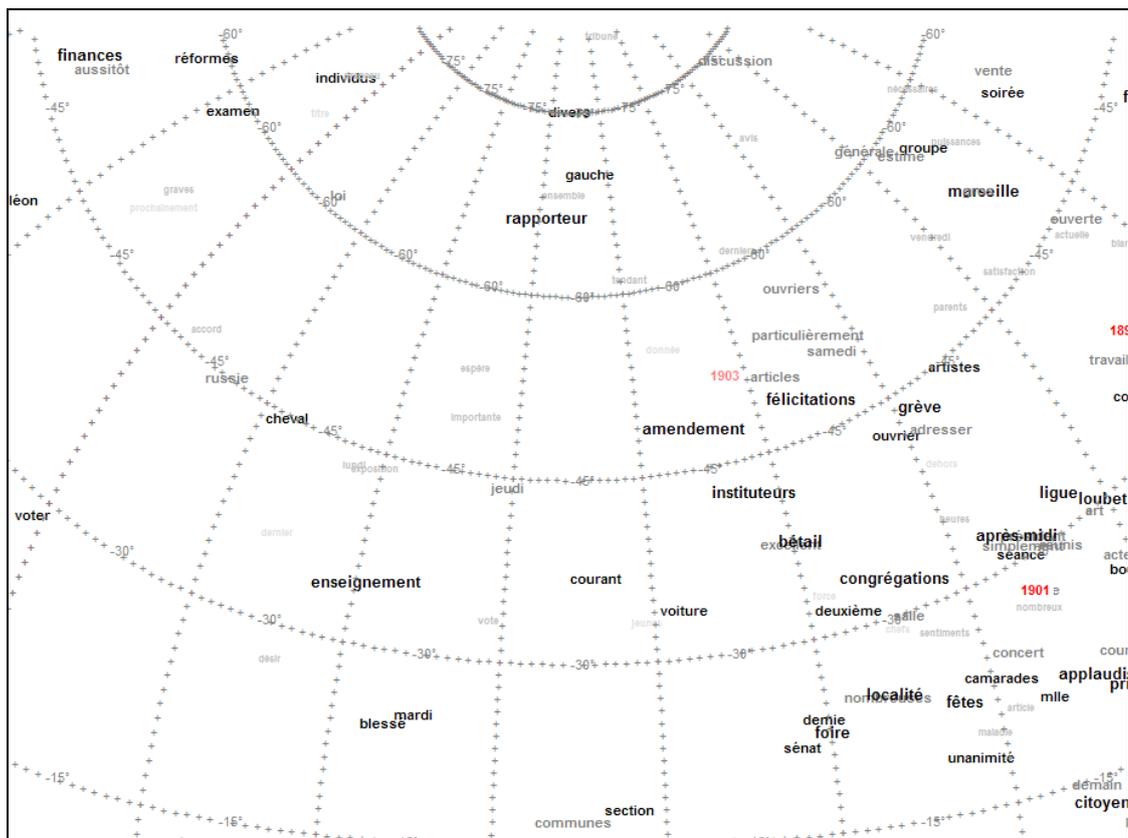


Figure IV-31 : Zoom sur la région 1903.

On remarquera tout d'abord l'item *boers* à proximité de l'année 1900 sur la figure IV-30. Le profil distributionnel de cet item, très nettement déficitaire jusqu'en 1899, connaît entre 1900 et 1902 un fort sur-emploi, comme on peut le constater sur l'histogramme ci-dessous. Les années 1900, 1901, 1902 voient en effet se dérouler la guerre anglo-boer (1899-1902), élément structurant de l'actualité de cette période.

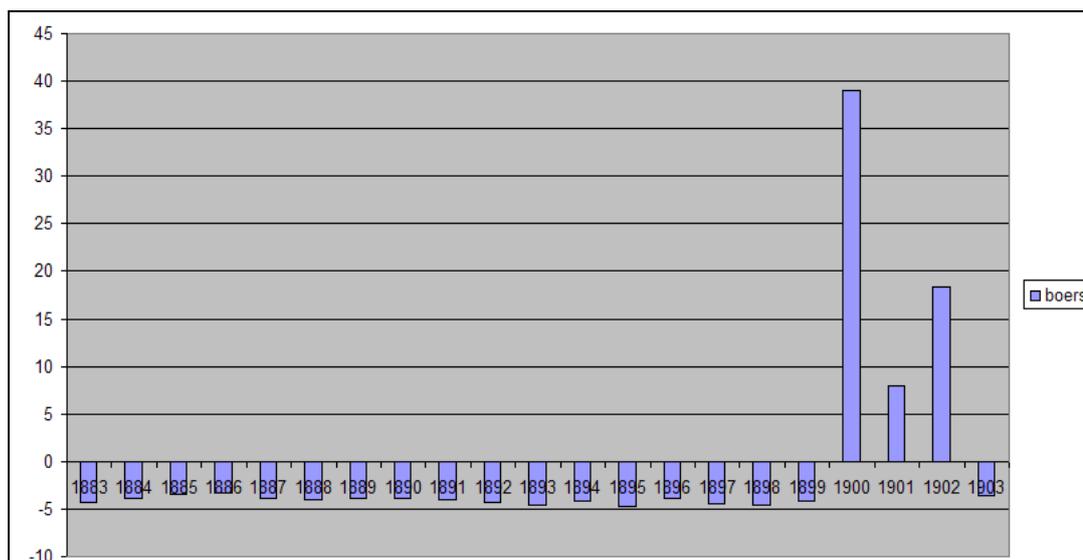


Figure IV-32 : Histogramme du profil distributionnel de la forme *boers* (en écarts-réduits).

Plusieurs formes relatives au domaine de l'enseignement figurent également sur ces zooms de la période 1900-1903 : *enseignement*, *instituteurs*, *élèves*, dont les profils distributionnels sont représentés ci-dessous.

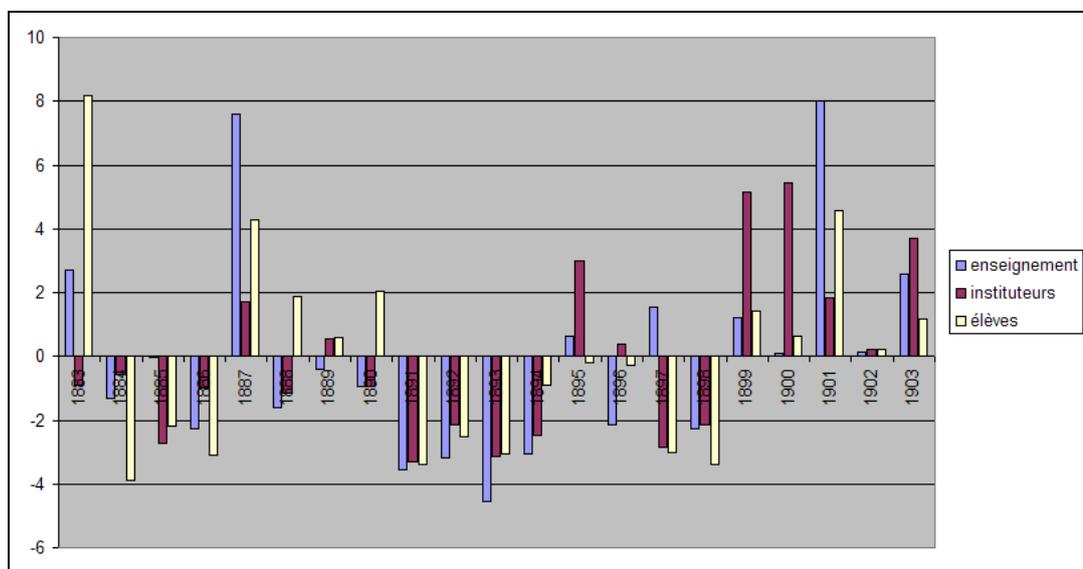


Figure IV-33 : Histogramme parallèle du profil distributionnel des formes *instituteurs*, *enseignement*, *élèves* (en écarts-réduits).

Dans la première moitié du corpus, la forme *enseignement* est employée de façon significativement positive en 1883 et 1887, pour ensuite connaître un profil très largement déficitaire de 1888 à 1896. Elle n'est à nouveau employée dans des mesures significatives qu'en 1901 et 1903. Les emplois positifs de la forme *instituteurs* sont quant à eux à trouver dans la seconde tranche diachronique du corpus, à savoir en 1895, 1899, 1900 et 1903. La forme *élève*, enfin, connaît son plus fort écart-réduit en 1883, puis témoigne d'un profil distributionnel déficitaire, à l'exception des années 1897, 1890 et 1901. La fortune de ces items traduit le débat sur l'enseignement sous la Troisième république, et son corolaire, le combat farouche mené par les anticléricaux contre les *congrégations*. L'histogramme suivant retrace l'essor du débat sur la question des congrégations en 1900 et du vote de la loi du 1^{er} juillet 1901 stipulant qu'aucune congrégation religieuse ne pourra désormais exister sans autorisation, loi qui résulte de longs débats au sein de la chambre et de la fermeture des écoles congréganistes en 1902 et qui reste d'actualité en 1903 :

Ch.IV (32), « La situation », (1901-03-28)

Ou bien que les membres des congrégations non autorisées ne pourront pas enseigner, ou bien - et je porterais volontiers le respectueux défi à mes adversaires de l'oser faire- il faudra inscrire ces paroles dans un texte de loi : «Art, 14. - Les membres des congrégations non autorisées pourront donner l'enseignement à la jeunesse. » (Vifs applaudissements à gauche et à l'extrême gauche.)

Présentée sous cette forme, la question est par là même dépouillée d'artifice. Il ne s'agit pas de savoir, entendez-le bien, si les catholiques ont le droit d'élever leurs enfants suivant leur conscience, si des laïques catholiques peuvent ouvrir des écoles : oui, ils le peuvent, et vous le savez. Il ne s'agit pas de savoir si les prêtres séculiers peuvent ouvrir des établissements d'instruction : - Ils le peuvent, et vous le savez. Il ne s'agit même pas de savoir si des congrégations autorisées qui se seront inclinées devant la loi pourront donner l'enseignement : - elles le peuvent, et vous le savez.

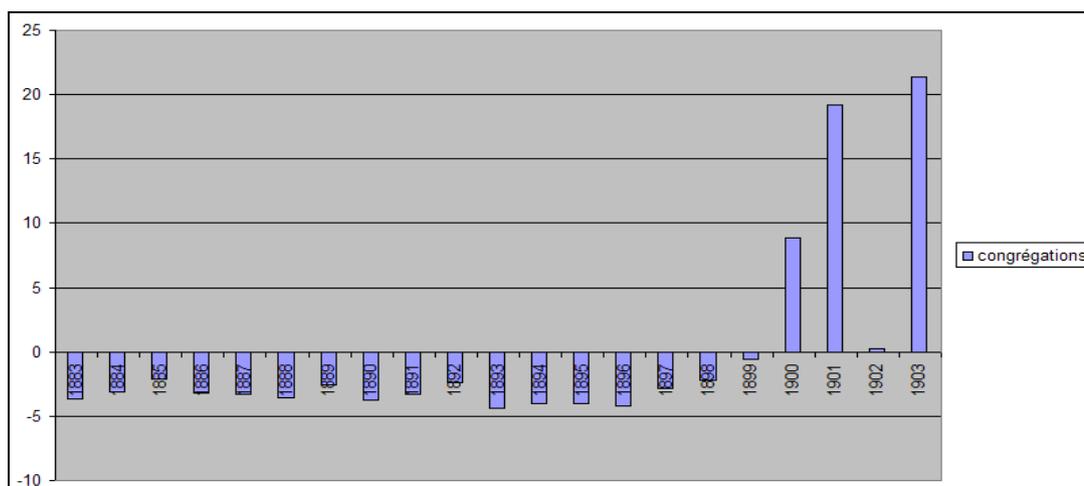


Figure IV-34 : Histogramme du profil distributionnel de la forme *congrégations* (en écarts-réduits).

Sur la figure IV-31, on trouvera enfin à proximité de l'année 1903 les items *ouvrier* et *grève*. Si l'on considère l'histogramme du profil distributionnel de la forme *grève*, on notera que son emploi est très majoritairement déficitaire dans la première partie diachronique du corpus (à l'exception de l'année 1890). La forme est en très net sur-emploi dans les trois dernières années du corpus, et tout spécialement en 1901, où elle réalise son plus fort écart-réduit.

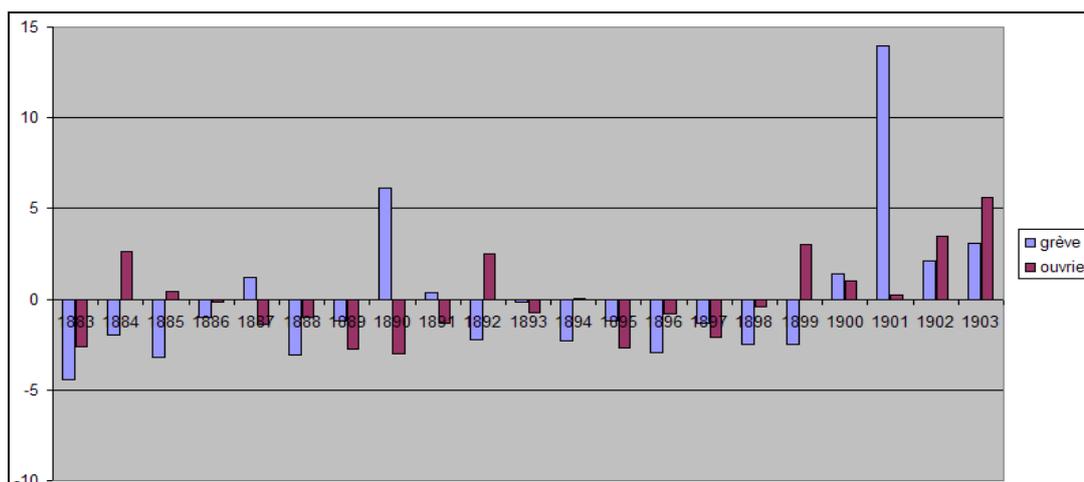


Figure IV-35 : Histogramme parallèle du profil distributionnel des formes *grève* et *ouvrier* (en écarts-réduits).

La consultation du contexte des occurrences de cette forme pointe que ces sur-emplois doivent être interprétés comme la trace discursive d'une série de grèves qui se déroulent en 1901 et en mars 1903. Elles ont lieu au niveau national : c'est par exemple la grève dite de Marseille en 1901, la grève des ouvriers journaliers des docks en 1903. Il s'agit également de grèves ayant cours au niveau international (grève de Prymoon en 1901, grève de l'antracite à New-York en 1903) et au niveau régional (grève des usines de Magnoncourt en 1903). Quant à la forme *ouvrier*, elle tend également à être sous-employée jusqu'en 1898 (à l'exception de 1884 et 1892) et s'avère fortement employée en 1902, et tout spécialement en 1903.

Ch.IV (33), Dépêches de nuit, Extérieur, Etats-Unis, « L'agitation gréviste », (1903-03-27)

New-York, 25 mars. — De toutes parts, on signale des troubles ouvriers. Il n'est pas douteux que les décisions de la commission sur la grève de l'antracite n'aient pour résultat d'encourager les grèves dans d'autres industries, puisque cette commission, tout en flétrissant les procédés du syndicat des mineurs, a accordé, en réalité, tout ce que demandait le syndicat.

Il est probable que la grève générale des ouvriers tisseurs dans le Massachusetts sera proclamée lundi prochain. Cette grève portera, dans la seule ville de Lowell, sur 18,000 ouvriers.

Ch.IV (34), Dépêches de nuit, Intérieur, « La grève du Havre », (1903-03-16)

Le Havre, 14 mars. La grève des ouvriers journaliers des docks et entrepôts continue. Aucune défaillance n'a été constatée ce matin dans les services de la voilerie, de la surveillance et du matériel des docks. Les grévistes sont calmes. Aucun incident.

Cet après-midi, une délégation des ouvriers des docks est allée demander au maire de provoquer une conférence avec le sous-préfet, pour s'expliquer sur les causes de la grève. La salle Francklin, ancienne Bourse du travail, avait été demandée ce soir par les camionneurs et ouvriers des docks. Le maire s'était presque engagé à l'accorder, quand des affiches, placardées par la Fédération, ont modifié sa décision.

Ces affiches annoncent que, si satisfaction n'est pas donnée aux ouvriers des docks, la grève générale sera déclarée dans le port et que cette grève comprendra les ouvriers journaliers du port, voiliers, charbonniers et ouvriers des magasins publics, syndiqués et non syndiqués. [...]

4.2.3. Remarques conclusives sur l'examen de la ventilation du vocabulaire en diachronie

Ce premier volet d'exploration, qui s'attache à observer la ventilation du vocabulaire en diachronie nous a permis de prendre connaissance des premières spécificités de notre corpus.

Nous avons pu ainsi noter la très nette division de notre corpus en deux tranches diachroniques, à savoir 1884-1896 et 1897-1903, ainsi que le profil saillant de l'année 1883. Cette distribution diachronique du vocabulaire pointe l'évolution d'un discours journalistique à travers lequel se repèrent les grands événements de la Troisième République, qu'il participe à construire à son échelle au gré de pratiques de sélection et de hiérarchisation de l'information.

Cette phase pré-exploratoire est également porteuse de pistes de recherches quant à l'écriture journalistique du *Petit Comtois* : que la forme lexicale *dépêches* soit sur-représentée en 1883 est une trace significative de la stratégie inaugurale du quotidien trouvant dans la monstration d'une technologie novatrice un argument commercial et idéologique. Autre indice à même de nous guider dans notre éclairage des productions discursives journalistiques de la fin du XIX^e siècle, le sur-emploi de l'expression *le bruit court* pendant la première année de parution du *Petit Comtois*, qui ouvre la voie à un chantier de recherche sur le traitement de l'information en diachronie. Il s'agirait notamment d'éclairer les pratiques journalistiques de traitement de l'information au carrefour du XIX^e et du XX^e siècle à la lueur des stratégies que le locuteur-journaliste adopte lorsqu'il met en scène des événements qui, énonciatifs ou non, renvoient toujours au discours d'autrui. L'intérêt socio-historique de cette classe d'analyse nous paraît d'autant plus grand que la diachronie de notre corpus recouvre une période charnière entre un journalisme dit « d'opinion » et un journalisme dit « d'information ».

Plus que des résultats, l'examen de la ventilation du vocabulaire met ainsi en évidence des pistes de recherche pour la poursuite de l'analyse. Au

terme de cette classe d'observations, nous disposons en effet de nouveaux modes de partitionnements endogènes, constitués non par rapport à un savoir *a priori*, mais à partir du texte lui-même. Si ces indications pré-exploratoires déterminent ainsi le contour d'un trajet de lecture objectivé, elles ne nous permettent d'accéder qu'à une appréhension du *signe* dans les grandes sections du corpus, en l'occurrence de sa diachronie. Nous délaierons temporairement l'approche macro-distributionnelle pour nous tourner vers un autre palier d'exploration : celui de la co-occurrence, et plus précisément au palier de la *co-occurrence généralisée*, telle que proposée par J. M Viprey (1997).

4.3. Analyse de la micro-distribution du vocabulaire du *Petit Comtois*

Parvenue à ce moment de notre recherche, nous déploierons une procédure d'analyse qui prend pour objet le niveau micro-distributionnel, c'est-à-dire le réseau des co-occurents de chaque forme lexicale. L'enjeu de cette classe d'explorations est de mettre au jour, de façon synthétique et globale, l'organisation *non-séquentielle et réticulaire* (Viprey, 2006d: 81) du texte. Dans l'esprit de la linguistique distributionnelle inaugurée par Z. Harris (1952), nous visons ainsi à mettre en relief des unités de forte équivalence distributionnelle, ce que J.-M. Viprey nomme une *isotropie* et qu'il définit comme un réseau co(n)textuel commun, ou si l'on préfère, un réseau de profils lexicaux collocatifs. Il nous semble important d'insister sur le fort positionnement méthodologique qui préside au concept d'*isotropie* :

Le terme d'*isotropie* n'a pas été choisi – ni construit – au hasard de l'étymologie, seulement parce qu'il renvoie au préfixe *iso* et au radical *tre(o)p* (inclinaison/ation). Il est mis en compétition et/ou en coopération avec le terme et la notion d'*isotopie*. Compétition, s'il s'agit de l'emploi totalement perverti qu'en fait la critique académique : un équivalent jargonnant du *champ lexical* ou *notionnel*, déjà si confus, si peu opératoire et si trompeur. Coopération, complémentarité, si l'on suit le travail de (re)définition qu'a mené Rastier depuis vingt ans [...] (Viprey, 2006d : 78).

Le concept d'isotropie, au même titre que la démarche d'observation généralisées des collocations, – par opposition à la recherche des co-occurents d'un terme pivot –, s'ancre dans une perspective méthodologique fondamentalement heuristique. *A contrario* d'une projection de catégories du *lexique* construites *a priori*, ayant pour conséquence une négation de la matérialité singulière d'un texte, le concept d'*isotropie* reflète une approche centrée sur le vocabulaire dont l'organisation émerge du texte lui-même.

4.3.1. Description de la configuration fine du vocabulaire du corpus

Sur le plan méthodologique, l'examen de la co-occurrence généralisée passe par la construction d'une vaste matrice où figurent en lignes et en colonnes les items les plus fréquents du corpus. En l'occurrence, la visualisation présentée ci-dessous résulte du croisement des 1000 premiers avec à leur intersection le nombre de co-occurrences lignes-colonnes dans une unité de contexte déterminée. Le paramétrage de la zone à explorer autour des formes interrogées susceptibles de fournir une activité pertinente sur l'activité cooccurrentielle a été fixé à un empan de 15 mots à gauche et à droite dans les limites de la phrase. Ce paramétrage, comme tout paramétrage, est arbitraire, puisque nul ne voudrait soutenir que l'activité co-occurrentielle s'arrête dans ces bornes artificielles et typographiques. Il nous semble néanmoins justifié par la gestion équilibrée qu'il instaure entre le silence et le bruit dans une perspective d'abord lexicale (et non par exemple micro-syntaxique). La proximité isotropique entre les items est, dans cette perspective, l'indice de profils lexicaux proches. Selon le même principe, le plus ou moins fort éloignement (jusqu'aux positions antipodiques) de deux items traduit une dissimilarité des co(n)textes d'occurrences.

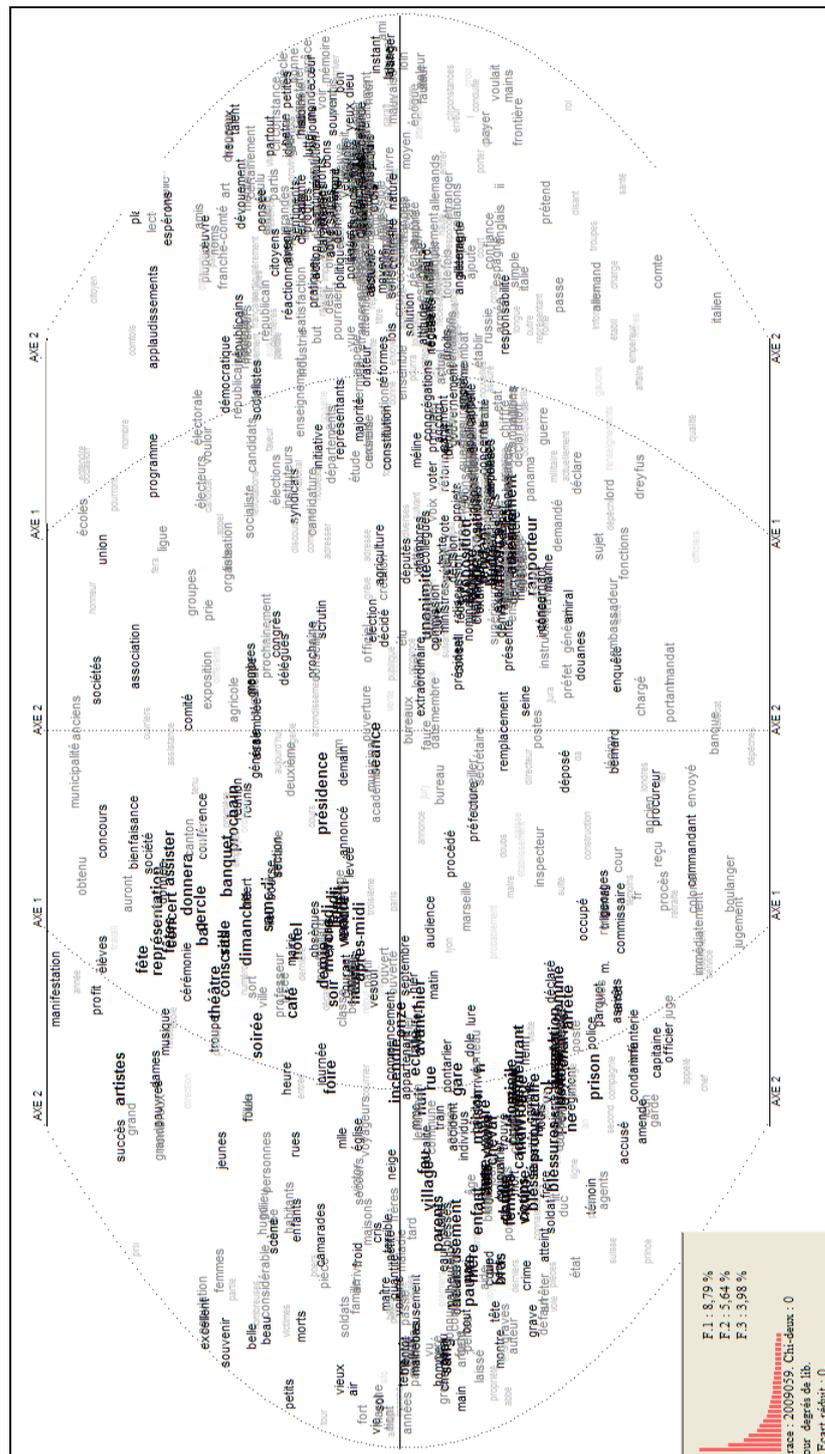


Figure IV-36 : Projection AFC de la micro-distribution du vocabulaire du corpus *Petit Comtois* (1883-1903).

Sur le facteur 1, qui représente 8,79% de l'inertie, un clivage oppose la moitié gauche du planisphère à sa moitié droite, qui semble pour sa part privilégier le vocabulaire politique. La moitié gauche de la géode est configurée par les facteurs 2 (5,64%) et 3 (3,98%) qui l'un et l'autre coopèrent à individualiser deux nuages de points²⁰ (ou *pôles isotropiques*) dans l'hémisphère gauche. Dans l'hémisphère droit, la redondance des deux facteurs est également sensible et conduit à un résultat analogue. Ces nuages se donnent à lire comme des sous-structures complexes du vocabulaire, qui donnent un aperçu thématique de l'ensemble du vocabulaire, tout en laissant apparaître les contraintes linguistiques du corpus. Pour étudier plus finement les isotropies observables sur la figure IV-36, nous observerons les zooms centrés de chacun de ces pôles.

Isotropie A : Une consistance lexico-thématique qui évoque les faits-divers et leur passage en justice

Nous envisagerons tout d'abord l'isotropie repérable dans le quart-ouest sud de l'hémisphère droit du planisphère, que, pour plus de commodité, nous nommerons isotropie A.

Les constituants de cette région dénotent le répertoire des faits-divers :

- *Cris, petit, terrible, neige, feu, localité, incendie, bientôt, sortir, pauvre, mari, mère, coup, pied, bras, mère, accusé, blessures, prison, police, parquet, arrêté, nommé, déclaré, enfant, chute, fils, femme, aussitôt, cheval, veuve, maison, hôpital, homme, perdu, sang, chercher, bout, malheureusement, bas, voisins, quantité, maître, cultivateur, cadavre, employé, vol, police, parquet, arrêté, nommé, déclaré, né, tête, ouvrier, sieur, demeurant, atteint, grave, trouvait*

²⁰ Nous admettrons qu'une partie d'un nuage de points et, elle aussi, un nuage de points en un sens moins contraignant que nous utilisons ici.

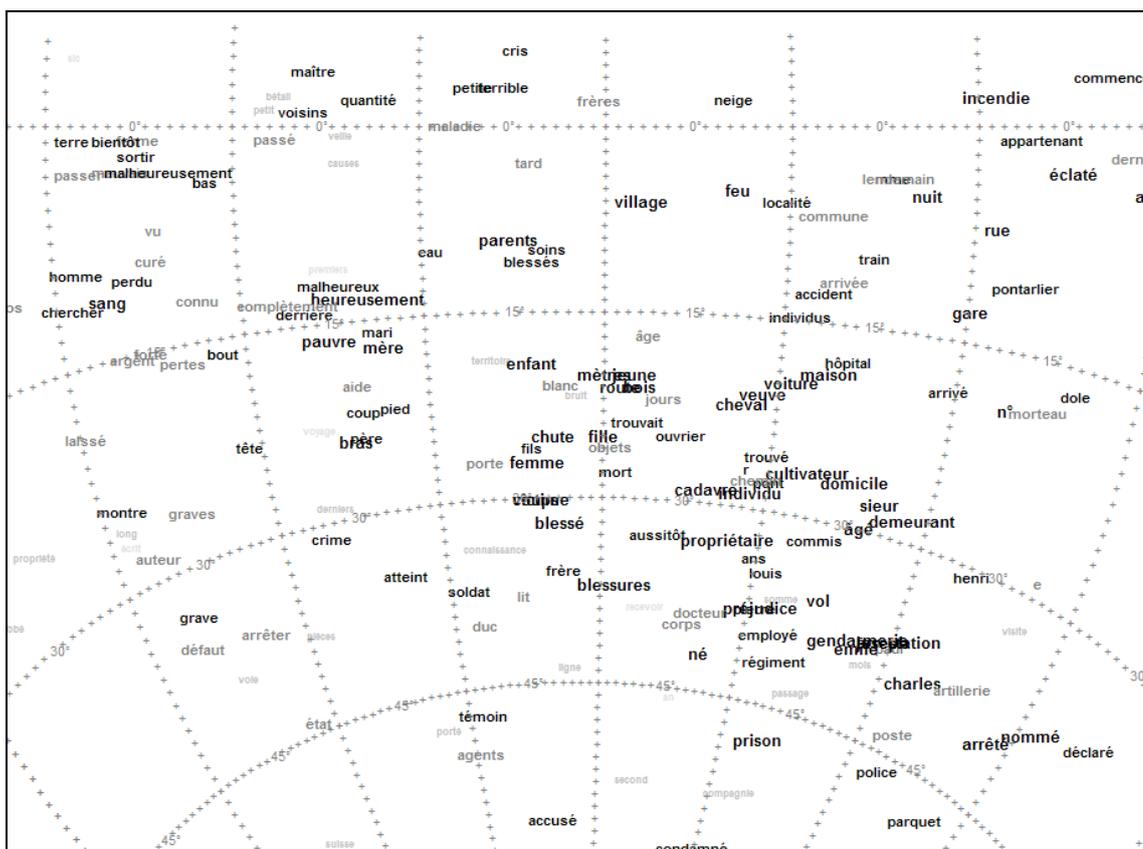


Figure IV-37 : Zoom sur la région A.

Cette isotropie présente la particularité d’être composée majoritairement de substantifs référant à des humains, qu’ils soient désignés par leur sexe ou leur rapport de parenté, (*femme, fille, fils, frère, mari, mère, père, parents, enfants, veuve*) ou qu’ils soient désignés par leur statut socio-professionnel (*ouvrier, cultivateur, propriétaire, soldat, employé, curé*). On retrouve parallèlement le lexique du corps humain, comme l’attestent les items *tête, bras, sang*, les formes *blessures, blessés, cadavre, mort*, qui se rapportent à une dégradation de celui-ci plus ou moins définitive. Un certain nombre de substantifs évoquent un schéma d’événements, à l’instar de *chute* et *incendie*. Le champ judiciaire est plus qu’évoqué dans cette région au travers des items *prison, police, parquet, accusé, témoin, crime*. On devine à travers la proximité des items *nommé, arrêté, déclaré*, la forme des minutes du tribunal. On notera la présence d’adjectifs subjectifs affectifs tels que *malheureux, terrible*, et

d'adjectifs évaluatifs tels que *grave* et *petit*, pouvant en contexte être également affectifs. Des adverbes sont aussi présents dans cette région : *malheureusement*, *heureusement*, des adverbes de temps tels qu'*aussitôt*, *bientôt*. L'adverbe *complètement* évoque quant à lui la réalisation complète et achevée d'un processus. Enfin, on relève des verbes présents sous leur forme infinitive, comme *chercher*, *arrêter*, *commencer*, mais sont aussi conjugués à l'imparfait (*trouvait*), qu'on peut légitimement supposer être la trace d'une dimension narrative.

L'isotropie B : réseaux de sociabilité et agenda

L'isotropie B, située dans le quart est nord de l'hémisphère droit, est centrée sur les termes *fêtes* et *concert*, entourés par d'autres items du même répertoire des événements culturels festifs (*fête*, *représentation*, *cérémonie*, *bal*) ainsi que des éléments verbaux évoquant les contraintes syntaxiques spécifiques de l'agenda (*donnera*, *assister*, *donnée*). Au sud de la zone, se concentrent des marqueurs chronologiques (*samedi*, *dimanche*). Dans la partie sud-est de la région, on notera la concentration de formes évoquant plus spécifiquement des événements culturels liés à l'activité politique (*réunion*, *conférence*, *banquet*, *réunis*, *section*).

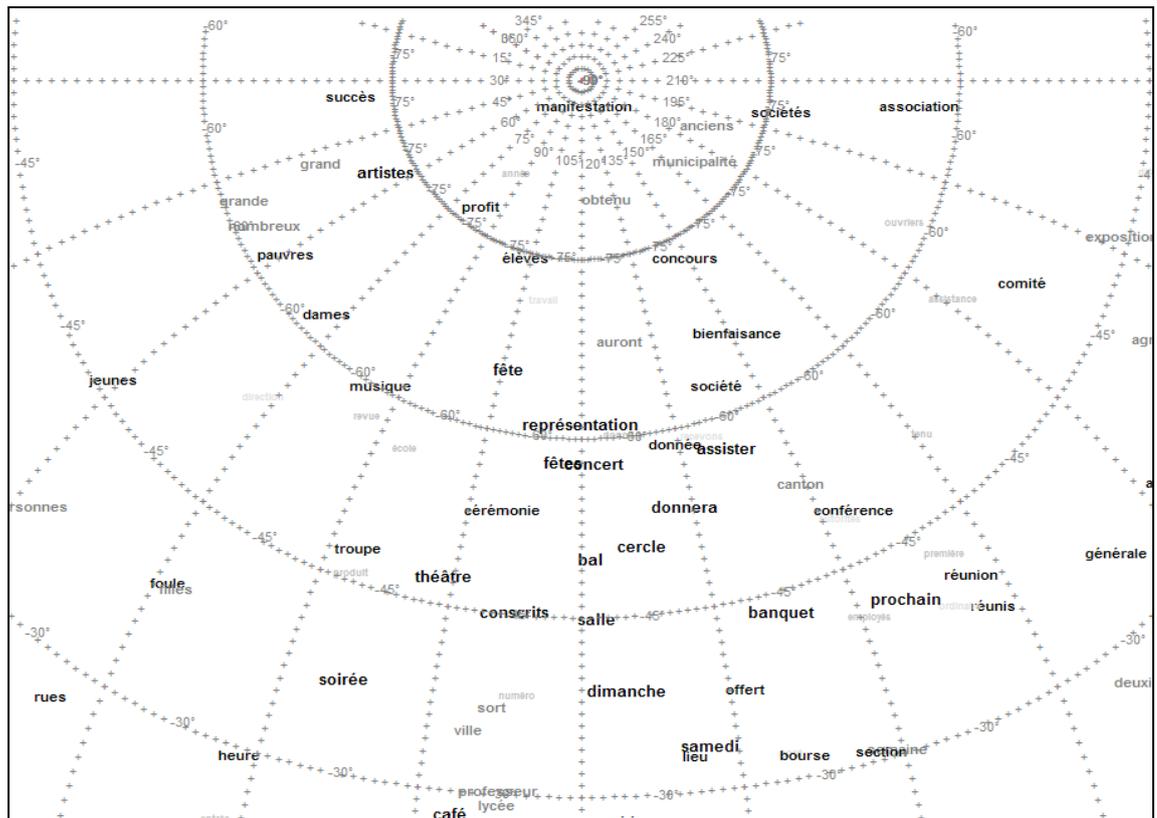


Figure IV-38 : Zoom sur la région isotropique B.

L'isotropie C : le pôle de la vie législative

Située dans l'hémisphère droit, dans le quart-ouest sud, l'isotropie C évoque le répertoire de la vie législative. Les items *rapporteur*, *décret*, *examen*, *sénat*, *rapport*, *unanimité*, *commission*, *députés* en témoignent explicitement, tandis que l'ambiguïté dont sont chargés certains substantifs virtuellement polysémiques (*chambres*, *chambre*) est levée par leur proximité distributionnelle avec ces derniers items. L'activité de la tribune parlementaire se donne à lire à travers les items *discussion*, *décision*, *interpellation*. Les items *agriculture*, *marine*, *colonies* et les items relatifs au domaine financier (*budget*, *dépenses*, *finances*, *crédit*) semblent les traces des débats des sessions parlementaires. Un ensemble de verbes à l'infinitif tels que *voter*, *examiner*, *présenter* et de verbes à l'indicatif présent comme *propose* et *adopte*, confirment enfin la relation entre l'isotropie observée et la vie législative. On remarquera par

ailleurs que certaines lexicalisations se devinent par la proximité de leurs constituants : ainsi les items *président* et *conseil* semblent pouvoir être interprétés, sans extrapolation, comme la trace du groupe nominal « le président du conseil ».

Liste des items composant l'isotropie C :

- *Demain, séance, préfecture, déposé, remplacement, seine, extraordinaire, président, présenté, décidé, élection, unanimité, commission, lecture, ministres, fédéral, nomination, chambre, examiner, interpellation, adopté, proposition, budget, décret, relatif, finances, rapporteur, marine, ministère, propositions, texte, vote, collègues, députés, agriculture, voter, discussion, chambre, projets, loi*

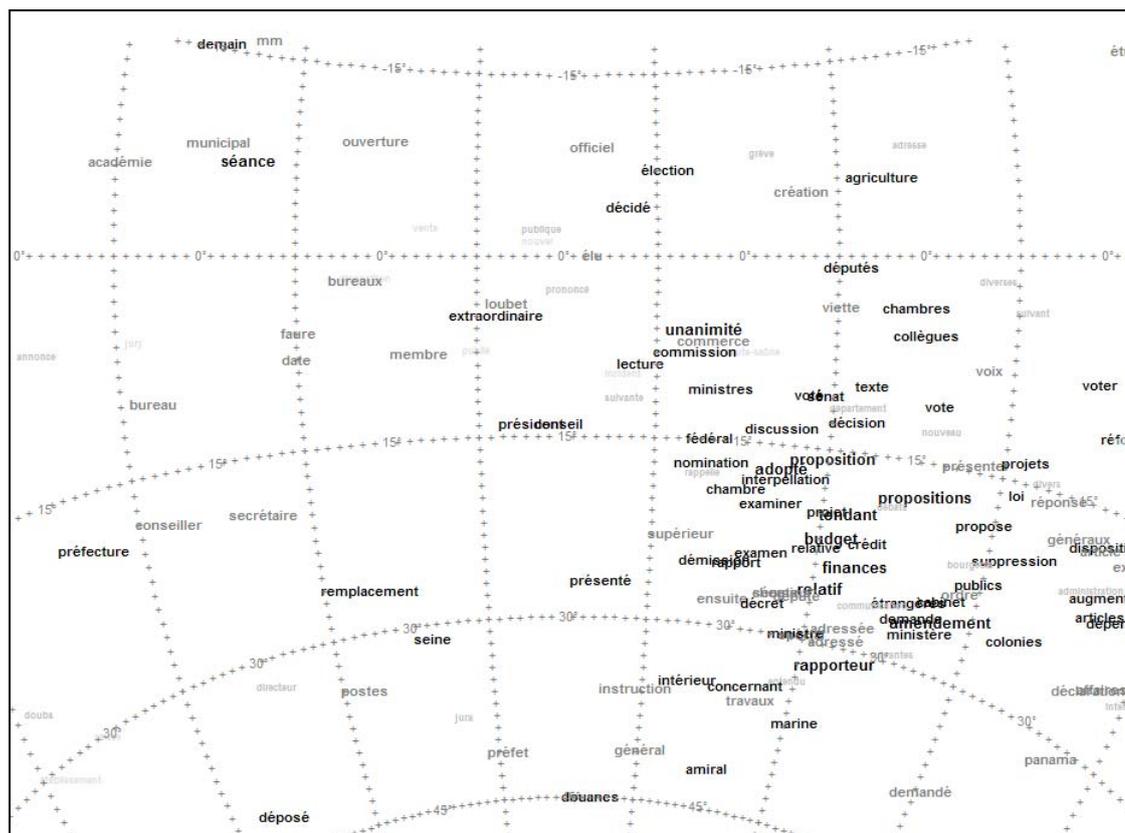


Figure IV-39 : Zoom sur l'isotropie C.

L'isotropie D : la sphère du politique

Située dans le quart-ouest nord de la moitié droite du planisphère, l'isotropie D est composée des items suivants :

- *Veux, solution, réformes, solution, puissances, égard, pense, moyens, opinion, lutte, toujours, histoire, monde, cœur, souvent, bon, dieu, instant, adversaires, démocratie, socialistes, cléricaux, progrès, volonté, pensée, dévouement, patrie, lutte, histoire, monde, lutte, révolution, bons, yeux, force, dieu, croire, crois, sentiments, est-ce, intérêts, politique, politiques, reconnaître, orateur, vérité, contraire, eglise, force*

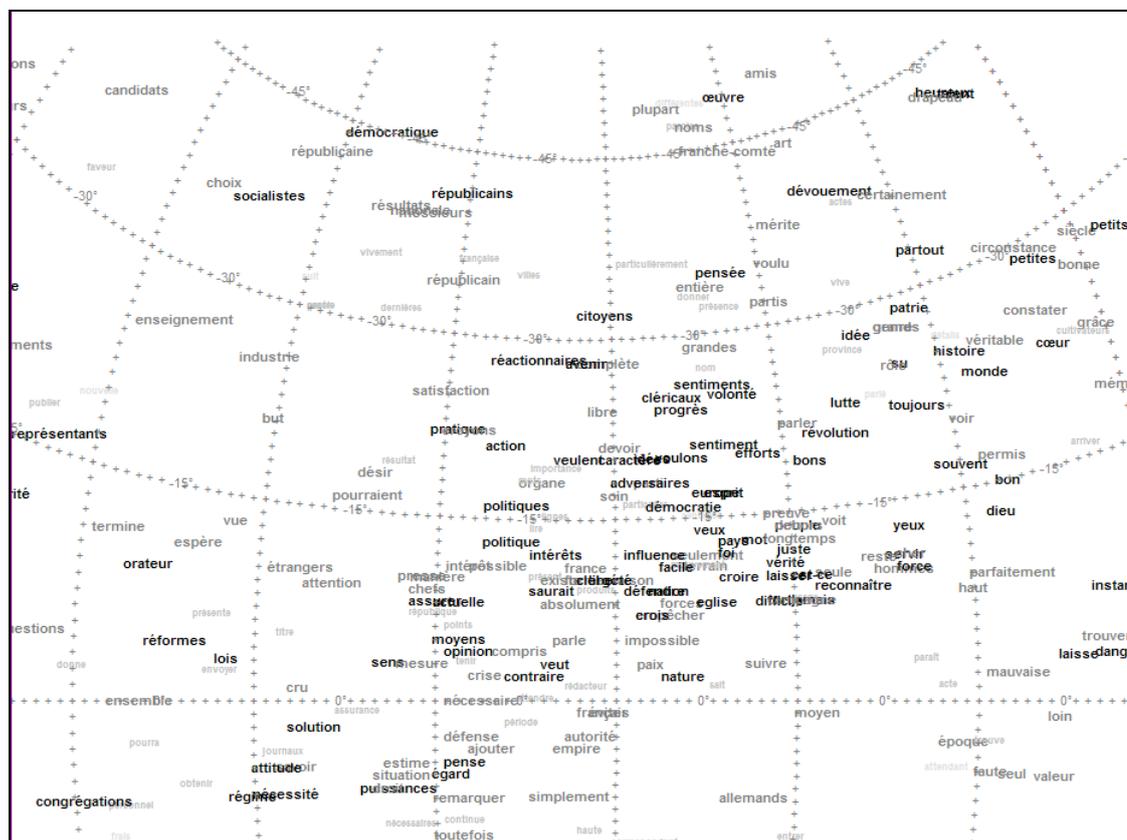


Figure IV-40 : Zoom centré sur l'isotropie D.

Il est ainsi patent qu'affleurent à sa surface les traces d'un discours politique, défini à la suite de J. Duchastel et V. Armony comme « la représentation de l'espace, de la communauté, des rapports sociaux et du rapport de l'individu à la société » (Duchastel & Armony, 1993 : 162). On notera en effet ici la référence aux grands ensembles sociétaux, parmi lesquels *pays, europe, puissances, monde, citoyens, patrie*, couplés avec les désignants politiques *socialistes, réactionnaires, cléricaux, républicains*, tandis que le terme de *politique* apparaît par deux reprises, au singulier et au pluriel. Ces désignants entrent en réseaux avec un ensemble de formes lexicales que nous envisagerons comme autant de lexicalisations de notions-concepts et notions-principes (Guilhaumou, 2003) (*progrès, patrie, démocratie*). Par ailleurs, on voit se dégager de ce pôle une dimension rhétorico-stylistique forte, décelable à travers la prédominance d'un ensemble de verbes conjugués (*crois, voulons, saurait, pense, laisse, veut*) ou à l'infinitif (*laisser, reconnaître, assurer, défendre, servir*), ainsi que d'adverbes (*partout, toujours, souvent*). Plutôt que de procéder à une analyse approfondie de chacun des items de ce pôle isotropique, nous aimerions pointer les pistes d'analyse suggérées par la similarité des profils des formes en présence.

Ce pôle nous invite tout d'abord à questionner la rhétorique du discours politique rencontré dans les colonnes du *Petit Comtois*. Une première entrée de cet axe de recherche pourrait être l'item *vérité*, qu'on observe sur le pôle D. Celui-ci renvoie à un trait très caractéristique de l'ethos mis en scène tant dans les professions de foi adressées aux électeurs que dans les articles d'opinion émanant de la rédaction, qui se targuent de dispenser la *vérité*. Cette dernière est d'ailleurs posée comme un *devoir*, terme qui est le plus fort co-occurent de la forme. Que la co-occurrence des deux items renvoient à la structure *devoir la vérité à quelqu'un* ou *devoir à la vérité de + verbe à l'infinitif*, - particulièrement récurrente sous la plume de J. Gros -, elle est la trace d'une mise en scène de l'ethos du locuteur qui se présente sous les traits de l'honnêteté personnifiée.

Ch.IV (35), « Le droit de réponse est-il le droit à l'insulte ? », (1902-03-18)

[...] Nous vous devons la vérité tout entière. Nous venons vous la dire. Avant moi, mes collègues vous ont retracé l'œuvre néfaste du ministère de Défense républicaine désorganisant successivement l'armée, la marine, les finances. Je vous demande la permission d'y revenir à mon tour. C'est en vain que nos adversaires essaient d'assimiler la situation actuelle aux premiers temps de la fondation de la République ou aux heures troublées du boulangisme. [...]

Ch.IV (36), « La lutte électorale en Allemagne » (1887-01-24)

[...] Nous devons à la vérité de dire que, même dans le peuple, on n'est plus dupe de la misérable comédie que les conservateurs ont joué jusqu'à ce jour, en présentant les « armements de la France » comme une menace. On sait fort bien, dans les provinces allemandes, que la France ne fait que suivre l'exemple de l'empire germanique, et qu'elle ne demanderait pas mieux que de vivre en paix et en bonne intelligence avec ses voisins d'outre-Rhin. [...]

D'un côté le martèlement de « la vérité » au gré de syntagmes péremptoires (*en vain, indiscutablement*) et des mises en relief par des phrases présentatives. De l'autre la dénonciation incessante d'un discours mensonger visant à *égérer l'opinion*. Il s'agit, pour les républicains, comme l'indique le premier verbe co-occurent de la forme (écart-réduit : 25,02)²¹, de *rétablir* une vérité dénaturée et de lever le voile sur le jeu des ennemis de la République :

Ch.IV (37), Chronique de la semaine, (1883-08-19)

[...] Néanmoins, comme il y a partout des gens bien informés, on a fait courir le bruit que Léon XIII était mécontent de ce que M. Grévy se fut permis d'avoir quelques paroles de critique sur l'attitude hostile des évêques. Il est assez peu probable que les diplomates du Vatican, qui sont renommés pour leur finesse, aient laissé percer leurs sentiments particuliers en cette circonstance. La vérité probable, c'est que de part et d'autre on cherche un *modus vivendi* qui mette fin à la guerre de l'Eglise contre les institutions. M. de Freycinet l'avait cherché avant le cabinet actuel, et, si j'ai bonne mémoire, il fut dans cette occasion fort joliment « roulé », comme on disait partout ailleurs qu'à l'Académie.[...]

Le débat n'est pas dans la prétendue persécution dont l'Eglise se plaint aujourd'hui ; la vérité est que, depuis le pontificat de Pie IX, l'Eglise est entrée dans une voie d'hostilité contre la société moderne, qui amènera fatalement celle-ci à dénoncer les concordats et à proclamer la séparation de l'Eglise et de l'Etat. [...]

²¹ Désormais dans ce chapitre et le suivant, la mention de nombres positifs ou négatifs entre parenthèses renvoie à l'écart-réduit.

Ch.IV (38), « Le droit de réponse est-il le droit à l'insulte ? », (1902-03-18)

[...] La vérité est toute autre : la République est, aujourd'hui, indiscutablement fondée et le peuple français, sincèrement rattaché à cette forme de gouvernement, n'entend pas que des agitateurs viennent lui substituer aucune dynastie. Où est Boulanger? Où est le prétendant? Où est le chef militaire menaçant les institutions républicaines? Le seul danger couru actuellement par la République vient du ministère au pouvoir, qui compromet par ses origines et par ses agissements les bases même de l'accord sur lequel la République fut fondée.

Pour justifier sa véracité, le discours politique républicain s'orne des traits rhétoriques de l'évidence. À *l'obscurité profonde*²² apportée dans le débat par ses adversaires, le discours républicain répond par l'ostension de l'évident, du clair, du naturel et du bon sens :

Ch.IV (39), « Petite correspondance », (1884-01-30)

Il est clair, en effet, que les déclamateurs qui décrivent journallement l'état de nos finances, sans souci de ruiner notre crédit et de tuer la confiance ; qui, à propos des luttes où l'honneur français est engagé, cherchent à égarer l'opinion en lui faisant peur des Allemands; qui, au risque de faire le jeu d'un Chinois charlatan, vilipendent le gouvernement de leur pays, il est clair, disons-nous, que ceux-là n'aiment pas la France. [...]

Ch.IV (40), Chronique régionale, Doubs, (1883-08-03)

[...] Une nation qui ne possède pas le gouvernement adapté à son état politique et social est fatalement condamnée à se voir ballotée de restauration en révolution, triste spectacle offert par notre histoire de 1789 à 1870. Bon gré, mal gré, la France est une démocratie, est le simple bon sens montre clairement que la République est le gouvernement naturel d'une nation démocratique. [...]

À la dénonciation du travestissement de la vérité que nous venons d'évoquer dans *Le Petit Comtois*, s'ajoute celle de la *spoliation lexicale* (Angenot, 1982 : 41). Tandis qu'on le taxe d'être autoritaire, conservateur ou encore nationaliste, *Le Petit Comtois* n'a de cesse d'essayer de « prendre la langue »

²² Cf. « Les financiers de la droite », (1888-01-29) : [...] Il faut entendre chaque année, lors de la discussion générale de la loi de finances, certains de ses orateurs exposer, avec une ardeur toujours nouvelle, des griefs qui ne varient jamais. M. d'Aillières, M. Daynaud, M. de Soubeyren et quelques autres ont le monopole peu enviable et apportent invariablement dans le débat une obscurité profonde C'est là un spectacle que nous voyons en ce moment se renouveler ! Nous n'en parlerions pas s'ils n'avaient déposé une proposition destinée à égarer l'opinion publique et à compliquer la situation financière de la France. [...]

(Bonnafous, Tournier, 1995 : 68), pour imposer le seul sens authentique des mots : le sien.

Ch.IV (41), « Nos mandataires », (1895-02-23)

[...] La lutte des classes, c'est là, nous ne l'oublions pas, le danger de toute démocratie ; pour l'éviter, il faut donner satisfaction à ce qu'il y a de légitime dans ses revendications. Comment cette démocratie pourrait-elle devenir conservatrice, dans le sens vrai et juste du mot, si on ne lui donne quelque chose à conserver ? [...]

Ch.IV (42), « Le droit de réponse est-il le droit à l'insulte ? » (1902-03-18)

[...] Par contre, si, pour être nationaliste, il suffit d'aimer l'armée, de demander qu'elle soit respectée dans ses chefs comme dans ses petits soldats, de défendre la France, de protester contre l'introduction de la politique dans l'armée, d'être contre les internationalistes, contre les syndicats occultes, d'être pour la nation, pour la France, pour le peuple, en un mot contre les étrangers; [...] — eh ! bien, oui, nous sommes dans ce sens nationalistes, nous ne récusons pas le mot et nous nous en faisons gloire !

C'est dans cette perspective qu'il convient d'éclairer la présence sur le pôle isotropique de la forme *vouloir* au présent de l'indicatif et à la troisième personne du singulier. Loin d'exprimer ici une volonté, la forme, qui « co-occure » avec le verbe *dire*, est employée pour « faire la loi » linguistique (Bonnafous & Tournier, 1995 : 68) :

Ch.IV (43), « Notre politique » (1883-08-12)

Qu'est-ce que cela veut dire, autoritaire ? Je défie d'abord quiconque voudra prendre nos votes de nous apporter une seule loi libérale que nous n'ayons pas votée.

Ch.IV (44), « La République et la loi », (1888-01-27)

Quand on prétend être l'Etat, on peut dire aussi qu'on est la loi. Le règne du bon plaisir est fini. Il ne reviendra pas. La France vit actuellement sous un régime de justice. Car République veut dire Liberté, et Liberté c'est Justice. [...]

Il ne s'agit ici bien entendu que d'une première approche d'un trait rhétorico-stylistique, qu'il conviendrait d'approfondir dans le cadre d'une étude ultérieure. Dans la continuité d'une observation de la « prise de la langue » par une formation discursive, ce pôle isotropique pointe une autre piste de recherche : celle de l'étude des désignants socio-politiques tels qu'ils circulent dans *Le Petit Comtois*. Prendre pour objet d'étude ces désignants, c'est s'intéresser aux

formes par lesquelles la rédaction du *Petit Comtois* définit son républicanisme et combat les forces politiques en présence. À titre d'exemple, le premier désignant socio-politique que nous souhaiterions examiner est d'ailleurs celui-là même de *républicains*. S'il est un mot dont *Le Petit Comtois* ne cesse en effet de dénoncer le détournement et de contester le sens, c'est bien celui-là. C'est du moins ce que suggèrent les principaux co-occurents de l'item sur l'ensemble du corpus, reflétant une conception dichotomique du terme : il est ainsi des *vrais* républicains (+49,8 pour 46 co-occurrences) et des *faux* républicains (+10,93 pour 17 co-occurents). Un retour au texte permet de préciser les contours de cette opposition et montre que les contextes d'emploi de ces désignants axiologisés se concentrent exclusivement sur la seconde moitié de notre corpus. Comme on peut le constater ci-dessous, la dialectique des *vrais* et des *faux* républicains ne s'impose qu'à partir de 1893, où *vrais* devient le premier co-occurent de la forme, à raison d'un écart-réduit de 52,91 pour 43 occurrences. Dans la première partie du corpus, les premiers co-occurents de *républicains* sont majoritairement des substantifs (*candidats, journaux*) qu'il qualifie ou avec lesquels *républicains* entre en réseau (*réactionnaires, sièges*), voire des adjectifs substantivisés tels que *modérés*.

De 1883 à 1892				De 1893 à 1903		
EC.RED	OCC.	FORME		EC.RED	OCC.	FORME
43,85	54	candidats		52,91	43	vrais
29,7	17	modérés		39,23	27	modérés
26,78	536	les		31,3	16	progressistes
24,57	27	réactionnaires		26,42	672	les
21,48	11	sièges		25,38	15	conservateurs
20,35	11	nuances		24,22	138	tous
17,88	16	sénateurs		22,3	13	démocrates
17,85	43	journaux		20,68	11	monarchistes
17,4	10	comités		19,97	38	union
16,93	33	députés		19,84	23	sincères
16,89	79	tous		19,52	21	réactionnaires
13,58	13	conservateurs		19,48	11	journalistes
13,49	22	union		19,07	46	républicains
13,44	11	principes		17,23	8	nuances
12,77	12	veulent		16,89	19	cléricaux
12,47	9	sincères		16,69	30	électeurs
12,44	23	électeurs		15,46	402	des
11,39	12	conseillers		15,03	16	groupes
11,36	18	arrondissement		14,8	18	démocratique
11,18	11	groupes		14,56	17	faux
10,26	40	leurs		14,33	11	ralliés
10,25	91	ont		14,3	14	auxquels
9,9	9	cléricaux		13,79	16	radicaux
9,54	237	des		13,76	20	vieux
9,47	14	délégués		13,67	24	candidat
9,46	8	élus		13,6	28	groupe
9,02	9	monarchistes		12,99	16	candidats
8,85	54	ils		12,89	13	réaction
8,8	33	république		12,41	8	nez
8,78	17	canton		11,97	114	aux
8,76	15	candidat		11,67	18	bons

Tableau IV-3 : Principaux co-occurents de *républicains* sur les périodes 1883-1892 et 1893-1903.

Seul l'item *sincères* est dans cette perspective strictement adjectival. Figurant, du moins partiellement, l'opposition qui se dessine sur la seconde moitié du corpus, l'utilisation de cet adjectif insiste déjà sur l'authenticité, la réalité et la

vérité du républicanisme défendu par *Le Petit Comtois*. Sont en effet qualifiés de *sincères* les républicains dévoués aux principes de la Révolution par opposition au « libéralisme d'apparat » (« Autoritaires et Libéraux », 1883-08-09) et/ou aux « républicains équivoques en quête d'un mandat public » (1891-01-27), autrement dit intéressés. La qualification du républicanisme suivant cet « axe de la vérité et de l'authenticité », marginale dans la première moitié du corpus, se développe surtout à partir de 1893. L'intensification de la qualification positive du républicanisme trouve, de façon générale, son origine dans la menace de l'identité républicaine par la politique d'apaisement menée dès 1892. Le désignant *vrais républicains* résonne tout d'abord, dans ce contexte, comme une dénonciation du ralliement des catholiques à la République suite à l'encyclique papale du 20 février 1892 qui les y invite :

Ch.IV (45), « M. Dequaire à Marnay », Chronique régionale, (1897-02-23)

[...] Depuis que le pape d'Italie en a donné l'ordre, cela à la honte d'une partie des Français qui ont accepté cet ordre, tout le monde est ou se dit républicain ; dans ces conditions, il faut des signes de ralliement auxquels on reconnaîtra les vrais républicains. [...]

La rédaction du *Petit Comtois*, à l'instar de ses lecteurs et correspondants radicaux, réserve ainsi le désignant de *vrais républicains* aux anticléricaux :

Ch.IV (46), Chronique régionale, Doubs, Montbéliard, (1899-02-22)

Le cléricalisme, ennemi de Gambetta, ami de Méline, a miné l'opinion publique et les vrais républicains constatent avec douleur le recul vers le fanatisme et la barbarie que les Croix proclament à son de trompe.

Si l'adjectif *vrais*, associé à *républicains*, est exclusivement antéposé en vue de déterminer ce dernier, son antonyme, *faux*, témoigne de conditions d'emplois syntaxiques sensiblement différentes. Sur les 17 co-occurrences des items *républicains* et *faux*, seules six occurrences correspondent au désignant *faux républicains*. Ce terme ne désigne d'ailleurs pas les catholiques ralliés, mais les « républicains de gouvernement », membres et partisans du gouvernement Méline, comme on peut l'observer à la lecture de l'exemple ci-dessous :

Ch.IV (47), « M. Dequaire à Marnay », Chronique régionale, (1897-02-23)

[...] En ma qualité de président du vaillant comité radical du canton de Marnay, je croirais manquer à mon devoir de démocrate convaincu si je ne venais, à cette tribune, remercier les braves lutteurs du canton de Marnay qui, par un effort commun, n'ont pas hésité un instant à fonder notre belle association radicale devant les agissements louches et rétrogrades des faux républicains ou républicains de gouvernement. [...] A notre appel, en face du danger qui nous menace, notre comité s'est triplé en trois mois et c'est bientôt à la tête d'une véritable armée de radicaux sincères que nous ferons triompher, ici, les principes de la Révolution française, en face des faiblesses des faux républicains qui, actuellement, détiennent le pouvoir.

Aux républicains anticléricaux, s'opposent en fait dans notre corpus les *républicains à faux nez*, par lequel le radical *Petit Comtois* réfère aux catholiques ralliés.

Ch.IV (48), Chronique régionale, Baume-les-Dames, (1897-02-23)

[...] Nous avons déjà assez de peine à tenir tête, dans les trois départements, à la politique équivoque des républicains à faux nez qui ont entrepris de livrer la République aux influences cléricales, en s'assurant, par des avantages personnels, le concours de certains radicaux, conscients ou non du rôle qu'ils jouent.[...]

Ch.IV (49), (1898-02-21)

[...] Eh bien ! nous y sommes dans la politique d'aventures; et toute la bourgeoisie « bien pensante », en commençant par les républicains à faux nez jusqu'aux cheveu-légers de la vieille monarchie, est derrière le ministère qui a provoqué ce gâchis ; ce qui nous autorise à dire que les pires révolutionnaires, en France, sont ces prétendus conservateurs qui nous trouvent trop avancés parce que nous sommes pour une République qui serait énergiquement républicaine.

Ce désignant, dont nous avons trouvé trace dès 1885 dans l'ouvrage de Raoul Luçay (1885 : 22), est construit à partir d'un substantif désignant un élément cocasse de déguisement du visage. Le thème du déguisement est d'ailleurs particulièrement sollicité pour dénoncer la sensibilité républicaine affectée par les réactionnaires et les cléricaux :

Ch.IV (50), (1900-01-18)

[...] il est républicain parce que, voulant être candidat contre M. Mercier, il jugea alors prudent de cacher ses opinions et de se présenter comme républicain. Hélas, pauvre *Dépêche*, ce masque de républicanisme ne colle pas ; le réactionnaire dépasse. [...]

Ch.IV (51), Chronique régionale, Doubs, Baume-les-Dames, (1900-03-28)

[...] il faut bien que la *Dépêche* se rende à l'évidence, sa clientèle électorale est faite des débris des anciens partis hostiles à nos institutions démocratiques, dont le masque républicain a pu faire des dupes, mais qui, aujourd'hui, ne trompe plus que ceux qui veulent être trompés. Vous êtes dans l'engrenage de la réaction cléricale ; elle ne vous lâchera pas et vous y passerez tout entier. [...]

Ch.IV (52), « Les jésuites », (1897-02-02)

[...] Le petit volume de Gueneau sera le collyre qui éclaircira la vue souvent troublée de nos braves campagnards et qui leur permettra de reconnaître dans le hobereau, dans le richard qui sollicitera leurs votes l'an prochain, le Jésuite déguisé, le chat-fourré cléricale qui fait patte de velours pour mieux croquer les pauvres petites libertés que nous avons eu tant de peine à obtenir.[...]

Pour aller plus en avant concernant la désignation des *adversaires* politiques, il convient de remarquer que ceux-ci sont représentés sur le pôle isotropique D à travers les formes *cléricaux* et *réactionnaires*. Il est intéressant de remarquer le contraste entre le nombre important d'occurrences de ces formes, à savoir respectivement 337 occurrences et 398 sur l'ensemble du corpus, avec la relative rareté des désignants socio-politiques proches issus d'une procédure d'agglutination : on ne dénombre en effet qu'une seule occurrence de *clérico-monarchique* et de *clérico-libérale*. Le terme de *réaction cléricale* ne se rencontre pour sa part qu'à hauteur de 16 occurrences. La proximité des profils micro-distributionnels des items *cléricaux* et *réactionnaires*, signalant l'équivalence de leur contexte d'emploi est principalement trace de l'assimilation des deux groupes politiques dans le discours du *Petit Comtois*. Nombreux sont les énoncés où les deux désignants se succèdent soit par juxtaposition, – en produisant un effet d'accumulation des menaces politiques pesant sur la République –, soit par coordination, de sorte à rendre compte de leur alliance historique récurrente :

Ch.IV (53), « Les jésuites », (1897-02-02)

Cléricaux, réactionnaires, conservateurs, tout le monde se dit républicain : c'est le mot d'ordre soufflé par les Jésuites.

Ch.IV (54), Chronique régionale, (1900-01-04)

[...] Ainsi qu'il arrive toutes les fois que les réactionnaires et les cléricaux, jettent le masque et s'unissent ouvertement pour combattre les institutions

démocratiques, les républicains éprouvent, aujourd'hui le besoin de s'unir, eux aussi, contre ces ennemis de toute liberté. [...]

Co-occ. de "réactionnaires"			Co-occ. de "cléricaux"		
EC. RED	OCC.	FORME	EC. RED	OCC.	FORME
40,41	28	cléricaux	39,61	28	réactionnaires
33,07	21	feuilles	35,73	15	royalistes
29,91	45	républicains	31,83	16	monarchistes
21,06	35	journaux	27,8	10	ralliés
19,64	315	les	22,17	314	les
15,85	10	sénateurs	21,14	18	bons
13,6	12	candidats	19,44	28	républicains
12,37	192	des	10,25	17	journaux
10,06	12	élections	9,65	39	leur
9,03	11	candidat	9,11	148	des
8,65	14	amis	8,93	50	aux
8,31	51	sont	8,9	25	leurs
7,92	206	et	8,09	188	et
7,32	32	tous	7,04	42	sont
7,22	23	leurs	6,94	118	que
6,75	10	députés	6,73	59	se
6,53	8	haute-saône	6,57	23	contre
6,47	18	république	6,53	45	ont
5,73	29	ils	5,86	27	ils
5,69	8	électeurs	5,81	21	nos
5,54	95	qui	5,24	9	eux
5,23	39	aux	5,13	14	république
5,21	9	républicain	4,9	51	ne
5,14	43	ont	4,81	8	amis
5	10	voix	4,42	9	jamais
4,76	20	nos	4,24	77	qui
4,75	9	eux	4,22	36	plus
4,4	22	dont	4,04	20	tous

Tableau IV-4 : Principaux co-occurents de *réactionnaires* et de *cléricaux*.

Ce sont ces alliances « coupables » que *Le Petit Comtois* dénonce également par le terme *amis*, qui est un co-occurent privilégié des désignants *cléricaux* et *réactionnaires*, comme on peut l'observer ci-dessus.

Ch.IV (55), « La droite et la loi militaire », (1889-01-24)

Nous allons plus loin et nous disons qu'en sa qualité de moraliste il est tenu de donner l'exemple du devoir; or, il nous semble que c'est un des premiers devoirs, un devoir sacré pour tout citoyen, de servir son pays, de le défendre, et s'il le faut, de verser son sang. Mais le clergé et ses amis les réactionnaires ne l'entendent pas de cette oreille. Habitué à jouir d'une foule de privilèges que leur a concédés la monarchie, ils n'admettent pas que le régime républicain les soumette au droit commun.

On remarquera, parmi la liste des co-occurents présentés ci-dessus, l'item *bons*, qui, à dix-huit reprises dans notre corpus, est antéposé à *cléricaux*. Des principes de bonté prônés par les catholiques, les républicains en tirent argument ironiquement : ils font résonner la dimension axiologique intrinsèquement positive de l'adjectif, dimension axiologique que la mise en contexte renverse :

Ch.IV (56), « La croisade contre le bureau de bienfaisance », (1890-01-13)

[...] Ces bons cléricaux sont toujours intéressants à observer ! Sous le nom de Pharisien, ils ont autrefois tué Jésus qui avait le tort, paraît-il, de dire trop de vérités. Puis comme le souvenir de celui-ci continuait à être populaire, malgré un supplice infamant, ils l'ont fait dieu pour avoir occasion de lui consacrer un culte qu'ils exploitent en conscience depuis dix-huit siècles.

Ch.IV (57), « Coups de crosse dans l'eau », (1892-01-27)

Ah ! oui, sans doute, nos bons cléricaux consentiraient à faire leur paix avec la République, si celle-ci, de son côté, voulait bien abroger la loi sur l'enseignement, la loi qui impose le service militaire aux séminaristes, et si elle laissait carte blanche aux congréganistes de tout poil. C'est-à-dire, en d'autres termes, que si la République voulait bien consentir à renier ses principes, à n'être plus la République, messieurs de l'épiscopat lui feraient la gracieuseté de ne plus la traîner dans la boue et de ne plus la pourfendre à coups de crosse.

Au cours de notre examen des réseaux paradigmatiques de l'item *cléricaux*, nous avons été interpellée par le désignant *honnêtes gens*, comptant près de 174 occurrences dans notre corpus²³. L'évolution sémantique de cette expression pendant la Révolution a été retracée par M. Dorigny (1985). Ce dernier montre comment, d'un terme positivement connoté, désignant sous

²³ Cette séquence graphique complexe n'ayant pas été définie comme une unité linguistique et intégrée au dictionnaire de notre corpus, ses occurrences sont en quelque sorte atomisées dans l'index hiérarchique.

l'ancien régime le comportement social d'un homme aisé, riche et instruit, *honnêtes gens* devient un qualificatif dépréciant dès le printemps 1792, lorsqu'il entre en opposition avec le terme *sans-culottes*. D'après M. Dorigny, ce renversement sémantique date plus précisément du discours du 28 juin 1792 de Lafayette, qui, s'érigeant en représentant des *honnêtes-gens*, condamne l'invasion des Tuileries :

Il semble que le discours de Lafayette ait eu pour résultat de provoquer le rejet de la formule « honnêtes-gens » par tous ceux qui à cette date, ne voyaient plus d'issue que dans un renversement du trône (Dorigny, 1985 : 85).

Un siècle plus tard, l'expression se révèle encore un lieu de tension sémantique et de lutte politique. C'est du moins ce qui transparaît de ces contextes d'emploi dans notre corpus. Deux sens nous paraissent cohabiter : en premier lieu, conformément à son contenu sémantique prévalant à partir du printemps 1793, il désigne principalement les opposants à la République, et plus précisément, les royalistes :

Ch.IV (58), Lettres de Paris, (1885-01-30)

[...] En attendant, il est très curieux d'examiner l'évolution que les ex-monarchistes des 24 et 16 mai tentent en ce moment vers les doctrines libérales. Ces honnêtes gens comptent bien, paraît-il, se poser désormais en défenseurs des libertés publiques et en appeler de leur défaite du 25 janvier au suffrage universel. Ils espèrent qu'en ayant l'intransigeance pour auxiliaire dans cette campagne, ils parviendront à se faire prendre au sérieux comme libéraux. [...]

Ch.IV (59), (1885-01-26)

[...] On a sorti alors la « monarchie démocratique » de M. Hervé, le rédacteur en chef du *Soleil*, et ces honnêtes gens se sont trouvés pris tout à coup d'une tendresse inexprimable pour le suffrage universel. Ils ne songeaient plus à le traiter dédaigneusement de « loi du nombre » comme autrefois.

En second lieu, on relève dans un nombre plus restreint d'énoncés un emploi de l'expression insistant sur sa dimension culturelle et non politique. *Honnêtes gens* renvoie au modèle de l'homme cultivé, du gentilhomme, qui fait écho à la définition qu'en donne Furetière²⁴ à la fin du XVIII^e siècle :

²⁴ « Honnête homme, se dit d'un homme de bien, d'un gentilhomme, qui a pris l'air du monde, qui sait vivre, qui a du mérite et de la probité ».TLFI en ligne.

Ch.IV (60), Revue de la Santé, (1883-08-05)

[...] Ah ! c'est sur ces dernières surtout que l'hygiéniste doit appeler l'attention des honnêtes gens. Car santé, vigueur corporelle et mentale, ne vont pas sans les bonnes mœurs, comme c'est facile à démontrer. [...]

À partir de 1889, on observe une intensification des emplois du désignant, qui, de surcroît, se pare de guillemets quasi-systématique, ce qui témoigne d'une prise de distance avec la matérialité de la séquence. Manifestement, cette évolution d'emploi du désignant est liée à la tenue des élections législatives, et à l'intensification des polémiques entre partisans d'une République fragilisée, plus que jamais menacée par la coalition du clergé, des royalistes et des nationalistes. Il nous semble donc voir dans la massification des emplois du désignant une intensification des attaques contre les royalistes d'une part, et dans les guillemets un écho au discours nationaliste. Les nationalistes, le général Boulanger en tête, ont tiré parti des affaires politiques et financières qui se sont succédées depuis 1887 pour faire de l'« honnêteté » leur cheval de bataille. Dans le programme boulangiste de 1888²⁵, le slogan « Boulanger, c'est l'honnêteté » est ainsi martelé, tandis qu'il déclare vouloir débarrasser le peuple de la « voyoucratie » pour instaurer la « République des Honnêtes gens »²⁶.

Ch.IV (61), « Appel à l'énergie », (1889-01-04)

L'un des organes officieux de l'orléanisme se plaint avec amertume de ce qu'il appelle « le découragement des honnêtes gens. » Nous n'avons pas besoin de dire que dans l'esprit de la feuille en question, « les honnêtes gens » ce sont les royalistes.

Donc, si nous en croyons ce journal, les « honnêtes gens » découragés « s'abandonnent, comptant sur l'intervention surnaturelle de quelque hasard heureux beaucoup plus que sur eux-mêmes. » [...]

Ch.IV (62), (1889-01-07)

²⁵ cf. Gallica en ligne :

<http://gallica.bnf.fr/ark:/12148/bpt6k1128115.r=%22G%C3%A9n%C3%A9ral+Boulanger%22.lan&FR.textePage>

²⁶ Guyot, Y., 1988, Le boulangisme. Disponible en ligne :

<http://gallica.bnf.fr/ark:/12148/bpt6k75075h.r=voyoucratie.langFR>

[...] Dans une réunion qu'ils viennent de tenir pour prendre une détermination en vue de l'élection de la Seine, les amis de Philippe Vif, suivant leur habitude, ont longuement discoursé pour ne rien dire qui vaille. [...]

Au fond, il s'agissait de savoir s'ils se déclareraient pour ou contre M. Boulanger.

Naturellement, ces aigrefins n'ont dit ni oui ni non.

M. Boulanger est l'ennemi de la République : donc il faut le ménager.

M. Boulanger a chassé nos princes et travaille probablement pour son propre compte et non pour celui de M. le comte de Paris : donc il ne faut pas le soutenir trop énergiquement.

Tel est le raisonnement que ces « honnêtes gens » se sont fait [...]

Ce qu'il nous semble intéressant de remarquer, c'est ainsi la non-coïncidence du désignant à lui-même (Authier, Revuz, 1995) mise en relief par la formation discursive républicaine. Les républicains n'ont de cesse de dénoncer l'inadéquation entre le sens littéral de l'expression et sa connotation positive héritée de son sens antérieur à la Révolution, avec les qualités morales et politiques de leurs adversaires. Il convient d'ailleurs de remarquer que les groupes référés par le désignant *honnêtes gens* sont divers, et ne sont pas uniquement politiques, mais socio-politiques. Le désignant réfère ainsi parfois aux cléricaux :

Ch.IV (63), « La croisade contre le bureau de bienfaisance », (1890-01-13)

[...] Moyennant cela, nos bons cléricaux continueraient à proclamer qu'en dehors d'eux et de leurs amis, il n'y a pas sur la terre « d'honnêtes gens », attendu qu'ils sont les continuateurs du Dieu.

Ch.IV (64), Chronique régionale, Jura, (01-02-1897)

L'abondance des matières ne nous a pas permis de parler jusqu'ici du double procès en diffamation intenté par M. Trouillot, député, à la *Croix du Jura*. [...] Nous aurons d'ailleurs occasion de revenir sur ce procès qui est caractéristique de la manière de faire cléricale et qui montre à quelles armes ces « honnêtes gens » ne craignent pas de recourir pour avoir raison de leurs adversaires.

Ch.IV (65), Chronique régionale, Doubs, Pontarlier (1902-03-17)

[...] Une histoire assez amusante qui vient de se passer dans notre commune, montre sous son véritable jour la façon dont s'y prennent les femmes du parti dit « des honnêtes gens » pour récolter de l'argent devant servir à l'écrasement des républicains. Donc la femme d'un des chefs du parti cléricale se présente au domicile d'un de ses voisins et s'annonce comme faisant une quête pour les boers. [...]

Honnêtes gens qualifie également les républicains modérés, accusés de vouloir livrer la République aux cléricaux et aux monarchistes :

Ch.IV (66), Chronique régionale, Jura, (1898-02-16)

[...]La *Liberté* nous apprend, en effet, que la *Cloche* est l'organe de M. Waldeck-Rousseau. Allons! allons! même dans le parti des « honnêtes gens », il y a encore de beaux jours pour l'intrigue; et si ses derniers dilettante, si M. Labordère et Lelièvre ne sont pas contents, c'est qu'en vérité ils sont difficiles !

Ch.IV (67), « Comptes à régler », (1897-02-09)

Le temps et l'abondance des matières ne m'ont pas permis de donner immédiatement suite au projet que j'avais formé de rechercher ce qu'il peut y avoir de si extraordinairement vertueux dans la coalition Grillon-Moustier-Rimbaud-Chaudey-Vuillecard et Cie, pour autoriser tous ces braves gens à le prendre de si haut avec nous et à nous traiter si dédaigneusement [...]. Certes, s'il y a une politique bien réellement loyale, ce serait celle qui consisterait de la part de tous ces modérantistes à nous dire ouvertement, à nous autres radicaux: « Vous êtes trop avancés pour nous, et vous nous conduisez plus loin que nous ne voulons aller. [...] Mais loin de procéder ainsi, l'oligarchie dont nous parlons crée la *Dépêche*, et l'affublant dès le début d'un faux nez, elle impose à son nourrisson l'obligation de défendre d'abord le ministère radical Bourgeois, parce qu'il fallait, à la veille des élections municipales, sauver en premier lieu la mise de l'un des associés, le pseudo-radical Vuillecard, maire de Besançon. Puis, lorsque le tour fut joué, c'est-à-dire lorsque M. Vuillecard fut réélu au titre radical, la *Dépêche* tourna casaque et s'éprit subitement d'une belle passion pour le ministère Méline-Rimbaud, qui s'appuie sur la droite contre l'extrême gauche. Oui ou non, les faits sont-ils exacts? Et s'ils sont exacts, qu'y a-t-il donc de si étonnamment vertueux et de si profondément « honnêtes gens » dans toutes ces pantalonnades politiques? [...] Quel exemple de moralité politique donné aux populations, par ceux-là même qui se proclament « honnêtes gens » à l'exclusion de tous autres !

Parfois encore, et plus largement, *honnêtes gens* réfère à la société *bien pensante*, où le désignant a trait au niveau social et économique, mais aussi culturel :

Ch.IV (68), Chronique régionale, Flagey, (1902-03-05)

[...] Voulez-vous un spécimen de la vertu des « honnêtes gens » et de l'art avec lequel ils pratiquent cette charité chrétienne qui la fait aimer et respecter? Voici : Il y a quelque temps, un personnage très bien pensant de notre région vendait sur pied le bois d'une de ses propriétés [...]. Le coupeur, au lieu de brûler le bois mort comme cela se fait d'habitude, le mit en tas, le réservant ainsi pour son usage et se réservant de le rentrer chez lui son travail terminé. Il se croyait dans son droit ; mais le riche et bien pensant propriétaire ne l'entendait pas ainsi, et il en avait décidé autrement. Il prévint, en effet, son acheteur que, s'il enlevait le bois mort, il le poursuivrait en dommages-intérêts, préférant que ce bois fût

brûlé sur place. Il est vrai que le coupeur est républicain et par conséquent mal pensant aux yeux de son riche et pieux vendeur !

La surcharge sémantique du désignant génère une certaine confusion, perceptible dans les demandes de précisions sur son sens ainsi que dans les énoncés métadiscursifs visant à préciser le sens visé par le locuteur :

Ch.IV (69), Chronique régionale, Haute-Saône, (1894-02-20)

Votre correspondant termine en disant que les honnêtes gens soutiennent le curé. Tous les notables de l'endroit me chargent de vous demander ce que le susdit correspondant entend par honnêtes gens.

Ch.IV (70), « La fièvre du devoir », (1895-02-2)

[...] Les ministres d'aujourd'hui peuvent être des maladroits, ils l'ont déjà prouvé dans l'affaire Canrobert, personne ne peut dire du moins que ce ne sont pas d'honnêtes gens, nous parlons ici d'honnêteté dans le sens de sincérité politique, médiocres peut-être, assurément de bonne foi.

4.3.2. Synthèse sur la configuration fine du vocabulaire

L'observation de la structure fine du vocabulaire nous livre des indications précieuses sur la structuration lexico-thématique du corpus, qui rejoignent en partie les résultats de l'analyse du rubriquage précédemment menée (chapitre III). Les quatre grands pôles isotropiques distingués sont précieux pour appréhender le discours du *Petit Comtois*. L'un décrit en effet le discours journalistique qui fait la part belle aux familles événementielles du crime, du vol, de l'accident, qui sont les motifs obsédants de la presse du XIX^e siècle comme l'a montré D. Khalifa (1995). Le second met au jour la fonction d'agenda socio-culturel et mondain du quotidien, diffusant dans ses pages régionales les avis de réunions et les annonces de festivité. Le troisième pôle isotropique confirme pour sa part le rôle central de la vie législative dans la presse de la Troisième République régionale. Enfin, un dernier pôle isotropique révèle la prégnance du politique et de la défense de l'idée républicaine dans les colonnes du *Petit Comtois*.

De cette première description de la structure fine du vocabulaire de l'ensemble du corpus, est néanmoins évacuée toute dimension diachronique.

Or, il serait stimulant d'évaluer dans quelle mesure cette configuration lexico-thématique du discours du *Petit Comtois* varie, *a fortiori* lorsque l'analyse macro-distributionnelle du vocabulaire démontre un net clivage entre la première tranche diachronique du corpus et la seconde.

4.4. Analyse de la structure fine du vocabulaire en diachronie

4.4.1. Premières expérimentations

Pour ce faire, la première classe d'analyses mise en œuvre afin d'observer l'influence de la diachronie a consisté à étudier la configuration fine du vocabulaire telle qu'obtenue par les méthodologies de l'AFC dans diverses sections chronologiques du corpus. Sur le plan méthodologique, nous avons donc soumis, suivant le même principe qu'exposé précédemment, une matrice de co-occurrence des 1000 items les plus utilisés dans le corpus, pour un moment chronologique déterminé. L'AFC a permis de constituer les planisphères suivants, qui rendent compte de la configuration fine du vocabulaire en 1885, 1890, 1896 et 1900.

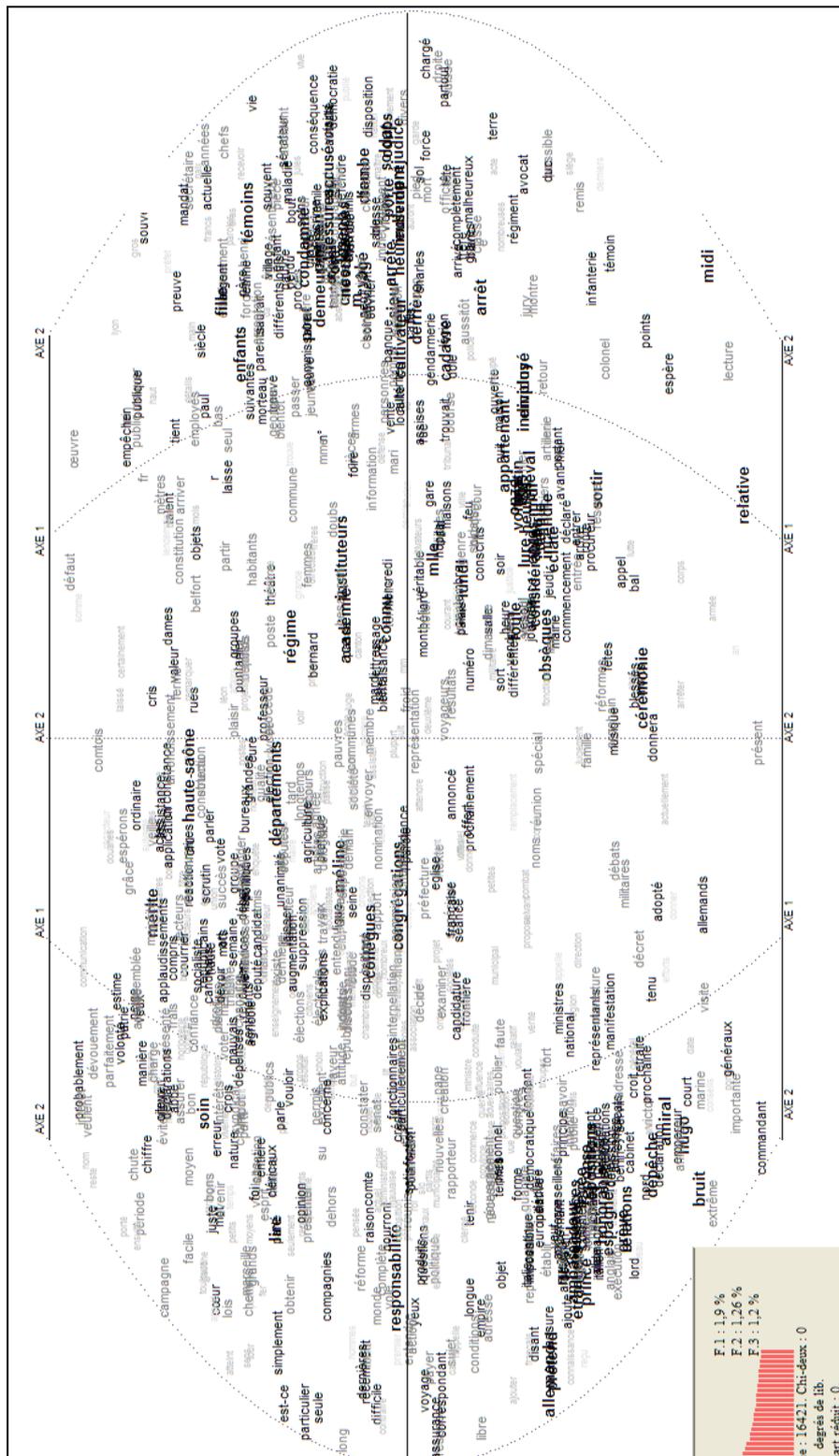


Figure IV-41 : Projection AFC de la configuration fine du vocabulaire de 1885.

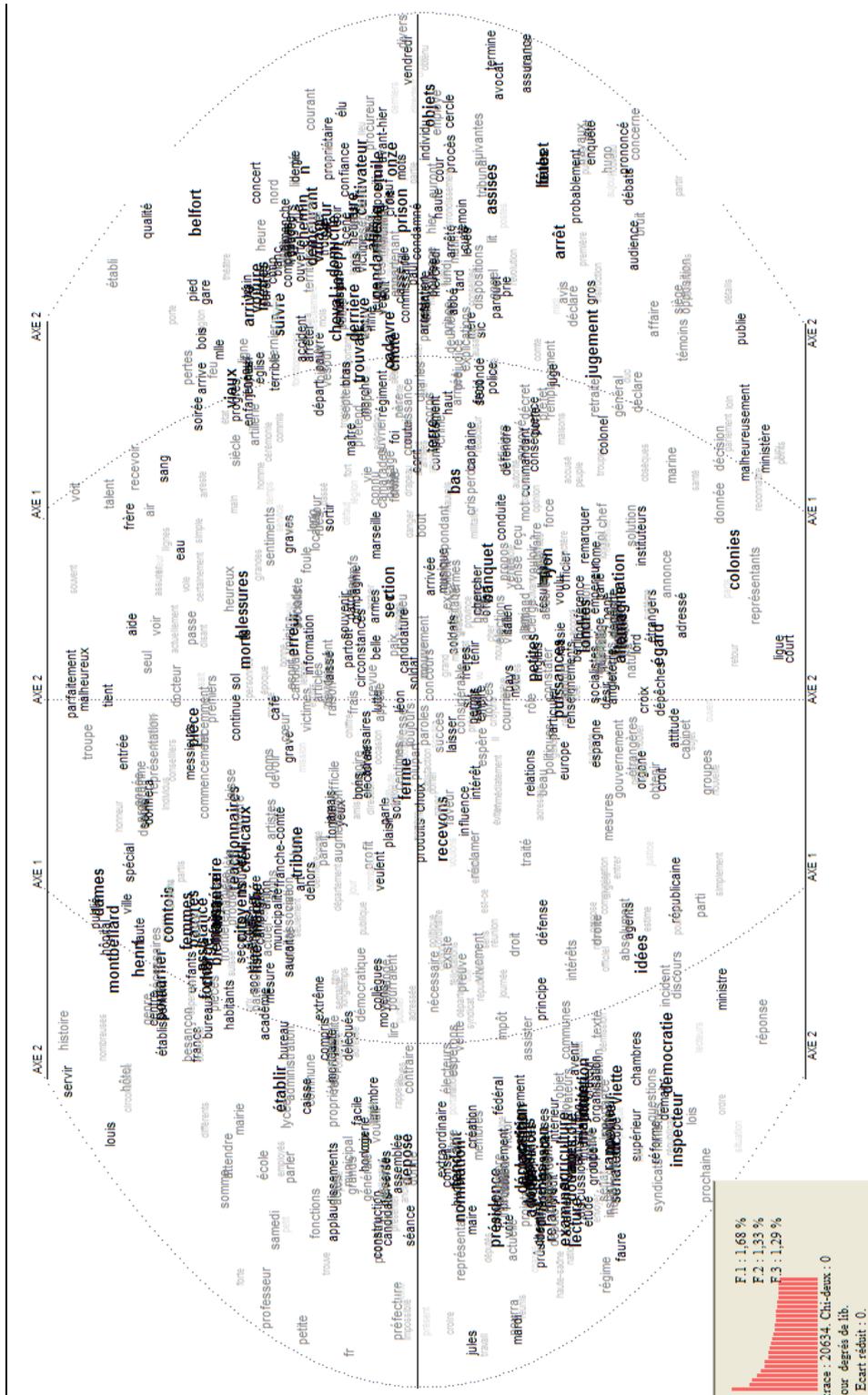


Figure IV-42 : Projection AFC de la configuration fine du vocabulaire de 1890.

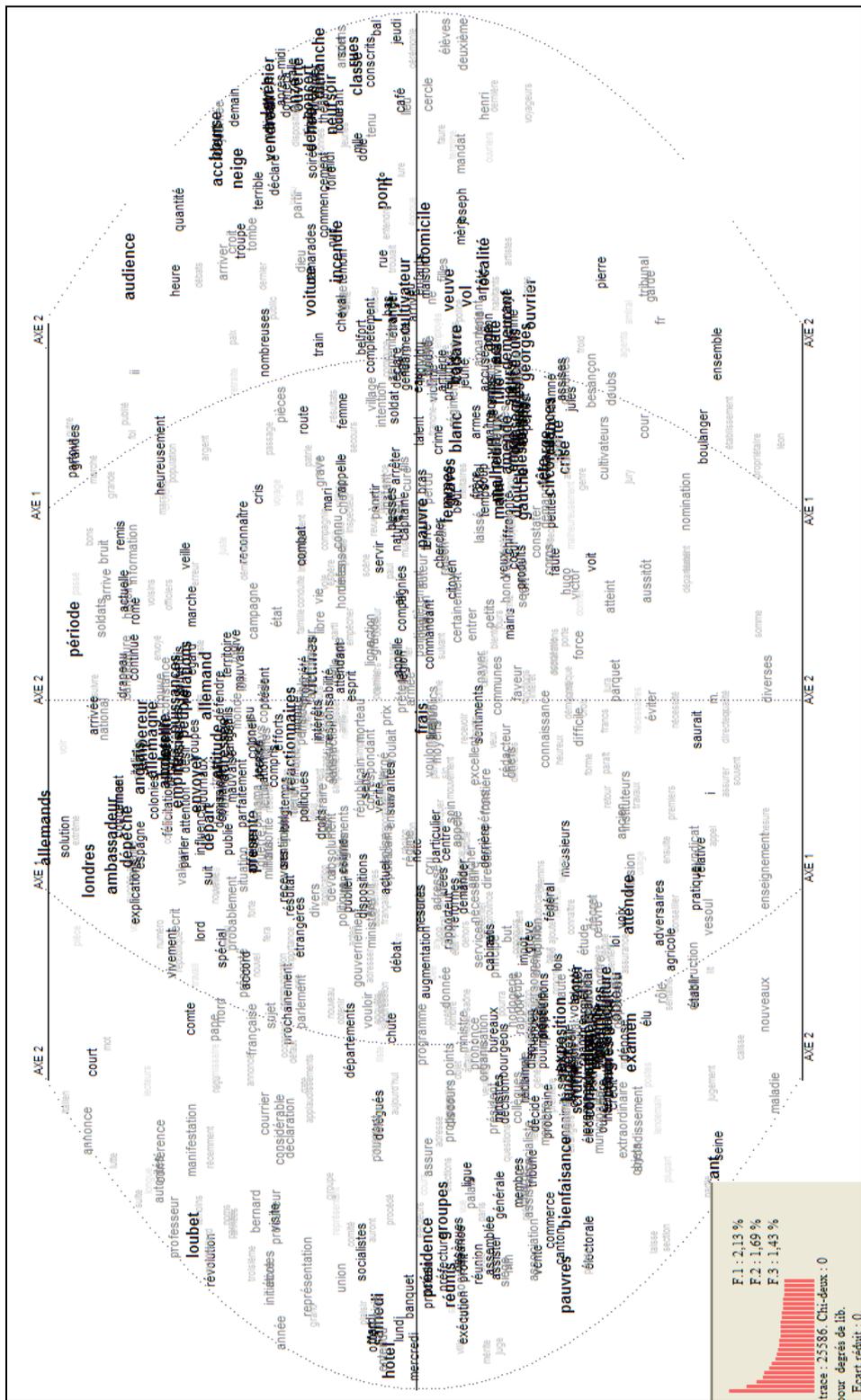


Figure IV-43 : Projection AFC de la configuration fine du vocabulaire de 1896.

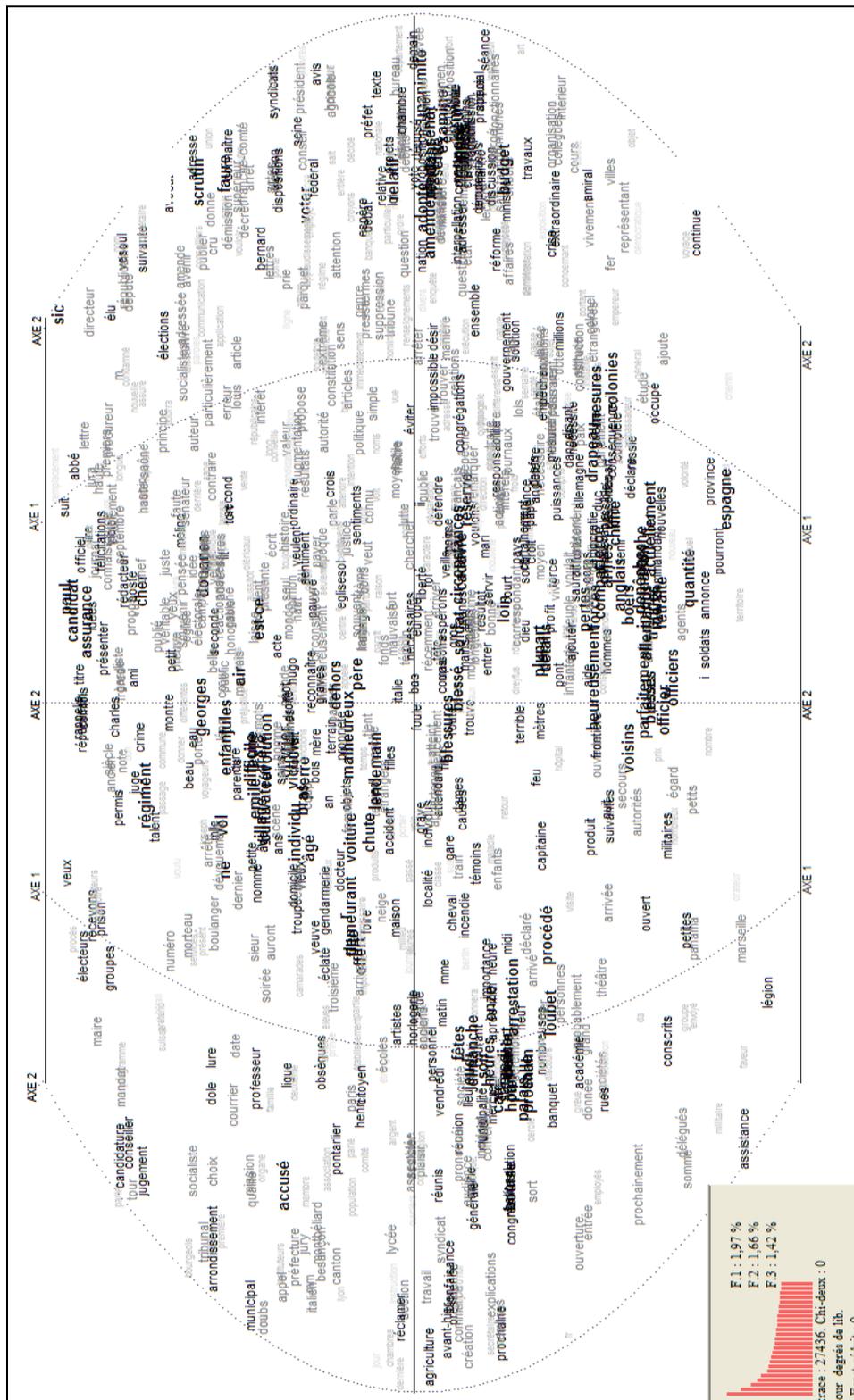


Figure IV-44 : Projection AFC de la configuration fine du vocabulaire de 1900.

On retrouve sur chacun de ces graphes, plus ou moins nettement, quatre pôles isotropiques. Si l'on procède à un zoom centré sur chacune de ces cartographies du vocabulaire, on observe une remarquable similitude avec la configuration du vocabulaire du corpus total. En témoignent, à titre d'exemple²⁷ les deux figures suivantes, qui sont deux zooms dans la figure IV-44, et qui évoquent très nettement les pôles isotropiques A et C de la figure IV-36 :

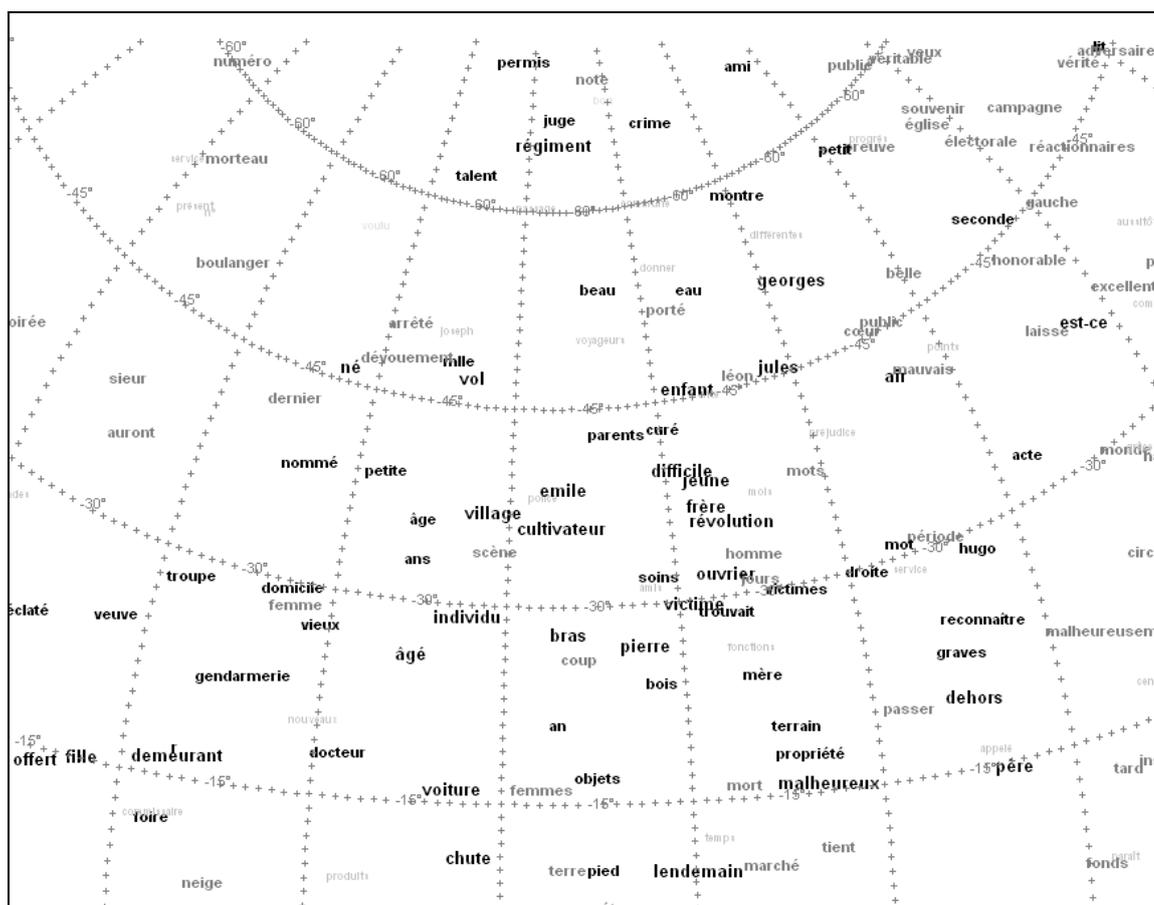


Figure IV-45 : Zoom centré sur l'isotropie A de la configuration fine du vocabulaire en 1900.

²⁷ On trouvera tous les zooms en annexe.

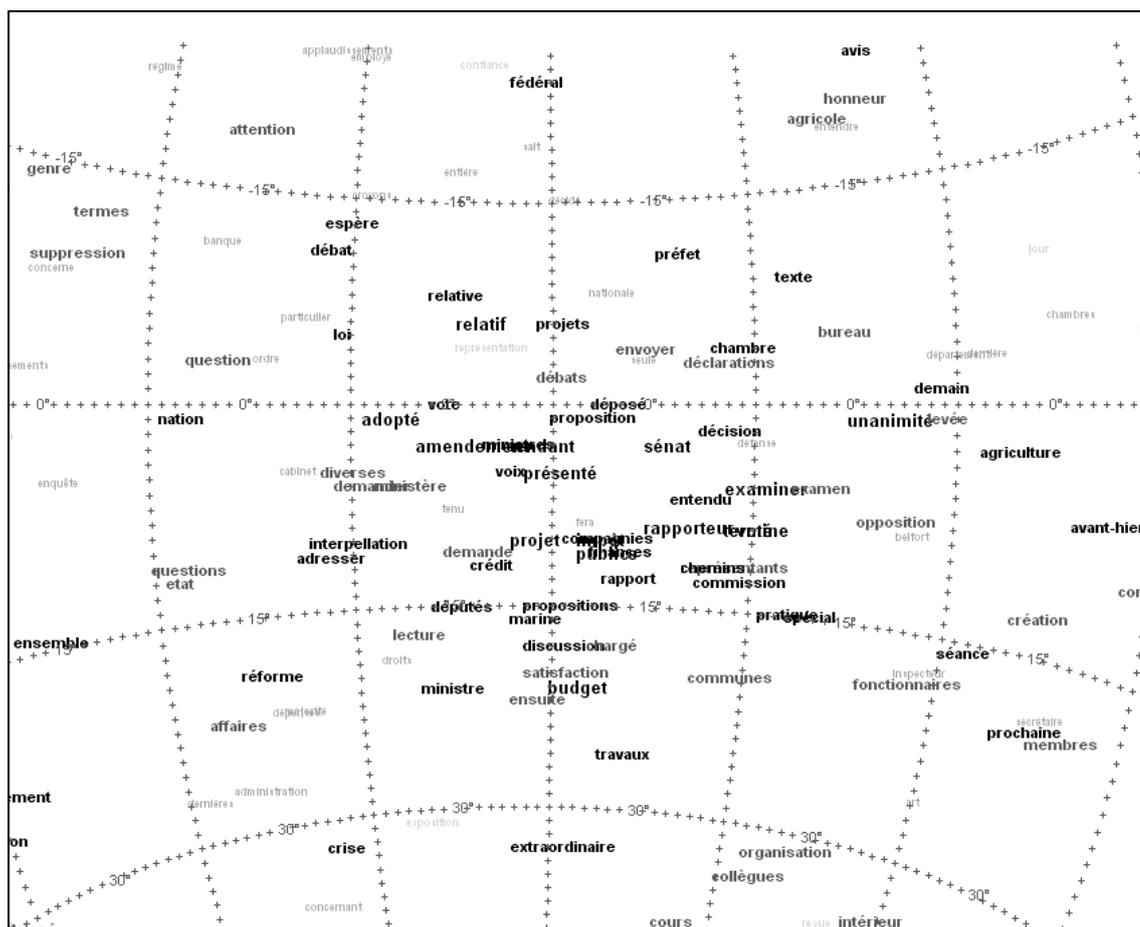


Figure IV-46 : Zoom centré sur l'isotropie C de la configuration fine du vocabulaire en 1900.

4.4.2. Comparaison simultanée

Afin de vérifier la pertinence des indications livrées par les analyses successives de la structure micro-distributionnelle du vocabulaire par année de parution, nous nous sommes tournée vers un mode d'observation AFC visant à évaluer d'un seul coup d'œil l'influence de la diachronie sur la configuration fine du vocabulaire.

À cet effet, nous avons divisé le corpus *Petit Comtois* en deux ensembles textuels équilibrés, l'ensemble A correspondant à la première moitié diachronique du corpus (1883-1893), l'ensemble B à la seconde (1893-1903).

Un planisphère AFC de la cooccurrence généralisée des 1000 premiers items du corpus pour chaque moitié diachronique a été élaboré. Nous avons ensuite attribué une couleur aux items selon leur appartenance aux pôles isotropiques de l'ensemble A ; puis nous avons projeté cette coloration sur le planisphère de micro-distribution de B, paramétrée à l'identique (liste d'items commune, empan cotextuel semblable).

Cette carte laisse apparaître une relative invariance de la texture lexicale du *Petit Comtois* où l'on retrouve de façon assez stable les pôles isotropiques observés sur l'analyse menée sur la totalité du corpus. Selon ce planisphère colorisé, il apparaît donc que la structure fine du vocabulaire ne subit pas une influence importante de la diachronie. La coloration des items selon leur position dans la première moitié du corpus recoupe en effet très majoritairement leur position dans la deuxième tranche du corpus, ce qui n'exclut cependant pas qu'un certain nombre d'items subissent des migrations relatives plus ou moins significatives. Seule une minorité de formes lexicales migrent de façon repérable entre la première tranche diachronique et la deuxième : c'est notamment le cas de l'item *ouvriers*, qui relève du pôle « rouge » sur la période 1883 à 1893 (quart-ouest de la moitié droite du planisphère) et du pôle « rose » de 1893 à 1903 (quart-ouest de la moitié gauche du planisphère). Que signifie une telle migration dans l'espace géodésique ? Comment l'analyser ?

Le cotexte étant le premier palier en vue d'une étude sémantique d'une forme, il est légitime de poser que les migrations observées dans l'espace de la classification de la structure co-occurentielle du vocabulaire sont l'indice d'une évolution sémantique de la forme considérée. Il s'agit donc de saillance sémantique. Pour illustrer notre propos, nous concentrons notre attention sur une forme pointée par les outils statistiques comme saillante, la forme *ouvriers*.

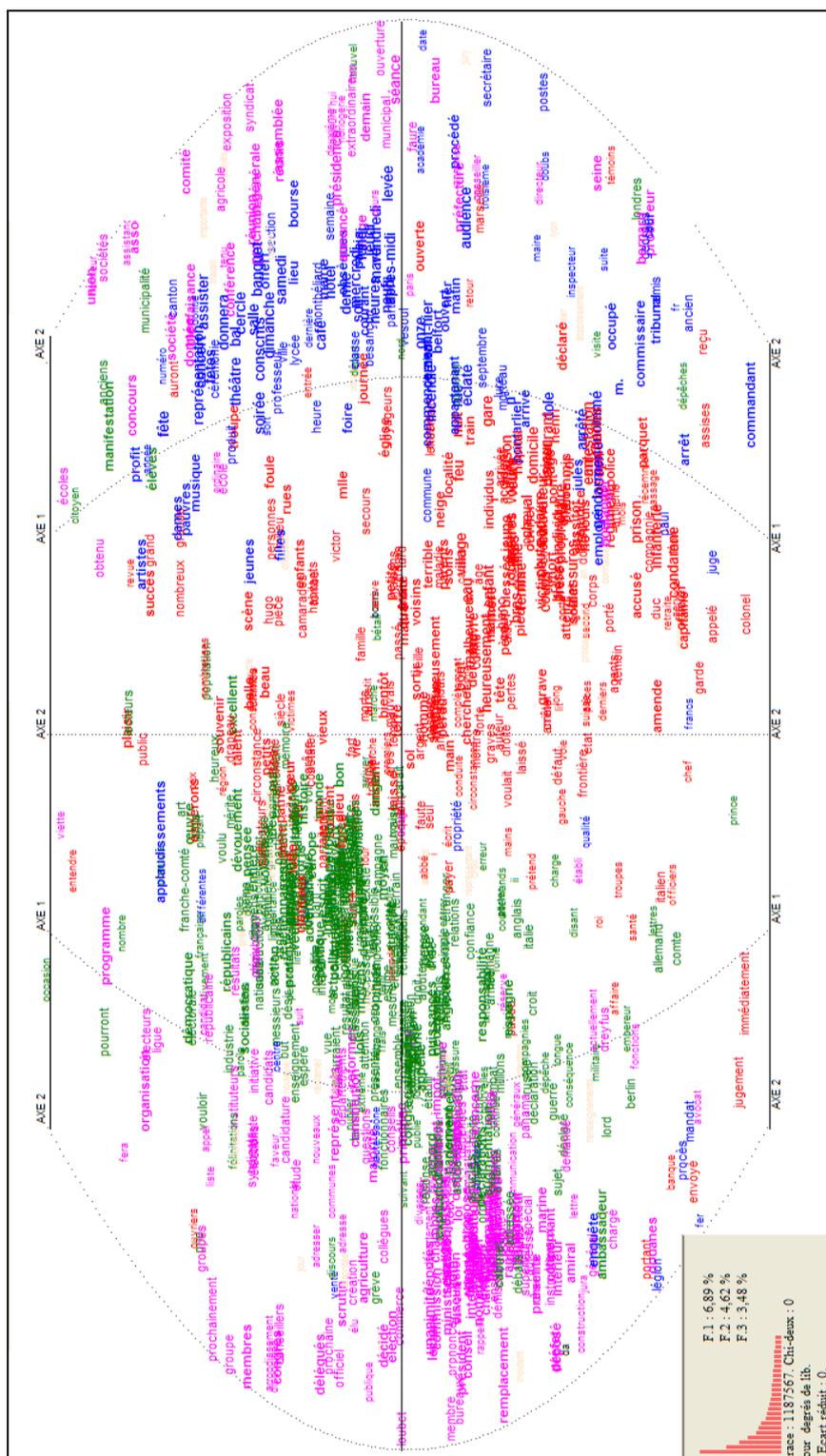


Figure IV-47 : Comparaison de la configuration micro-distributionnelle du vocabulaire de l'ensemble diachronique A (1883-1893) à l'ensemble B (1893-1903).

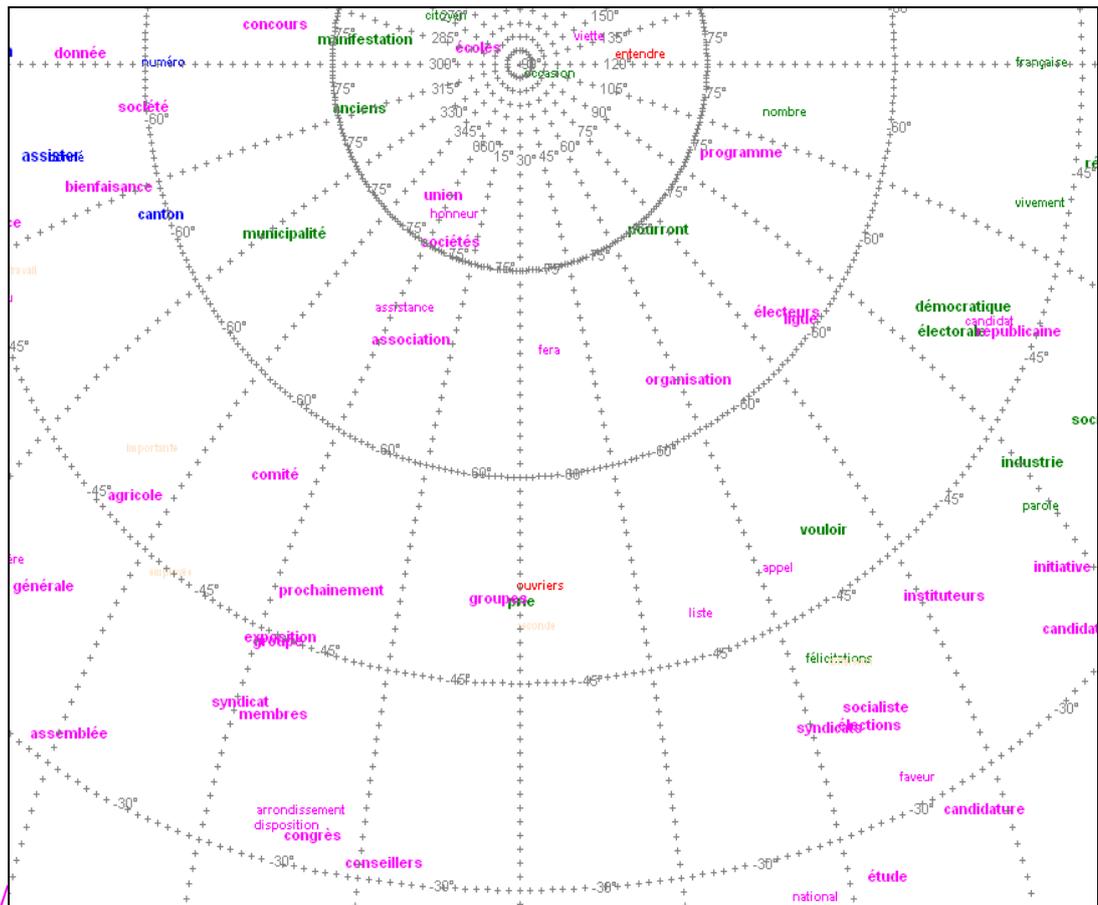


Figure IV-48 : Zoom sur la forme *ouvriers*.

La figure IV-48 met en relief la « migration » de l’item « *ouvriers* », qui se révèle confirmée par un examen du zoom centré sur sa position dans la configuration fine du vocabulaire de la période 1883-1893 (figure IV-49), où il a pour voisinage les formes évocatrices des faits divers *accusé*, *circonstances*, *témoin*, etc.

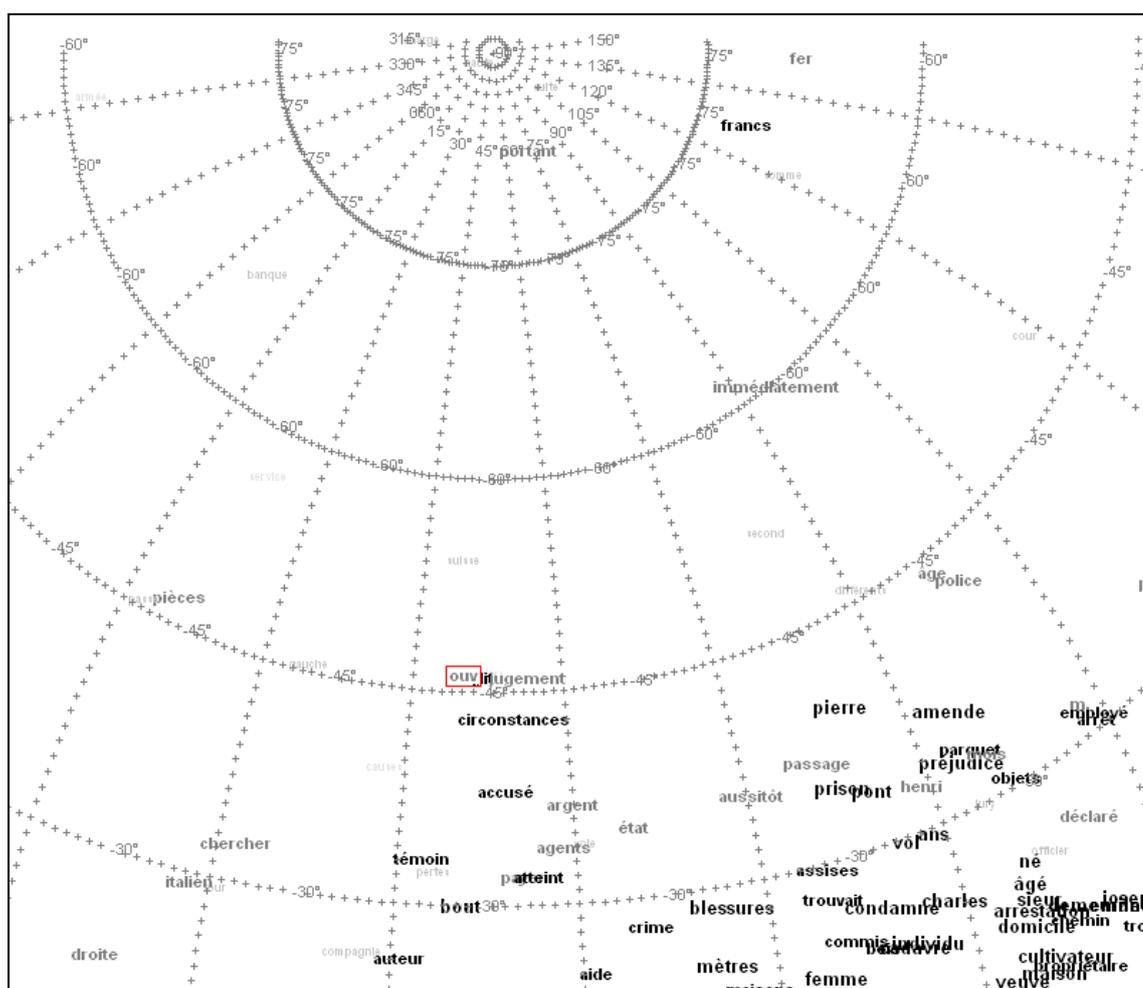


Figure IV-49 : Zoom sur *ouvriers* dans la configuration fine de A (1883-1893).

L'examen des co-occurents les plus forts de la forme *ouvriers* de 1883 à 1893 et de 1893 à 1903 permet de prendre la mesure de l'évolution de ses contextes d'emploi durant une période porteuse de transformations décisives pour la société ouvrière.

EC.RED	OCC.	FORME
53,97	36	patrons
44,15	86	travail
31,85	18	salaires
29,09	20	mineurs
26,45	22	grève
25,36	8	monteurs
22,96	9	travaillent
21,91	12	syndicale
21,18	11	mines
19,25	11	ateliers
17,18	8	salaire
16,71	13	fabrique
15,8	10	ouvrières
14,78	9	usines
14,7	14	italiens
14,19	11	syndicats
13,56	10	usine
12,31	8	grévistes
12,01	8	boîtes
10,5	13	employés
10,09	9	pain
9,18	14	délégués
8,93	49	sans
8,78	10	tarif
7,42	10	allemands
7,32	8	rues
7,02	8	aide
6,21	15	réunion
5,98	10	industrie
4,91	9	secours
4,83	13	prix

Tableau IV-5 : Liste des co-occurents forts de la forme *ouvriers* sur la période 1883 à 1892.

EC.RED	OCC.	FORME		EC.RED	OCC.	FORME
76,9	86	patrons		12,22	13	patron
49,82	65	grève		11,97	11	grévistes
46,84	39	mineurs		11,29	9	mutuels
36,84	23	syndiqués		10,89	15	fédération
35,98	47	syndicats		10,47	12	établissements
35,5	104	travail		9,63	20	secours
31,82	11	métallurgistes		9,12	12	villes
29,13	23	retraites		8,67	22	générale
27,88	47	employés		8,58	12	bourse
26,33	19	salaire		8,52	21	groupe
25,31	8	arsenaux		8,4	9	délégation
25,05	12	manufactures		8,33	14	industrie
22,12	39	syndicat		8,17	22	travaux
21,9	12	corporation		8,08	15	profit
20,53	11	travaillent		7,93	8	accidents
20,08	14	conciliation		7,74	8	renvoi
19,47	12	salaires		7,57	12	art
19,46	16	usines		7,39	9	délégués
19,35	11	chômeurs		7,34	9	demandes
19,28	9	quais		7,27	11	retraite
19,28	12	syndicale		7,18	23	français
18,46	23	usine		6,93	9	travailleurs
18,18	13	chômage		6,62	10	sort
17,95	10	chantiers		5,95	8	augmentation
17,67	8	manufacture		5,92	24	etat
16,35	17	port		5,83	8	cultivateurs
14,45	13	ouvrières		5,46	9	réunis
14,42	8	maritimes		5,44	11	assister
13,93	10	ateliers		4,91	9	organisation
13,76	8	forges		4,75	30	loi
13,01	12	arbitrage		4,65	9	nationale
12,82	12	industriels		4,46	10	force
12,41	13	mines		4,2	12	intérêts

Tableau IV-6 : Liste des co-occurents forts de la forme *ouvriers* sur la période 1893-1903

On remarquera qu'en tête de ces listes, tant sur la période 1883-1892 que sur la période 1893-1903, figure en commun le co-occurent *patrons*, témoin de la dialectique du capital et du travail sous-jacente. Au-delà de cette constance, la plupart des co-occurents de la forme divergent d'une tranche diachronique à l'autre et reflètent la naissance d'un mouvement revendicatif ouvrier de masse,

qui subit la dépression économique de fin de siècle et dans lequel se répandent les idées révolutionnaires.

Le voisinage lexical de l'item *ouvriers* donne tout d'abord à lire la *grève* comme fait social qui se massifie de la première tranche diachronique à la seconde. Indice des mutations sociétales à l'œuvre à la fin du siècle, l'apparition dans la deuxième partie du corpus en tant que co-occurent de la forme *ouvriers* du terme *chômeurs* est des plus intéressante à observer. Les 47 occurrences du terme *chômeurs* se répartissent exclusivement entre 1895 et 1903, et s'imposent pour référer à ceux qu'on désignait par le syntagme nominal *ouvriers sans travail* de 1883 à 1892.

Ch.IV (71), Dépêches de nuit, « Les réunions publiques », (1884-01-15)

Vous trouverez dans les journaux du matin des détails sur les réunions bonapartistes et sur la réunion des ouvriers sans travail. Cette dernière cause un très vif scandale. Le *Temps*, qui l'apprécie dans un article spécial, donne les noms de six personnes arrêtées à la sortie.

Ch.IV (72), Dernière heure, « Manifestation pacifique », (1902-03-02)

Rouen, 1^{er} mars. Trois cents ouvriers chômeurs manifestèrent pacifiquement cet après-midi; la police les dispersa.

Pour sa part, le substantif *chômage*, dont on dénombre 25 occurrences de 1883 à 1892, ne devient un élément privilégié de l'environnement lexical de l'item *ouvriers* qu'à partir de 1893 :

Ch.IV (73), Chronique régionale Doubs, « Libéralité », (1902-03-26)

La loge maçonnique de Besançon a versé à la recette du bureau de bienfaisance, à la mémoire de M. Belin, enterré civilement le 17 mars courant, une somme de 100 fr. pour les ouvriers victimes du chômage.

La commission administrative du bureau de bienfaisance adresse ses remerciements aux généreux donateurs.

L'environnement lexical de l'item fait la part belle aux formes d'organisation du mouvement ouvrier dont l'identité se cristallise à la fin du XIX^e siècle. La comparaison des deux listes est à ce titre très significative. Sur la période 1883-1893, au 8^{ème} rang des co-occurents d'*ouvriers*, figure la forme *syndicale*, qui précède très nettement l'item *syndicats* dans l'ordre des co-occurents

d'ouvriers sur la diachronie considérée. Cette indication reflète donc le passage d'une forme embryonnaire d'organisation, celle de la *chambre syndicale*, seule structure institutionnelle alors permise par le régime politique de 1883 à 1892, à la création de *syndicat* dont la forme s'impose à partir de 1893.

Ch.IV (74), Chronique Régionale, Jura, (1884-01-16)

Les membres de la chambre syndicale des ouvriers travaillant l'article de Saint-Claude sont convoqués en assemblée générale, pour le dimanche 27 janvier courant, à deux heures et demie précises, au siège social, café Regad.

Ordre du jour : Compte rendu des Opérations de l'année 1883.

Renouvellement du bureau.

Les nouvelles adhésions seront reçues.

La tendance s'inverse dans la deuxième tranche diachronique du corpus. La forme adjectivée « syndicale » a cédé la place au substantif, dont on relève un nombre conséquent d'occurrences, tant au pluriel qu'au singulier, dans le voisinage d'*ouvriers*.

Ch.IV (75), Chronique régionale, Doubs, Besançon, « Ouvriers en cuirs », (1894-02-03)

La commission administrative du syndicat des ouvriers en cuirs est priée d'assister à la réunion qui aura lieu aujourd'hui 3 courant, à huit heures du soir, Bourse du travail, rue des Chambrettes. Urgence.

Ch.IV (76), Chronique régionale, Doubs, Besançon, (1900-01-17)

La commission du syndicat des ouvriers maçons de Besançon invite tous ses adhérents syndiqués et non syndiqués à assister à la conférence faite par M. Martin, inspecteur du travail, qui aura lieu le samedi 17 mars, à huit heures et demie du soir, salle Saint-Esprit.

M. l'inspecteur du travail se bornera surtout à expliquer la loi sur les accidents et à bien faire comprendre tous les moyens dont on peut user pour obtenir ce à quoi l'on a droit en cas d'accident.

Vous êtes priés, camarades, de vouloir bien assister en nombre, car la conférence aura un caractère très instructif pour les travailleurs qui tiennent à se renseigner complètement sur la nouvelle loi relative aux accidents du travail.

Le secrétaire, F. L.

Cette organisation du mouvement ouvrier dans la diachronie 1892-1903 se donne également à lire via la présence des items *corporation*, *fédération*, *organisation*, *bourse et secours* dans le voisinage lexical de la forme *ouvriers*.

Ch.IV (77), Chronique régionale, Doubs, Montbéliard, (1898-02-08)

Je vous dirai que depuis longtemps déjà nous avons nos regards tournés vers cette région industrielle et laborieuse, souhaitant ardemment que des circonstances particulières nous permettent d'y venir apporter notre expérience d'organiseurs des syndicats ouvriers. Malheureusement, jusqu'ici l'organisation des groupements ouvriers de notre ville en Fédération, en Bourse du travail, et la lutte incessante soit pour le maintien de leurs conditions de travail ou leur amélioration avaient absorbé tous nos instants, toutes nos ressources. [...]

Cette institutionnalisation s'accompagne d'une diversification des revendications ouvrières présentées d'une tranche diachronique à l'autre. Il est tout d'abord intéressant de relever l'inversion du rapport singulier/pluriel pour la forme « salaire » de la première tranche diachronique à la seconde, ce qu'on pourrait supposer être un indice de la constitution identitaire du groupe ouvrier.

Ch.IV (78), Dépêches de nuit, « Désordres à Decazeville », (1886-01-28)

Une grève accompagnée de désordres graves vient d'éclater dans le centre minier de Decazeville (Aveyron). Le motif de cette grève vient de la question des salaires et du ressentiment de certains ouvriers contre le sous-directeur de l'exploitation. [...]

Ch.IV (79), Chambre des députés, (1896-01-22)

Séance du 21 janvier.

[...] Pour remédier aux causes du chômage, il faut réglementer la journée de travail, fixer un minimum de salaire et faire aux ouvriers des conditions normales. Il faut constituer une caisse de secours et d'assurance contre le chômage. L'Etat et les communes devraient donner l'exemple. [...]

Ch.IV (80), Chronique régionale, Doubs, Montbéliard, « Constitution d'un syndicat ouvrier », (1898-02-08)

[...] Pour relever à un taux légitime le salaire de toutes les parties de notre industrie, de tous les ouvriers, de ceux-là même qui sont dans les ateliers, aussi qui, comme nous, ne sont pas assez rétribués, il nous faut cette union, cette solidarité morale et matérielle, le dévouement, l'énergie de vous tous, sans s'occuper des conséquences qui en résulteront et qui seront, j'en suis certain, non seulement à l'avantage des ouvriers et des patrons, mais encore l'intérêt de tout le monde, de toutes les industries, à quelques branches qu'elles appartiennent.

Dans un interdiscours chargé de sens, il est tout aussi stimulant d'observer la relation de co-occurrence entre *pain et ouvriers*. Celle-ci renvoie en effet à la coordination « du pain et/ ou (du) X », mot d'ordre révolutionnaire du XVIII^e siècle étudié par J. Guilhaumou et D. Maldidier (1986). Réactivant la mémoire

révolutionnaire, ce mot d'ordre canalise les revendications à la fois symboliques et concrètes du mouvement ouvrier :

Ch.IV (81), Dépêches de nuit, « Agitation en Italie », (1889-01-20)

Rome, 19 janvier. Les manifestations de la misère se multiplient en Italie.

A Ferrare, les ouvriers parcourent les rues de la ville, demandant du travail ou du pain. On compte dans cette province plus de 7,000 ouvriers sans travail.

A Parme, les ouvriers ont également parcouru les rues, en silence, portant un immense écriteau sur lequel étaient écrits ces mots ; « Pain et travail. »

La police a arrêté plusieurs des manifestants. Les troupes sont intervenues. Des perquisitions ont été faites dans plusieurs auberges.

Propriété, de même que *ouvriers*, est un item dont le profil collocatif varie fortement de la première tranche diachronique à la seconde, comme le suggère la migration de la forme de presque 90 degrés.

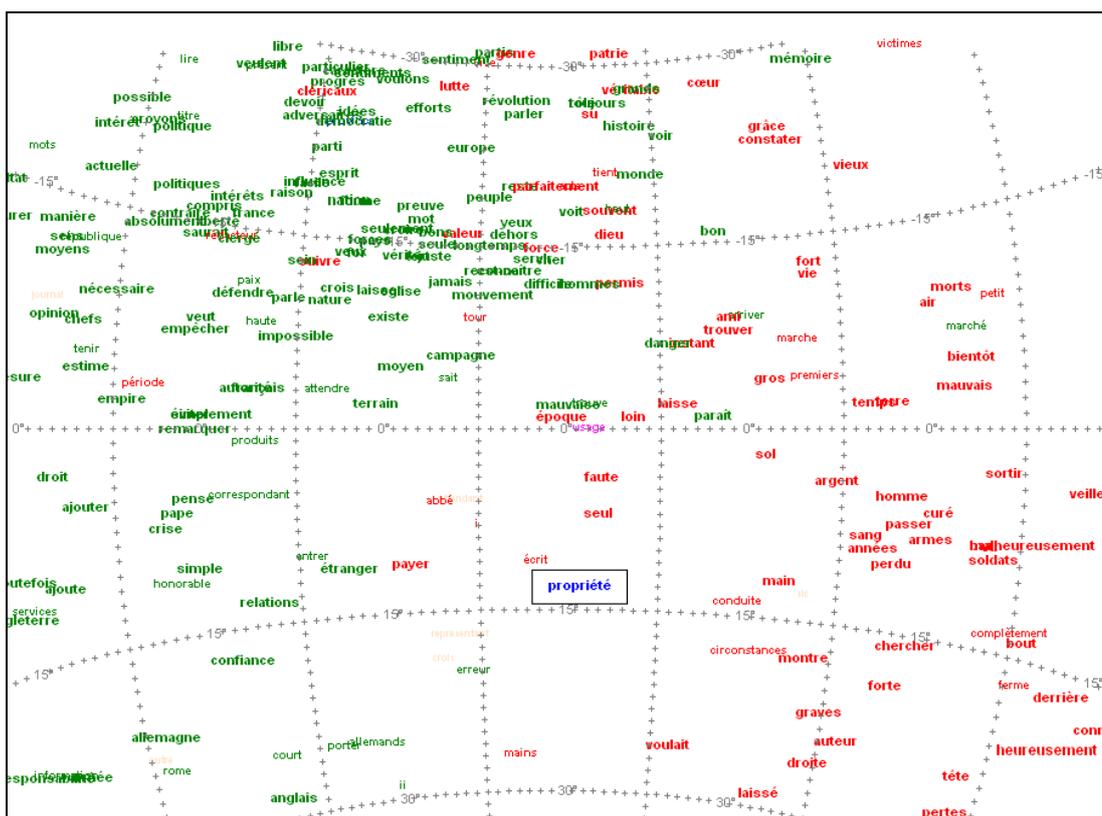


Figure IV-50 : Zoom sur l'item *propriété*.

De 1883 à 1892, les 128 occurrences de *propriété* ont principalement pour environnement lexical les formes *propriété*, *incendie* et *lieu*. Ces indications nous suggèrent d'interpréter grossièrement *propriété* comme synonyme de *domaine* ou de *demeure*.

EC.RED	OCC.	FORME
25,85	10	propriété
11,66	9	incendie
10,55	153	la
5,09	162	de
4,48	35	dans
4,13	10	ou
2,79	31	une
2,5	6	lieu
2,45	12	aux
2,35	38	a
2,16	13	cette
2,11	16	sur

Tableau IV-7 : Les co-occurents de *propriété* de 1883 à 1892.

La consultation des contextes d'emploi de l'item confirme cette première piste. Utilisé principalement dans le rapport de faits-divers, incendie ou découverte macabre, *propriété* désigne très massivement, de 1883 à 1892, un domaine ou une demeure :

Ch.IV (82), Chronique régionale, Doubs, Besançon, « Marchaux. Incendie », (1891-01-09)

Un incendie a éclaté, le 2 janvier, dans la propriété de MM.Cailloz, Claude-Joseph, et Mignot, Jean-Pierre, cultivateur à Marchaux. Les deux maisons continues et les récoltes ont été complètement la proie des flammes.

Les pertes causées par ce sinistre, dont les causes sont inconnues, peuvent être évaluées à la somme de 10,000 fr. pour la maison Cailloz et à 4,500 fr. pour la maison Mignot.

Il n'y a pas eu d'accident de personne.

Ch.IV (83), Chronique régionale, Cour d'assises du Doubs, (1891-01-06)

Le 9 juin 1890, vers cinq heures du soir, deux habitants de la commune de Chantrans, les sieurs Rozet père et fils, revenant de leur travail, aperçurent derrière un buisson, dans une propriété sise au lieu dit le Pré au Tartre, un cadavre qu'ils reconnurent pour être celui du sieur Magnenet, maçon, demeurant à Chantrans. Le corps, portait deux blessures, l'une au bras droit, l'autre dans le dos, paraissant avoir été produites par le passage d'un projectile de forme ronde. Le sieur Magnenet, qui tenait encore une serpe à la main, semble avoir été surpris au milieu de son travail ; une traînée de sang indique qu'il avait dû faire quelques pas avant de tomber.

Manifestement, sur cette période, la référence à la propriété s'inscrit du côté des faits divers et ne fait pas l'objet d'un débat dans les colonnes du *Petit Comtois*. Ce débat est au contraire éludé. Les seules références à la propriété, en tant que principe d'un système économique, sont liées à un procès à l'encontre du journal *La République démocratique et sociale*, accusée du délit de provocation au meurtre, pour avoir cité le célèbre mot de P.-J Proudhon et avoir déclaré « qu'une exécution de la bourgeoisie est nécessaire ». *Le Petit Comtois* condamne ces « théories » qu'il juge « insensées » et s'insurge contre la faiblesse des sanctions à l'égard de M. Tallandier, directeur du journal, de laquelle découle « un encouragement à de plus grandes violences de langage, en attendant qu'il se rencontre quelque nouveau Curien pour passer à la propagande par le fait. » (Bulletin, 1884-01-23) :

Ch.IV (84), « Pro domo », (1884-01-10)

[...]Ce journal [*La République démocratique et sociale*], qui paraît depuis peu de jours, est poursuivi pour excitation au meurtre et outrage aux bonnes mœurs.

L'excitation au meurtre résulte d'un passage dans lequel M. Morphy, reprenant le mot de Proudhon : « La propriété, c'est le vol, » ajoute : « Mort aux voleurs ! » et déclare qu'une exécution de la bourgeoisie est nécessaire :

« Avec la chimie, il n'y a plus de minorité, et les apôtres des revendications du peuple sauront bien, grâce à la dynamite, vaincre la bourgeoisie, sa puissance, sa police et ses troupes. »

On trouvait généralement que l'extrême gauche montrait trop de tendresse pour les dynamitards. Il paraît indispensable, aujourd'hui qu'un membre de ce groupe a laissé émettre des théories aussi insensées dans son journal, que l'extrême gauche désavoue une bonne fois les partisans de la propagande par l'action. [...]

Les co-occurents de « propriété » privilégiés de 1893 à 1903 sont tout différents et nous laissent apercevoir un parcours interprétatif particulièrement suggestif du terme. Au vu de ces indications statistiques, c'est un autre contexte d'emploi qui transparait : le discours du *Petit Comtois* réfère à la propriété en tant que socle d'un système politique que le courant socialiste se donne pour mission de supprimer.

EC.RED	OCC.	FORME
74,84	9	bâtie
57,03	17	individuelle
34,29	10	industrielle
27,46	16	propriété
18,34	319	la
15,75	10	suppression
5,43	23	sa
4,9	263	de
4,82	11	loi
4,16	15	dont
4,15	10	france
4,14	17	ou
3,83	8	droit
3,57	8	non
3,52	8	toute
3,3	104	et
3,09	46	dans
3	14	ils
2,94	51	une
2,85	57	en
2,73	38	qu'
2,6	54	il
2,47	26	sur
2,39	54	un
2,34	38	pour
2,31	67	des
2,18	13	ces

Tableau IV-8 : Les principaux co-occurents de propriété de 1893 à 1903.

Les co-occurents privilégiés de *propriété*, à savoir *individuelle* et *suppression*²⁸, ainsi que l'accroissement très net de son nombre d'occurrences dans la deuxième tranche diachronique de notre corpus, où l'on dénombre 226 cas d'emploi de la forme contre 128 de 1883 à 1892, reflète l'intense circulation des théories socialistes durant cette période, du moins dans *Le Petit Comtois*. Qu'il s'agisse de le définir et de le défendre, comme dans les passages (85) et (86), ou de le railler lors d'un rapport de « réunion socialiste », comme dans l'extrait (87), le principe de la *suppression* de la propriété *individuelle* est abondamment évoqué de 1893 à 1903, reflétant les débats de l'espace public :

Ch.IV (85), « Une réponse », (1893-01-28)

[...] Je protesterai d'abord contre l'épithète de socialiste décernée à l'école catholique. Le mot socialiste ne doit s'adresser qu'aux écoles qui ont comme base doctrinale la suppression de la propriété individuelle et son remplacement par la propriété collective. Ainsi les vrais socialistes sont les partisans de la doctrine de Karl Marx. Les écoles qui maintiennent la propriété individuelle, en remaniant plus ou moins profondément l'ordre social actuel, ne sont pas socialistes : leurs adeptes sont des économistes plus ou moins avancés. [...]

Ch.IV (86), Chronique régionale, Doubs, Besançon (1898-01-01)

Le groupe communiste nous prie de publier le document suivant : [...] Notre programme, qui se résume en quelques mots, est le même : Transformation de la propriété individuelle en propriété commune. — Les instruments de travail, à la disposition des travailleurs. — Gouvernement direct du peuple par lui-même. — Lutte des classes. — Entente internationale des travailleurs. Quel est le socialiste, le militant qui pourrait ne pas l'accepter et ne pas s'y rallier ? Chacun de nous ne l'a-t-il pas défendu et ne le défend-il pas chaque jour contre les calomnies des adversaires du socialisme ?

Ch.IV (87), Chronique régionale, Haut-Rhin, « Réunion socialiste », (1893-01-01)

[...] Le citoyen Roussel est collectiviste et pour lui c'est le patronat qu'il faut avant tout faire disparaître. Les ouvriers se formeront partout en syndicats et supprimeront la propriété individuelle. La révolution de 1889 a été faite au profit des bourgeois. Celle qui se prépare se fera au profit des ouvriers.

En ce qui concerne le mot révolution, M Roussel tient à s'expliquer. Par le mot révolution, il n'entend pas la descente dans la rue, les barricades, les coups de

²⁸ Bien qu'il présente un écart-réduit très fort (34,29), l'item *industrielle* ne connaît que 10 occurrences, dont 9 sont concentrées dans un même article, visant à présenter l'« Association des ingénieurs-conseil en matière de propriété industrielle » (Chronique régionale, (1903-03-25). Aussi, tendrions-nous à nuancer la significativité de cette co-occurrence. Il en va de même pour la forme *bâtie*, renvoyant à la réforme législative de la loi sur la *propriété non bâtie*.

fusil. La révolution dont il veut parler se fera pacifiquement. Pendant qu'il y était, M. Roussel aurait bien dû nous dire comment, puisque cette révolution dont il parle doit se faire pacifiquement, comment il arrivera « pacifiquement » à supprimer la propriété individuelle; car M. Roussel conviendra bien avec nous que ceux qui possèdent ne se laisseront pas déposséder sans crier comme des paons et qu'il faudra peut-être user de moyens un tantinet violents pour leur faire rendre gorge. Il est bien fâcheux que M. Roussel ne se soit pas expliqué davantage sur ce point. [...]

L'évolution du profil de l'item « citoyen », migrant d'approximativement 60 degrés sur le troisième facteur, reflète le trajet de cet item dans la deuxième partie diachronique de notre corpus.

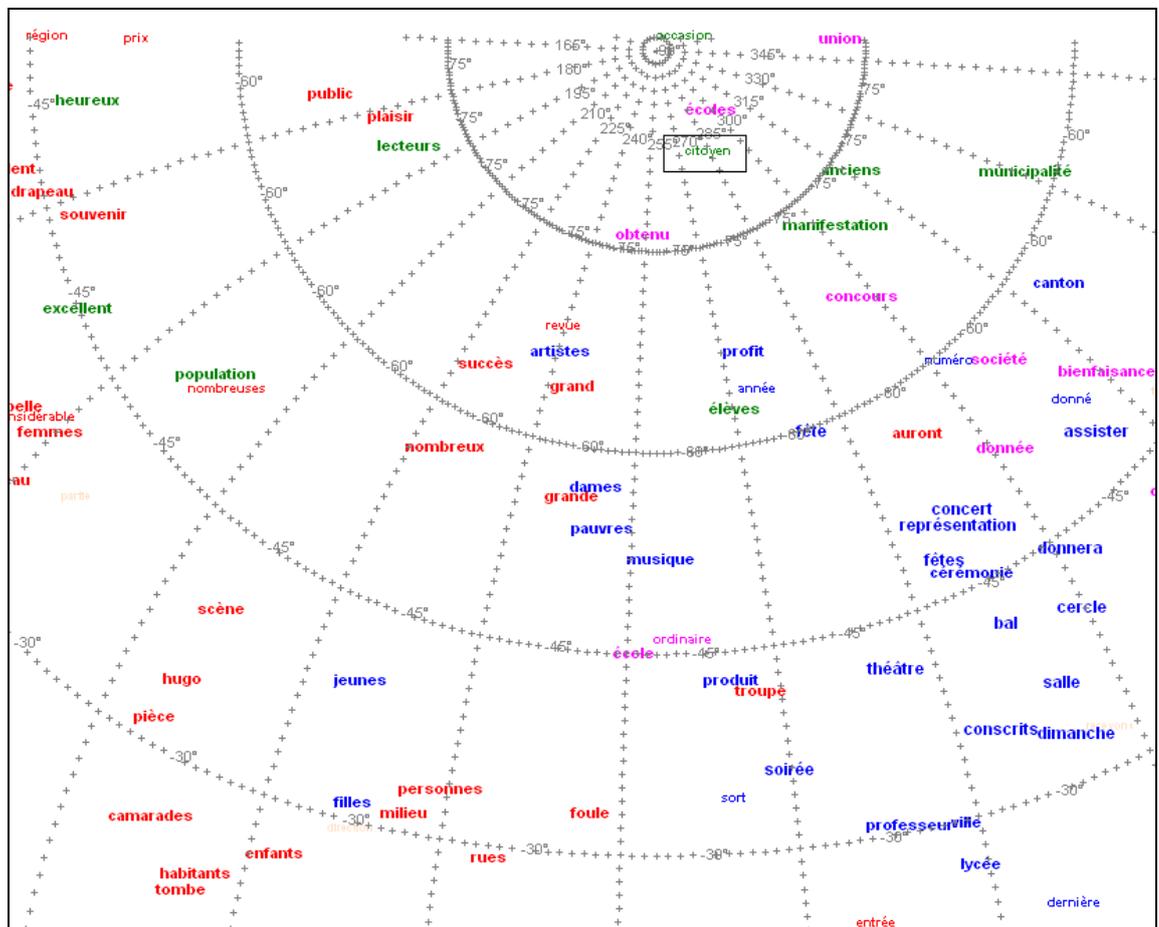


Figure IV-51 : Zoom sur l'item *citoyen*.

De 1883 à 1892, les 108 occurrences de l'item *citoyen* ont pour principaux co-occurents l'adjectif *bon* (18,79) et le substantif *homme* (11,41). Le citoyen se présente ici avant tout comme membre actif de la République : on remarquera d'ailleurs que les *bons citoyens*, auxquels *Le Petit Comtois* rend hommage, sont avant tout des personnalités politiques ou publiques républicaines.

Ch.IV (88), Chronique régionale, Saint-Hilaire, (1883-08-17)

[...] Républicain sincère, il est mort en bon citoyen, en honnête homme, emportant tous nos regrets.

Cet administrateur dévoué a rencontré en maintes circonstances de nombreux obstacles et une résistance opiniâtre, dont vous connaissez la source, et dont il n'a pu triompher qu'à force d'énergie et de persévérance. [...]

Ch.IV (89), « Le centenaire et nos écoles », (1888-01-27)

[...] Sans instruction, l'enfant ne peut devenir un homme vraiment utile, un citoyen digne de ce nom. Songeons donc à nos enfants ! Sans instruction, un peuple ne peut s'élever jusqu'à la philosophie et jusqu'à la liberté; il est fatalement la proie de ceux qui le gouvernent; il ressemble à ces esclaves que les anciens aveuglaient pour en faire des instruments plus dociles. [...]

Il transparaît là une conception de la citoyenneté dans laquelle le *citoyen* est aussi un sujet pourvu de droits comme le rappelle les passages suivants, dans lesquels l'adjectif indéfini *tout* met en exergue l'égalité prévalant entre citoyens :

Ch.IV (90), « Le parlementaire », (1889-01-23)

[...] Sous le régime parlementaire, tout citoyen a le droit d'élever la voix; il a des mandataires qui peuvent porter à la tribune ses doléances ou ses griefs; il a des journaux où il peut exprimer sa pensée, formuler ses revendications, saisir l'opinion de ses plaintes. [...]

Ch.IV (91), « Notre politique », (1883-08-12)

[...] Nous voulons pour tout citoyen la liberté de dire ce qu'il croit vrai, de le répandre, de le propager ; nous voulons la liberté sous toutes les formes; mais nous voulons l'autorité dans le gouvernement, parce qu'elle est la garantie de la liberté de chacun.

Nous aussi, nous voulons un gouvernement ferme et armé pour faire respecter la liberté et pour faire respecter la République par ses ennemis d'en haut et par ses ennemis d'en bas. [...]

De 1893 à 1903, outre un accroissement du nombre d'occurrences du terme par rapport à la première tranche diachronique, on perçoit dans le voisinage

lexical du mot *citoyen*, composé de nombreux noms propres, une profonde évolution du fonctionnement discursif du terme.

EC.RED.	OCC.	FORME	EC.RED.	OCC.	FORME	EC.RED.	OCC.	FORME
18,46	33	homme	9,86	13	donne	4,19	248	la
18,57	15	ligue	9,74	11	secrétaire	3,85	23	comme
16,83	8	tramu	9,46	8	prend	3,63	12	grand
49,13	15	biétry	8,86	14	famille	3,46	8	doit
55,43	8	barrot	8,76	8	pensée	3,45	372	de
26,24	31	droits	7,56	88	au	3,45	7	parti
17,39	22	parole	7,52	14	député	3,39	15	lieu
33,68	12	perrin	7,44	13	ensuite	3,24	16	nos
22,52	22	citoyen	7,15	9	père	3,19	20	faire
16,83	8	civiles	6,96	11	liberté	3,11	172	à
17,9	9	honnête	6,77	8	assemblée	2,98	86	un
46,07	8	chappuis	6,32	8	défense	2,84	9	droit
56,9	18	dequaire	6,13	10	dernière	2,77	7	honneur
26,55	17	beauquier	6,07	73	nous	2,76	28	aux
59,83	20	kermorgant	6,01	16	besançon	2,72	11	soir
18,03	16	socialiste	5,85	8	union	2,48	10	avons
69,3	16	broussouloux	5,83	11	réunion	2,43	7	membres
14,45	15	libre	5,7	8	doubs	2,36	71	qui
14,24	13	obsèques	5,68	9	chaque	2,35	6	mon
13,98	18	groupe	5,39	10	amis	2,35	6	voix
13,03	25	nom	5,25	178	et	2,34	8	février
12,86	10	fédération	5,22	9	publique	2,31	7	lettre
11,97	169	du	5,08	11	fois	2,27	8	fut
10,88	16	conférence	4,95	30	sa	2,2	9	ainsi
10,86	11	assister	4,87	9	comité	2,08	6	travail
10,66	264	le	4,8	12	petit	2,07	5	français
9,93	17	bon	4,73	30	ses			

Tableau IV-9 : Les principaux co-occurents de *citoyen* de 1893 à 1903.

De 1893 à 1903, le mot *citoyen* est convoqué comme appellation dans une proportion significative. On le rencontre principalement dans des rapports de réunions et de conférences socialistes, qui fleurissent dans les colonnes du quotidien, expliquant la prégnance des co-occurents *prend*, *parole*, *assister*, *conférence*, *assemblée*, et *réunion* :

Ch.IV (92), Chronique régionale, Doubs, Besançon, « La conférence du Kursaal », (1897-02-22)

Samedi, à neuf heures moins un quart du soir, le citoyen Dequaire, accompagné des citoyens: Beauquier, député du Doubs ; Baigue, adjoint au maire de Besançon ; Vaillandet, président de la Ligue des intérêts républicains, et Charboutet, prend place sur la scène.

Le citoyen Beauquier prend la parole et dit qu'il y a quelques mois il se trouvait à Saint-Claude (Jura), que dans cette localité, il existe un groupe socialiste révolutionnaire comme à Besançon et que le conférencier, le citoyen Dequaire, qui appartient à l'opinion radicale socialiste, fut écouté par tous avec plaisir et sans interruption, parce qu'on savait que le conférencier est un homme de foi, un propagandiste convaincu et honnête.

L'emploi de *citoyen* comme appellatif nous semble refléter la réactivation et la réappropriation de la figure du citoyen révolutionnaire par les socialistes, sur laquelle ne manque pas d'ironiser *Le Petit Comtois* :

Ch.IV (93), Chronique régionale, Haut-Rhin, « Réunion socialiste », (1893-01-01)

Après plusieurs petites réunions préliminaires qui se sont terminées, non sans discussions très aigres, M. Chappuis — pardon ! le citoyen Chappuis, l'apôtre du socialisme révolutionnaire à Belfort, a voulu frapper un grand coup. [...]

4.4.3. Remarques conclusives provisoires

La classe d'explorations que nous venons de mener nous a autorisée à observer l'évolution de la configuration fine du vocabulaire en diachronie et a pointé un faisceau de pistes réflexives stimulantes en prise, pour la plupart, avec les résultats antérieurs majeurs de l'analyse du discours du côté de l'histoire, appliquée à l'ensemble du long XIX^e siècle. En premier lieu, nous avons pu observer que l'influence de la diachronie sur la structure micro-distributionnelle n'était que très faible. Que les quatre pôles isotropiques observés à partir d'une étude de l'ensemble du vocabulaire du corpus se retrouvent aussi nettement de la première tranche diachronique du corpus à la seconde n'a pas manqué de nous interpeller. On aurait en effet légitimement pu attendre une évolution diachronique très prononcée de la configuration fine du vocabulaire, et cela, en raison de deux facteurs principaux. Tout d'abord, l'examen de la ventilation du vocabulaire en diachronie a pointé une évolution

diachronique très prononcée, dont on aurait pu attendre qu'elle ait une incidence sur le profil micro-distributionnel du vocabulaire du corpus. Ensuite, on aurait également pu supposer qu'un discours médiatique tel que celui du *Petit Comtois*, sur une diachronie somme toute assez large, soit caractérisé par une variation soutenue de ses thèmes et thématisations, résultant notamment plus ou moins directement de l'actualité ou encore de la radicalisation politique de la rédaction, radicalisation avérée dès 1891.

Qu'il n'en soit rien et que les pôles isotropiques observés témoignent d'une structuration thématique aussi stable nous amène à supposer l'existence d'une corrélation entre la configuration fine du vocabulaire et la structuration du quotidien en rubriques. Le rubriquage, en tant qu'il offre un cadre stable et récurrent à la mise en scène d'une actualité par essence fluctuante, doit donc être convoqué au titre de candidat comme facteur structurant de l'organisation réticulaire du texte.

En second lieu, nous souhaitons insister sur l'apport de la comparaison simultanée d'analyses micro-distributionnelles du vocabulaire du corpus de deux ensembles diachroniques distincts pour l'analyse du discours. Cette méthodologie exploratoire permet en effet de mettre en évidence, de façon endogène et non projective, l'évolution sémantique de mots qui se révèlent en l'occurrence des termes-clés du vocabulaire social et politique du XIX^e siècle.

4.5. Le rubriquage, facteur structurant de la configuration fine du vocabulaire.

4.5.1. Méthodologie d'analyse des dominantes lexico-thématiques des rubriques

Dans le but de confirmer l'hypothèse d'une corrélation entre la structuration de la configuration fine du vocabulaire et le système de rubriquage du *Petit Comtois* pointée par les manipulations précédemment exposées, nous

nous proposons de procéder à l'analyse des dominantes lexico-thématiques des rubriques.

Nous avons tout d'abord analysé la ventilation du vocabulaire dans les rubriques les plus importantes du *Petit Comtois*, en construisant une matrice croisant les 1000 formes lexicales les plus fréquentes en lignes avec les rubriques les plus importantes en colonnes.

Le graphique d'AFC résultant de cette manipulation (figure IV-52) laisse apercevoir sur le premier facteur, qui représente 38,77% d'inertie, un très net clivage entre, dans l'hémisphère gauche, l'ensemble des chroniques régionales et les rubriques *Faits divers*, *Tribunaux* et, dans l'hémisphère droit, un pôle regroupant l'ensemble des dépêches et de ses sous-rubriques (*Chambre des députés*, *Intérieur*, *Étranger*, *Sénat*). Le deuxième facteur affine cette configuration en installant dans la moitié gauche de l'hémisphère un gradient entre le groupe *Chronique Régionale*, *Chronique Régionale Besançon* d'une part, et d'autre part, les rubriques *Chroniques Régionale Haut-Rhin*, *Chronique Régionale Haute-Saône* et *Chronique Régionale Jura*, ainsi que *Faits-Divers* et *Tribunaux*. Dans la moitié droite du planisphère, le deuxième facteur intervient principalement pour individualiser la rubrique *Dernière heure* du groupe précédemment décrit, de même que le facteur 3. On notera que ce dernier, dans la lignée du clivage établi par le deuxième facteur dans la moitié gauche du planisphère, contribue à opposer la rubrique *Chronique Régionale Besançon* à la rubrique *Chronique régionale Jura*.

Ensuite, pour mettre en évidence l'existence de thématique(s) intrinsèque(s) à une rubrique donnée par l'interprétation de la structure micro-lexicale du vocabulaire du corpus *Petit Comtois*, nous avons utilisé comme indices de spécificités de chaque rubrique les écarts-réduits à l'équidistribution issus de l'analyse macro-distributionnelle précédente. Un code couleur a ensuite été attribué à l'indice de spécificité, suivant qu'il indiquait un moyen ou un fort (sur-) ou (sous-)emploi. Pour dégager les spécificités lexicales d'une rubrique, nous avons ainsi coloré les items du planisphère de la configuration fine du vocabulaire selon le code couleur suivant :

- En bleu, ont été codés les items fortement sur-représentés dans la rubrique,
- en turquoise, sont signalés les items moyennement sur-représentés ;
- les items fortement sous-représentés sont pour leur part signalés en rouge.
- Enfin, les items codés en rose sont à interpréter comme moyennement sous-représentés dans la rubrique.

4.5.2. Validation de la corrélation entre le système de rubricage et la configuration fine du vocabulaire

Nous avons remarqué que le pôle isotropique A dégagé lors de l'examen de la configuration fine du vocabulaire (figure IV-36) évoque assez explicitement le répertoire des faits divers. Rappelons que ce pôle s'articule autour de formes lexicales qui peuvent légitimement être interprétées comme des actants engagés dans l'action narrée (*femme, homme, mari, mère, fils, enfant, veuve, voisins, ouvrier, sieur, maître, cultivateur, employé, accusé*), des références au corps humain et à sa dégradation (*pied, bras, tête, blessures, sang, cadavre*). Il est aussi des toponymes évocateurs (*prison, parquet, maison, hôpital*) et des familles événementielles (*incendie, feu, accident, chute, crime, vol*), ainsi qu'un ensemble d'adjectifs affectifs et évaluatifs et d'adverbes

(*petit, terrible, grave, pauvre, malheureusement*) et de participe passé évoquant les procédures judiciaires (*arrêté, nommé, déclaré*) dont la co-occurrence résonne comme les traces d'une activité narrative. Pour infirmer ou valider notre hypothèse, nous avons ainsi projeté sur le planisphère de la configuration fine du vocabulaire les items sur-représentés de la rubrique *Faits-Divers*.

La visualisation (figure IV-53) permet alors de vérifier clairement la corrélation entre le pôle isotropique en question et cette rubrique : les formes lexicales spécifiques de cette case rédactionnelle forment un nuage de points très nettement localisé et délimité dans le quart-ouest sud de l'hémisphère droit du planisphère, qui recoupe parfaitement la zone du pôle isotropique A puisqu'on retrouve les items :

- *frère, veuve, mari, mère, fils, femme, frères, père, mari, famille, voisins, crime, victime, chute, mort, vol, incendie, accident, tête, sang, cadavre, arrêté, prison, individu, condamné, commissaire, matin, avant-hier, gare, rue, bout, coup, atteint, village morts coups, blessé, maison, domicile, lit, âgé, nuit, arrivée, commis, trouvé, trouvait, aussitôt, bientôt, argent, arrêter, malheureux, terrible, cris.*

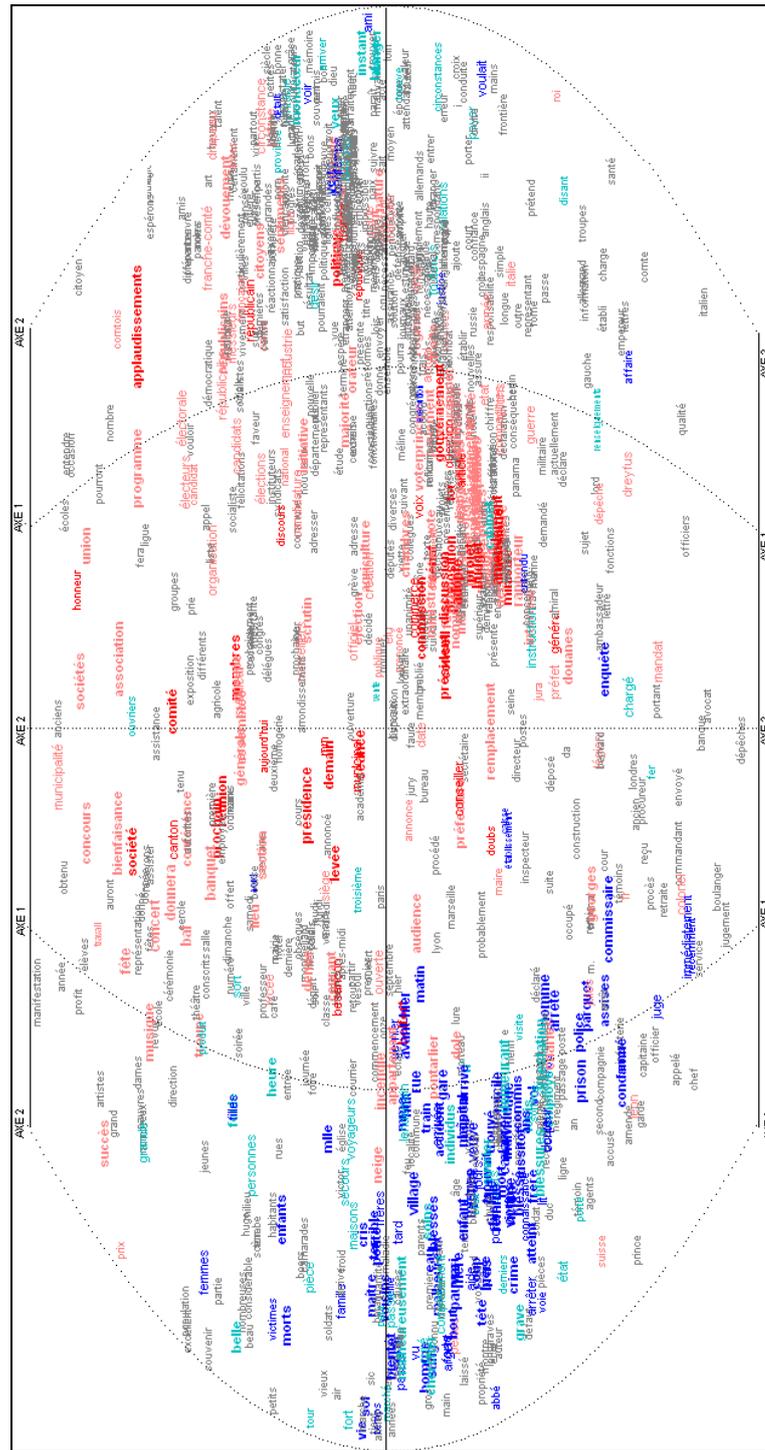


Figure IV-53 : Analyse micro-lexicale des 1000 premiers items du corpus *Petit Comtois* colorée d'après les suremplois dans « Faits-Divers ».

Pour sa part, le pôle isotropique B de la configuration fine du vocabulaire s'est révélé constitué de formes ayant valeur de marqueurs chronologiques (*jeudi, mercredi, dimanche, soirée, journée*) et d'items désignant divers événements culturels (*théâtre, fête, bal, représentation*), ainsi que des réseaux de sociabilité (*cercles, réunion, concours, bienfaisance, banquet, conscrits, café, foire*). Après analyses, il ressort que cette zone isotropique est fortement liée aux rubriques et sous-rubriques des pages régionales, qui seules, voient leur spécificités lexicales se concentrer dans le quart-est nord des hémisphères gauche du planisphère. Cette concentration est tout particulièrement dense si l'on examine les spécificités de la sous-rubrique *Chronique régionale Doubs Besançon* comme l'atteste la vue zoomée (figure IV-55) issue de la projection géodésique suivante (figure IV-54).

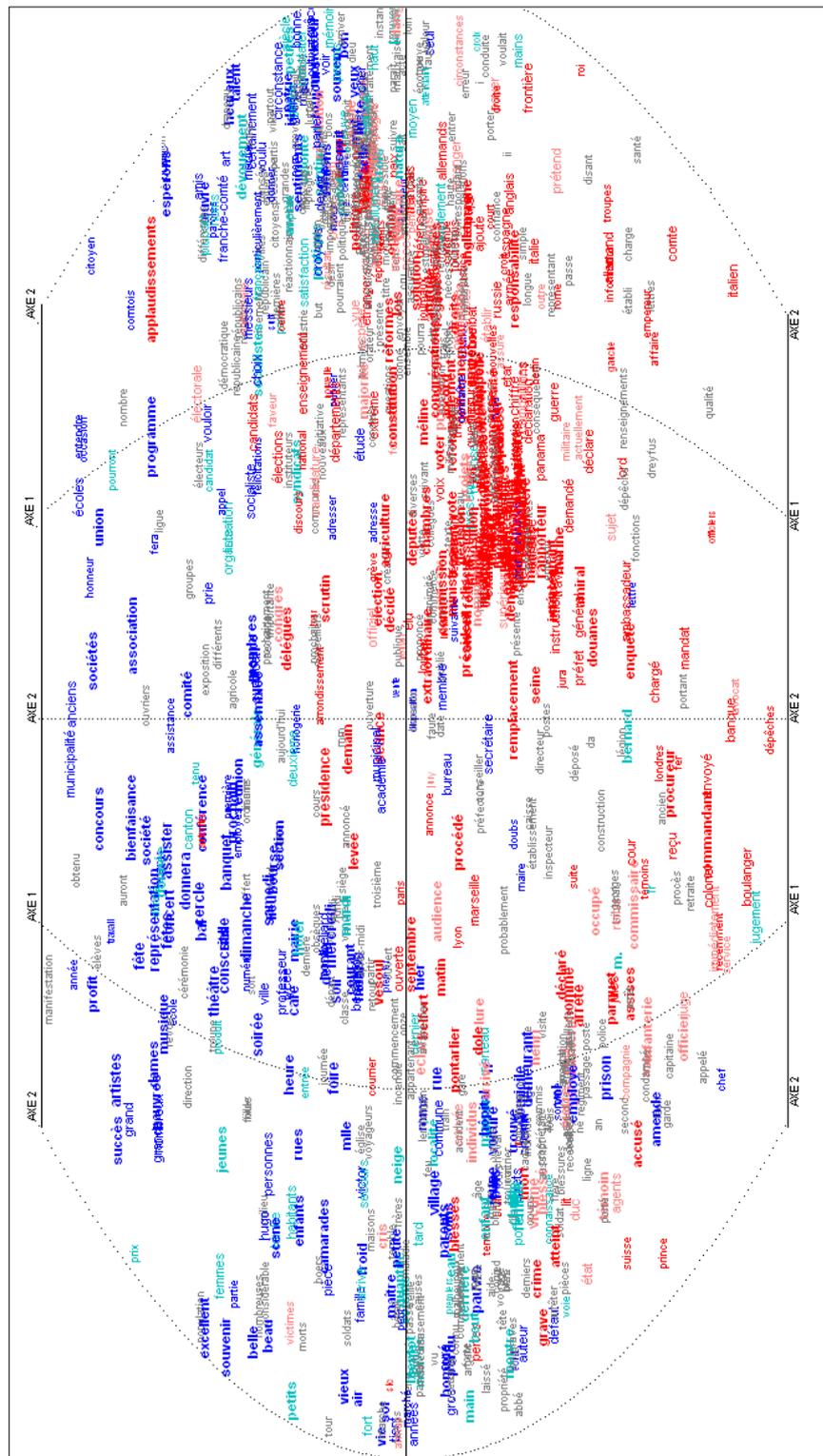


Figure IV-54 : Analyse micro-lexicale des 1000 premiers items du corpus *Petit Comtois* colorée d'après les suremplois dans « *Chronique régionale, Doubs, Besançon* ».

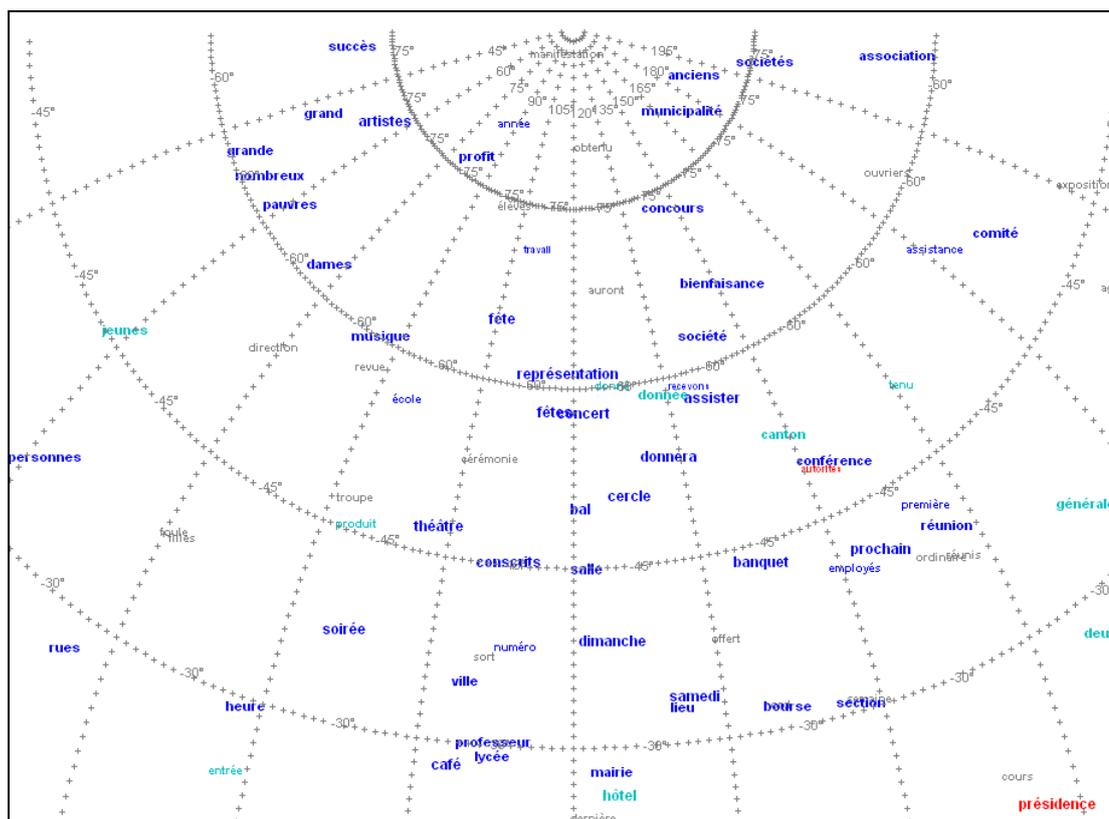


Figure IV-55 : Zoom sur la région B de la figure IV-53.

On peut y lire parmi les items les plus sur-représentés de la case des marqueurs temporels (*Samedi, heure, année, dimanche, soirée, prochain*), des toponymes (*ville, rues, café, mairie*), mais aussi des substantifs du champ de la vie culturelle et sociale (*musique, fête, concert, bal, cercle, banquet, bourse, section, réunion, comité, association, société, concours, conférence, pauvres, dames, artistes*).

Enfin, nous avons remarqué que le pôle C de la configuration fine du vocabulaire, situé dans le quart-ouest sud de la moitié droite du planisphère, regroupait un ensemble de formes lexicales du vocabulaire de l'activité législative. Parmi les formes qui structuraient cette zone isotropique, se trouvaient ainsi les items :

- *Demain, séance, préfecture, déposé, remplacement, seine, extraordinaire, président, présenté, décidé, élection, unanimité,*

commission, lecture, ministres, fédéral, nomination, chambre, examiner, interpellation, adopté, proposition, budget, décret, relatif, finances, rapporteur, marine, ministère, propositions, texte, vote, collègues, députés, agriculture, voter, discussion, chambre, projets, loi.

Si l'on compare ce pôle avec les items sur-représentés des *Dépêches de nuit*, on constate un parfait recoupement, comme on peut l'observer sur la figure IV-56, où sont colorés en bleu les items fortement sur-représentés dans cette case. Ces derniers correspondent à un ensemble de formes verbales conjuguées ou à l'infinitif (*déclare, examiner, rappelle, adopté, propose, voter, assure, termine*), à des substantifs, portant à la fois trace de l'activité législative et de ses objets (*examen, vote, nomination, voix, interpellation, discussion, proposition, question, budget, finances, loi, projets, réforme, dépenses, tribune, chiffre, état, déclaration, compagnies, millions, impôt, principe, accord, congrégations, mesures, texte, chambres, suppression, déclarations, majorité, guerre, marine*), ainsi qu'à des titres de fonction (*général, amiral, instruction, intérieur, travaux, marine, rapporteur, ministre, décret, député, députés, parlement, gouvernement, président, commission, ministres, collègues, sénat, orateur*).

Plus que sur ce recoupement, nous voudrions insister sur la configuration générale des spécificités lexicales des *Dépêches de Nuit*. La configuration diamétralement opposée des *Faits Divers* (figure IV-53) et des *Dépêches de Nuit* (figure IV-56) nous semble en effet un indice fort révélateur du lien étroit existant entre la consistance lexico-thématique des rubriques et la structure co-distributionnelle du vocabulaire du corpus : les sur-emplois propres aux *Faits Divers* se concentrent là même où les sous-emplois de *Dépêches de Nuit* s'inscrivent, et *vice versa*.

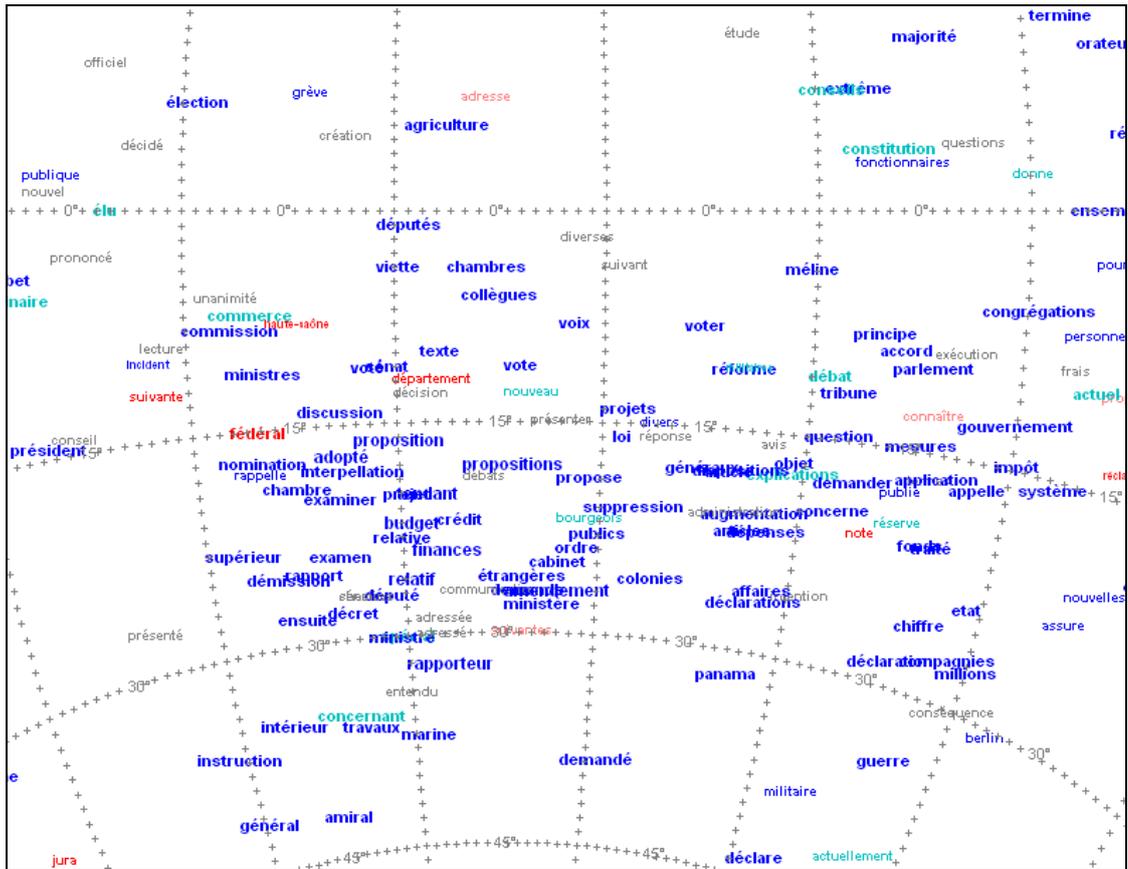


Figure IV-57 : Zoom sur les items sur-représentés de « *Dépêches de nuit* ».

Si les pôles isotropiques présentés précédemment étaient fortement structurés par une seule rubrique, le pôle isotropique D, d'après nos analyses, ne peut être éclairé à l'aune d'une seule case rédactionnelle, ce que nous attribuons à sa structure plus rhétorico-stylistique que lexicothématique et à la nature du quotidien de presse régionale du XIX^e siècle, où le discours politique ne saurait se déployer dans un seul site rédactionnel.

Il ressort de nos expérimentations que les spécificités lexicales d'une pluralité de rubriques, hébergeant des articles de commentaire, recourent, plus ou moins partiellement, ce pôle isotropique. À titre d'exemple, considérons les spécificités lexicales de la rubrique du *Bulletin*, correspondant à un article par lequel la rédaction met en lumière les événements politiques de la semaine et les commente.

Remarquons tout d'abord que les items sous-représentés de cette rubrique se répartissent de façon très nette dans la moitié gauche du planisphère. La plupart de ses items fortement (*opinion, sens, crise, adversaires, europe*) et moyennement sur-représentés (*cléricaux, pensée, action, politiques, république*) se concentrent quant à eux dans le quart nord-est de l'hémisphère gauche du plan.

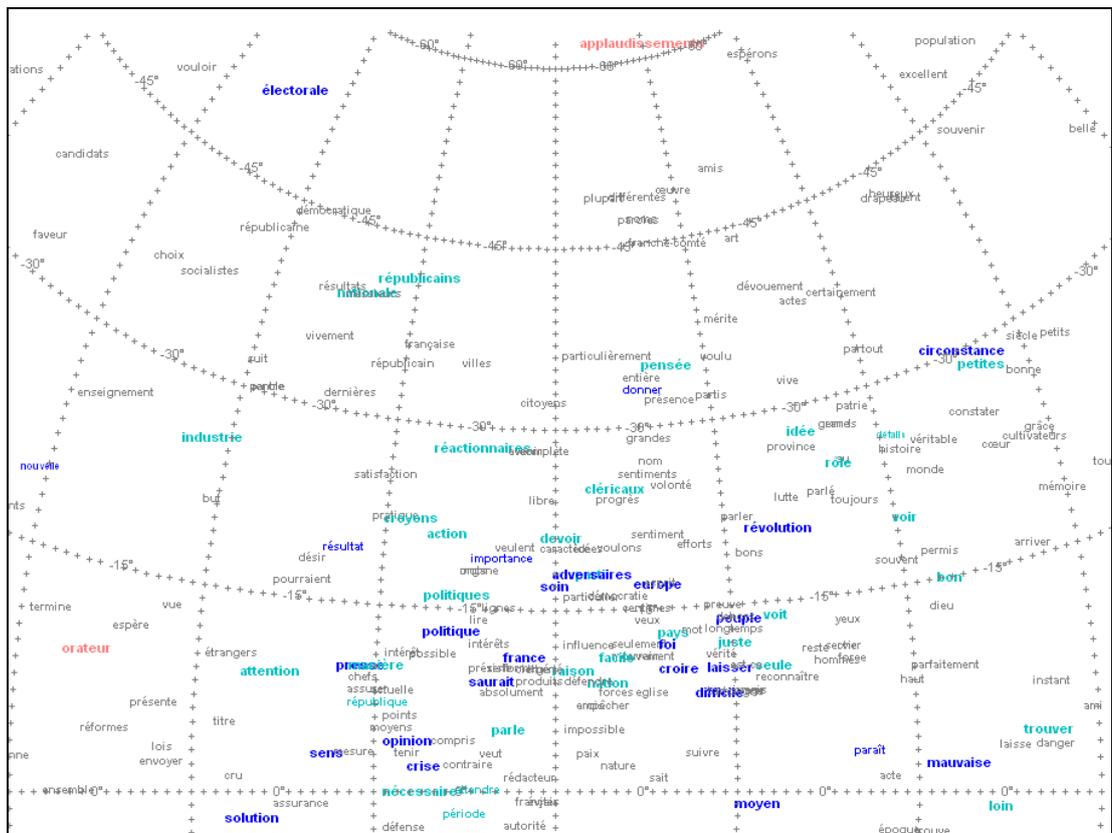


Figure IV-59 : Zoom sur la région D de la figure IV-57.

De même, les items moyennement et fortement sur-représentés de la case *Chronique de la Semaine*, regroupés dans l'hémisphère gauche du planisphère, recoupent également partiellement le pôle isotropique D, comme on peut l'observer sur les figures IV-60 et IV-61.

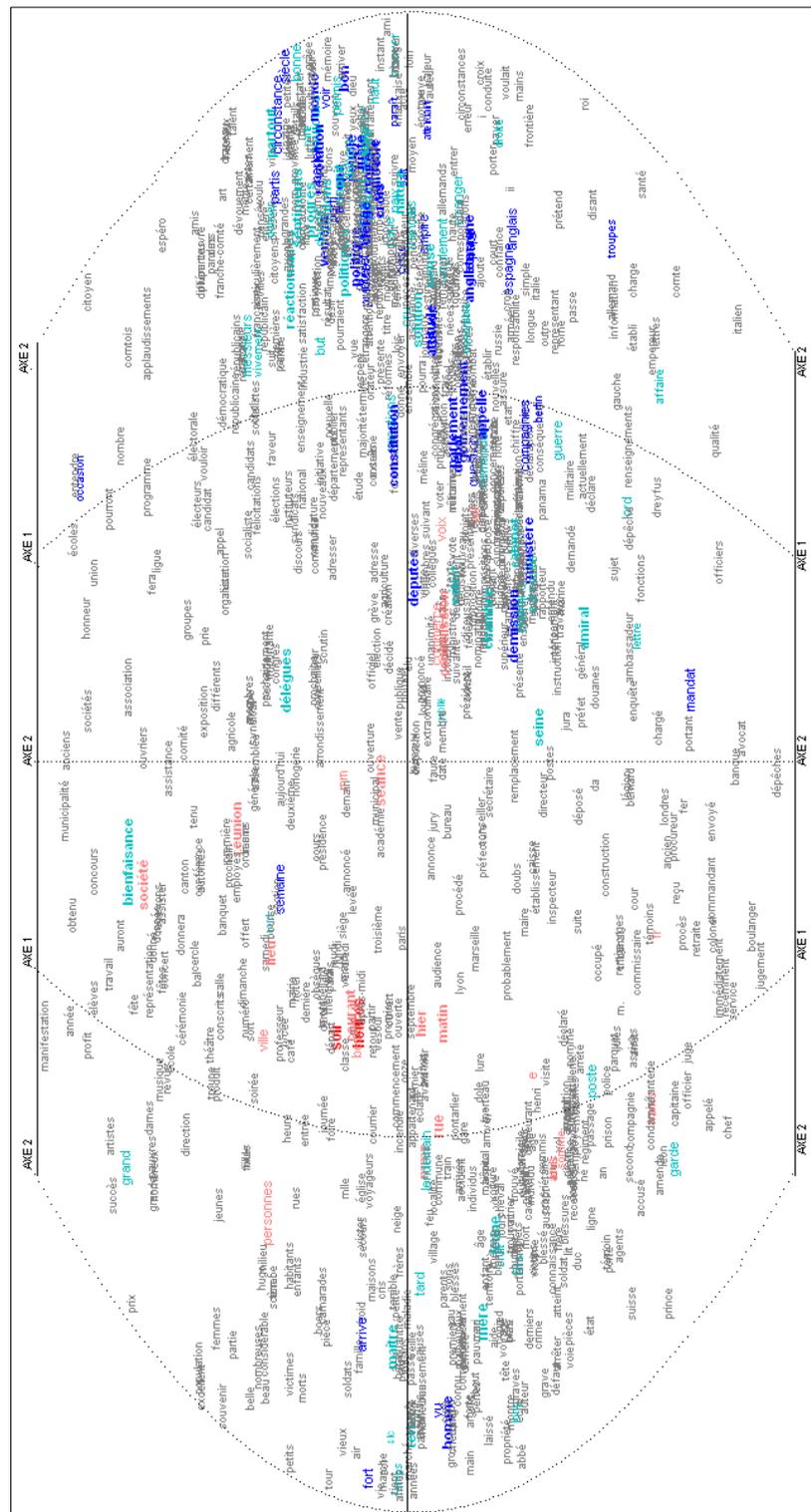


Figure IV-60 : Analyse micro-lexicale des 1000 premiers items du corpus *Petit Comtois* colorée d'après les suremplois dans « *chronique de la semaine* ».

vues sur les différentes dimensions du texte, auxquelles nous ont donné accès les différentes phases d'analyse mises en œuvre. C'est le potentiel heuristique de l'articulation des différents niveaux d'appréhension du vocabulaire que nous voudrions mettre en exergue.

Le premier volet d'expérimentations, axé sur le niveau macro-distributionnel, a fourni une vue synthétique de la distribution du vocabulaire dans les 21 sous-corpus diachroniques du corpus, pointant une sensible dichotomie diachronique (à l'exception notable de l'année 1883). Une division relativement nette se donne en effet à lire entre la première tranche diachronique du corpus et la seconde, à savoir 1884-1896 et 1897-1903, tout en marquant le profil saillant de l'année 1883. Cette dichotomie, faisant écho aux lignes dégagées lors de notre historique de la Troisième République, peut être envisagée comme l'indice de l'évolution des débats et des opinions se déployant dans cette conjoncture socio-historique. La proximité des années de la série 1884-1896 révélée par l'analyse de la distribution du vocabulaire indique, d'une part, le climat de tension entre la France et l'Allemagne, attisé par une série d'incidents diplomatiques tels que l'affaire Schnaebelé et le contexte colonialiste. Cette série s'unifie, d'autre part, à la lumière d'une poussée anticléricale sensible dès 1892, qui doit être envisagée comme la contestation de la *politique d'apaisement* menée par J. Méline, et de la succession d'affaires économique-politiques. Le regroupement de la série 1897-1903, d'abord indice d'une période d'instabilité ministérielle notamment alimentée par l'affaire Dreyfus, traduit les mutations des forces politiques en présence et l'intensification des débats autour de la question sociale, et plus généralement, le tournant radical de la Troisième République.

Notre second volet d'analyse s'est focalisé sur une approche de la co-occurrence généralisée dans le corpus : elle a autorisé de percevoir que l'organisation réticulaire du vocabulaire était structurée autour de quatre pôles isotropiques très distincts, indices des dominantes lexico-thématiques et rhétorico-stylistiques du *Petit Comtois*. C'est ainsi l'importance accordée aux

faits divers et à la vie législative qui se donne à lire, tandis que transparait la fonction d'agenda du quotidien régional, informant du programme des manifestations politiques, culturelles et mondaines. Enfin, le quatrième pôle isotropique, évoquant clairement le discours politique, s'articule autour du vocabulaire de l'idée républicaine.

En vue d'observer si cette configuration fine du vocabulaire connaissait des évolutions diachroniques analogues à celles que suggéraient les résultats de l'analyse macro-distributionnelle, nous avons ensuite entrepris une série d'examens visant à prendre la mesure de la variation des réseaux isotropiques du corpus au fil du temps. Contre nos attentes, l'éclairage de la configuration fine du vocabulaire par la diachronie a pointé une stabilité remarquable des pôles isotropiques, ne donnant guère prise notamment à la dichotomie de macro-distribution relevée *supra*. Ce résultat s'avère stimulant à plus d'un titre : d'une part, il indique que la dichotomie diachronique observée au niveau macro-distributionnel n'a aucune influence sur les réseaux lexico-thématiques du quotidien. Ce résultat est d'autant plus surprenant, d'autre part, en raison de la nature de notre corpus de presse quotidienne : celui-ci est en effet supposé rendre compte d'une actualité singulière, qui engendrerait une évolution des thèmes et des thématisations.

Parce que les pôles isotropiques évoquent très fortement le rubriquage du quotidien, nous nous sommes tournée vers une analyse des dominantes lexico-thématiques des rubriques. Des différentes classes d'analyse menées, il ressort une indéniable corrélation du rubriquage et de la structuration fine du vocabulaire. Les pôles à consistance lexico-thématiques se révèlent ainsi presque parfaitement recoupés par l'examen des spécificités lexicales des rubriques *Faits-Divers* et *Dépêches de nuit*, notamment.

Ces résultats nous amènent à réfléchir sur la notion même d'actualité et à observer derrière son apparence fluctuante une essence « récurrente ». Celle-ci nous semble à relier à l'inscription du discours du *Petit Comtois* dans une formation discursive républicaine qui détermine son contenu et ses formes

en retour. C'est tout le sens qu'il nous semble devoir accorder à la réflexion de Marc Angenot, qui souligne l'étroit rapport de l'actualité à l'idéologie en ces termes :

L'Actualité est ce qui produit la rencontre inattendue sur une table de dissection doxique du réel imprévisible et de la nécessité idéologique (Angenot, 1989 : 597-598).

L'actualité, en tant qu'événement discursif, ne peut être pensée indépendamment des déterminations de la formation discursive : celle-ci détermine ce qui sera événement et les ressorts de mise en scène de ce dernier, tant au niveau des structures narratives et des effets rhétorico-stylistiques que du lexique.

Chapitre V.

Approches du niveau morphosyntaxique sur le corpus *Petit Comtois*. Essai de typologisation textuelle.

En nous inspirant des nombreuses recherches encouragées par les travaux pionniers de D. Biber (1988), caractérisant des données textuelles à la lumière de leurs caractéristiques morphosyntaxiques dans une perspective comparative et/ou inductive (Rastier & Malrieu, 2001 ; Beauvisage, 2001 ; Poudat, 2006 ; Mayaffre, 2000, 2007a, 2007b), nous consacrerons le présent chapitre à une analyse morphosyntaxique des formes d'écriture du *Petit Comtois* regroupées selon leur rubrique d'appartenance. Il s'agira pour nous d'évaluer les apports d'une analyse prenant pour entrée le niveau morphosyntaxique à une exploration textuelle du discours du *Petit Comtois*.

Cette entreprise doit être entendue comme le préalable à la mise en œuvre future d'une observation diachronique des genres journalistiques dans le *Petit Comtois* et dans la PQR3R. Sur ce point, nous nous empresserons de préciser que si la rubrique reflète la division du travail et l'hétérogénéité des pratiques socio-discursives du journal, elle ne peut néanmoins être considérée comme un genre rédactionnel : tout au plus peut-elle constituer un indice de la genericité du texte si elle se réfère à lui dans son intitulé. Il ne s'agit donc pas

pour nous d'accéder au genre par la rubrique, mais d'envisager une méthodologie d'analyse, qui nous permette d'établir une typologie des rubriques du *Petit Comtois*. Cette expérimentation constitue le premier seuil d'une mise en place d'un dispositif d'observation des déterminations *locales* et *globales* (Rastier, F. & Malrieu, D., 2001 : 552) imposées par le genre.

Dans notre perspective, une étude visant à l'observation des genres exige, en premier lieu, que l'étiquetage morphosyntaxique convoqué dans ce cadre soit aussi fiable que possible, et, en second lieu, que le protocole méthodologique d'analyse soit impérativement robuste et transparent. Les expérimentations que nous allons présenter ici se fixent donc deux objectifs.

Le premier est d'évaluer la qualité d'un étiquetage morphosyntaxique opéré automatiquement, notamment convoqué par la *logométrie* (Mayaffre, 2000) sur un panel de données textuelles volumineux : il s'agira plus précisément dans le temps de cette étude d'identifier les erreurs d'un étiquetage automatique, pour, dans une contre-expérience future basée sur un étiquetage fin et contrôlé des mêmes données, mettre au jour l'impact d'un taux d'erreur traditionnellement jugé sans conséquence sur les résultats de l'analyse. Aussi nous consacrerons le premier temps de ce chapitre à exposer les résultats d'un étiquetage automatique, après avoir procédé à un bref état de l'art des entreprises de typologie textuelle à partir de variables morphosyntaxiques.

Notre second objectif est d'examiner les pistes méthodologiques de constitution d'un protocole d'analyse transparent et reproductible, suffisamment robuste pour supporter le traitement de la totalité des articles de notre corpus considérés comme autant d'individus au sens statistique, tel que l'exige une analyse parfaitement inductive des formes génériques du *Petit Comtois*¹. Dans

¹ Attribuer un genre d'appartenance à un article constitue une hypothèse initiale qui peut se justifier, et ce, tout spécialement dans le cadre d'une démarche exploratoire, où l'on procède par mouvements de réajustements. Certains types de discours semblent néanmoins se prêter plus ou moins bien à cette démarche. Dans le cadre de notre analyse de la presse écrite ancienne et régionale, nous sommes confrontée à un ensemble très hétérogène de textes ne s'accompagnant que de peu d'indices permettant de légitimer notre catégorisation. Nous jouons de la difficulté supplémentaire d'être coupée du contexte de production de ce discours et de ne pouvoir nous appuyer sur une méta-littérature qui lui soit contemporaine ; dès lors, catégoriser préalablement les articles du *Petit Comtois* nous paraît un exercice bien périlleux, nous

un deuxième temps, nous comparerons donc les résultats livrés par un examen basé sur un nombre restreint de variables ayant trait aux catégories grammaticales et sur un nombre plus conséquent de variables linguistiques.

5.1. Bref cadrage de la pratique de typologisation textuelle à partir de variables morphosyntaxiques

Initiées par D. Biber en 1988, les pratiques de typologisation textuelle à partir de corpus hétérogènes s'imposent aujourd'hui comme un chantier essentiel pour les linguistiques de corpus, l'univers de l'ingénierie documentaire, l'enseignement des langues, etc. Alors que les données numériques textuelles sont de plus en plus accessibles, il devient en effet crucial de pouvoir « maîtriser ce déluge de données hétérogènes » (cf. Illouz et al., 1999).

5.1.1. Méthodes inductives et analyses statistiques des données

On doit à D. Biber (1988) d'avoir ouvert la voie aux études de typologisations textuelles inductives dans son ouvrage *Variation between speech and writing*. Cherchant à vérifier la pertinence sur le plan linguistique de l'opposition établie *a priori* à partir de critères situationnels entre l'oral et l'écrit, D. Biber fonde son approche sur une analyse quantitative de traits linguistiques. Cette perspective repose sur le postulat que la *cooccurrence* de traits linguistiques définit une dimension linguistique, interprétée comme l'indice d'une dimension fonctionnelle :

Each dimension comprises an independent group of co-occurring linguistic features, and each co-occurrence pattern can be interpreted in functional terms (Biber, 1988: 14).

exposant aux risques de projection des catégories discursives du XXI^e siècle sur un discours antérieur, et à une démarche circulaire.

Sur le plan méthodologique, l'entreprise de D. Biber repose sur l'examen de 67 traits linguistiques, relevant majoritairement du niveau morphosyntaxique, dans un corpus composé de 481 textes normalisés en séquence de 1000 mots, issus de deux grands corpus de langue anglaise contemporaine (*The Lancaster_Oslo_Bergen Corpus of British English* ; *The London_Lund Corpus of Spoken English*), comptabilisant 23 genres différents (articles de recherche, reportages, conservations, nouvelles radiophoniques). Ces données, une fois étiquetées, sont ensuite soumises aux techniques de statistiques multivariées, et plus particulièrement à l'Analyse en Composantes Principales (ACP). Cette méthode mathématique d'analyse de tableaux à double entrée, parente de l'AFC, permet d'analyser et de visualiser les corrélations entre n variables sur un graphique à deux ou trois dimensions construit de manière à ce que la dispersion entre les données soit aussi bien préservée que possible. Cette procédure extrait la valeur maximum de la variance partagée parmi les variables pour chaque facteur. Ainsi, le premier facteur extrait la valeur maximale de la variance partagée, c'est-à-dire le groupe le plus important de cooccurrence des données. Le deuxième facteur extrait alors la valeur maximum de la variance partagée des variables après que le premier facteur ait été défalqué, et ainsi de suite. L'ACP permet ainsi de construire un ensemble de facteurs qui présentent la particularité d'être non corrélés entre eux, ce qui permet à D. Biber d'obtenir des dimensions les plus distinctes possibles. Les corrélations observées sur le graphique par des nuages de traits en cooccurrence, obtenues au terme de l'analyse, sont envisagées par D. Biber comme l'indice de dimensions fonctionnelles.

À partir de l'interprétation des facteurs constitués par l'ACP et de l'examen des contributions des variables, c'est-à-dire des variables présentant un fort poids dans la constitution de chaque facteur, D. Biber construit alors six dimensions textuelles, constituées comme suit :

- Dimension 1 : Involved versus Informational Production
- Dimension 2 : Narrative versus Non-narrative concerns
- Dimension 3 : Explicit versus Situation-Dependent Reference

- Dimension 4 : Overt expression of Persuasion
- Dimension 5: Abstract versus Non-abstract information
- Dimension 6 : On-Line Informational Elaboration²

Afin de vérifier la validité de ces dimensions, D. Biber calcule le poids de chaque trait linguistique en termes de rang de variation plutôt qu'en termes de fréquence absolue. Cette valeur standardisée, nommée « valeur du facteur » (*factor scores*) est obtenue par l'addition, pour chaque texte, des occurrences de chaque variable fortement contributive du facteur étudié.

La méthodologie de D. Biber s'avère innovante en cela qu'elle recourt aux méthodes d'analyses statistiques multivariées, auparavant réservées à l'analyse de données lexicales, pour mettre au jour des faisceaux de traits linguistiques corrélés et par suite des types de texte. Cette démarche a inspiré par la suite de nombreuses entreprises de profilage de textes reposant sur une analyse quantitative de critères morphosyntaxiques (Karlsgren, J. & Cutting, D., 1994 ; Kessler, B., Nunberg et al., 1997 ; G. Illouz, P. Lafon, *et al.* (2000) Rastier & Malrieu, 2001, Beauvisage, 2001). Il convient cependant de revenir sur l'aspect « inductif » même revendiqué par ces démarches. Si l'analyse statistique multidimensionnelle permet bien de faire émerger les variables discriminantes des textes eux-mêmes, il n'en demeure pas moins que les données soumises à cette procédure sont des « construits » : toute variable morphosyntaxique décèle une tradition grammaticale et morphosyntaxique sous-jacente. Si l'interprétation des facteurs et des dimensions textuelles est objectivée, aucune variable de l'analyse ne peut être posée comme purement objective, ce que rappelle très justement S. Branca-Rosoff (1999b) :

Le choix des paramètres est cependant posé comme naturel ; c'est-à-dire qu'il n'est pas explicité par l'analyste. Or, les traits retenus n'ont rien « d'objectif ». Leur sélection résulte des hypothèses du chercheur et de la tradition, plus précisément des bases typologiques qu'il juge intéressantes. De ce point de vue, il n'y a pas d'analyse purement immanente et on perçoit par exemple chez D. Biber l'influence des

² Notre traduction : D1 : Production induite/impliquée vs production informationnelle ; D2 : Préoccupations d'ordre narratif vs non-narratif ; D3 : Référence dépendant de la situation/inférable par contexte VS explicite ; D4 : Expression ouverte de la Persuasion ; D5 : Information abstraite vs concrète/non-abstraite ; D6 : Élaboration de l'Information en direct.

travaux sur la syntaxe de l'écrit ou des analyses énonciatives, sans que jamais ces sources soient mentionnées et discutées (Branca-Rosoff, 1999b : 15).

5.1.2. Genres et typologies textuelles

Les travaux pionniers de D. Biber n'accordent qu'une importance mineure aux genres, qui ne sont d'ailleurs envisagés qu'à titre de catégories sociales intuitivement pressenties par les locuteurs. *A contrario*, les entreprises de typologisation textuelle menées hors du courant anglo-saxon, et tout spécialement en France, ont pour pierre angulaire la notion de genre, à de rares exceptions près, telle que l'entreprise de profilage de textes conçue et réalisée par B. Habert, G. Illouz, P. Lafon, *et al.* (2000) qui s'attachent à « neutraliser » la généricité du corpus d'interventions radio-télévisées analysé³.

Postulant, dans le cadre d'une sémantique différentielle, que le genre constitue le niveau fondamental pour la classification des textes, F. Rastier et D. Malrieu (2001) interrogent la capacité d'une analyse morphosyntaxique à recouper automatiquement les catégories de genres, de champs génériques et de discours. Il s'agit pour les auteurs de vérifier, grâce à l'analyse multidimensionnelle, si un système d'oppositions de traits linguistiques légitime les catégories habituellement admises, et non pas de définir les genres. Entendant le genre comme une *interaction non-séquentielle et normée* de composantes autonomes d'ordre thématique, dialectique, dialogique et tactique, F. Rastier et D. Malrieu dévient à une analyse repliée sur le niveau de la phrase toute capacité à définir le genre :

Si l'approche empirique « inductive » peut aider à mettre en évidence des différences entre textes, nous n'attendons pas qu'elle *définisse* les genres textuels. Nous considérons en effet que les genres sont définis par l'interaction normée de composantes textuelles : nous jouons alors de difficulté, puisque les traits linguistiques dont nous disposons pour conduire l'étude présentée ici sont issus d'une analyse morphosyntaxique au palier de la phrase qui ne tient évidemment pas

³ Il conviendrait de remarquer également qu'une étude telle que Mayaffre (2004), réunissant des textes de nature très hétérogène, ne s'attache pas tant à observer des *genres* que les spécificités grammaticales des énoncés des locuteurs.

compte des contraintes des paliers supérieurs (Rastier & Malrieu, 2001 : 553).

Leur analyse se fonde sur un gigantesque corpus composé de 2541 ouvrages de textes intégraux, issus de quatre discours (juridique, scientifique, essayiste, littéraire) soit 164 millions de mots. D. Malrieu et F. Rastier procèdent à un étiquetage automatique de ce corpus en recourant au logiciel *Cordial*, qui leur fournit 251 traits (en grande partie morphosyntaxiques), pouvant être convoqués comme variables :

Les principales classes de catégories dont nous disposons sont les suivantes : les ponctèmes (que complète le décompte des dialogues, des paragraphes, des phrases, des incises, des propositions) ; les parties du discours et les pourcentages des sous-catégories pour chacune d'elles ; les personnes des verbes, des adjectifs possessifs et pronoms possessifs et personnels ; les temps verbaux ; les types de verbes, transitifs directs ou indirects, avec COD obligatoire, types de sujet ou de COD (abstrait/concret, animal, animé) etc. ; les types de noms : nom propre (humain, prénom, géographique, autre), nom commun (abstrait ou concret, animal, animé, humain, humanoïde), % de noms de lieu, de temps, de profession, noms composés, noms épithètes, de noms appartenant à un groupe nominal ; les types de propositions (principales, coordonnées, subordonnées et les types de ces dernières) ; enfin les types de compléments (Rastier & Malrieu, 2001 : 557).

Le score en pourcentage de chaque trait morphosyntaxique pour chaque ouvrage est alors répertorié puis soumis successivement à trois méthodes d'analyse, à savoir l'analyse univariée, l'analyse en composantes principales et la classification hiérarchique ascendante. Ces trois méthodes convoquées par F. Rastier et D. Malrieu s'avèrent présenter de très encourageants résultats. Elles convergent en effet pour faire émerger des différences notables entre discours, champs génériques et genres, révélant la forte variation des traits morphosyntaxiques.

5.2. Méthodologie générale

5.2.1. Recueil et sélection des variables

Pour étiqueter notre corpus d'articles de presse, nous avons recouru au logiciel *Cordial* conçu par la société *Synapse Développement*, qui s'impose, avec *TreeTagger*⁴, comme l'un des deux outils d'étiquetage automatique de référence dans le champ de l'analyse textuelle des données. C'est notamment sur les codes grammaticaux fournis par ce logiciel que s'appuient les études menées par F. Rastier et D. Malrieu (2001), T. Beauvisage (2001), D. Mayaffre (2004, 2007b), M. Kastberg (2003) et qui sont désormais conçus pour être implémentés dans *Hyperbase*.

Le logiciel *Cordial* livre le résultat de son annotation des données dans deux types de fichiers. Si l'utilisateur vise l'annotation d'un ensemble limité de textes, il peut faire le choix de les ouvrir un par un dans l'interface de *Cordial*, puis ordonner au logiciel d'annoter les données de chaque texte et de les sauvegarder dans un fichier propriétaire de type « sta ». Dans ce cas, *Cordial* crée un fichier relativement lourd, faisant synthétiquement état, sous forme de listes, des caractéristiques morphosyntaxiques du texte (en nombre absolu et pourcentage), de ses spécificités sémantiques mais également de ces traits stylistiques ainsi que de son niveau de langue. Dans les quelque 1390 lignes de résultats ainsi construites, peu se révèlent pertinentes pour l'utilisateur ne s'intéressant qu'à une analyse du niveau morphosyntaxique. Plus de 1025 lignes de résultats sont en effet relatives aux domaines ou thèmes (principal ou secondaire) des textes étudiés. Il convient de remarquer à ce propos que ces résultats sont fort peu probants et qu'ils dénaturent plus les thématiques observées qu'ils ne les décrivent. Il ne s'agit bien sûr pas ici d'incriminer le logiciel *Cordial*, mais d'évoquer les contre-performances de l'état actuel des techniques d'étiquetage sémantique automatique et son cadre conceptuel, consistant à accoler à chaque mot d'un texte une étiquette correspondant à son « sens ». La pertinence et la fiabilité des renseignements

⁴ <http://www.ims.uni-stuttgart.de/projekte/corplex/TreeTagger/>

relatifs au style et au niveau de langue du texte sont encore plus critiquables, totalement inadaptées qu'elles sont à l'analyse de textes antérieurs au milieu du XX^e siècle. Toute comparaison du texte à des données exogènes non identiques quant à leur diachronie ou leur genre est en effet une pratique très contestable, dont on peut légitimement remettre en cause la significativité. Connaître le pourcentage de mots du texte considéré appartenant au *Dictionnaire fondamental* de Georges Gougenheim, conçu en 1958, listant les 3500 mots les plus courants de la langue française supposés correspondre au vocabulaire de base de la langue française, n'est nullement pertinent dans le cadre d'une étude de la presse du XIX^e siècle. De même, une étude du lexique selon L'*Échelle Dubois-Buyse* de François Ters, Georges Mayer et Daniel Reichenbach, construite à partir des relevés de fréquence d'occurrences dans un corpus d'œuvres littéraires, d'articles de presse du journal *Le Monde*, de dépêches de l'AFP, de textes issus de sites et de forums de discussion sur Internet évaluant la maîtrise de l'orthographe du rédacteur du texte considéré, est une donnée totalement inutile dans notre perspective.

Si l'on se focalise sur les lignes de résultats relatifs au niveau morphosyntaxique sans même s'intéresser à la qualité de l'étiquetage effectué, les problèmes persistent. La conception grammaticale sous-jacente à certaines variables proposées dans ce type de fichier se révèle pour le moins contestable. Parmi les résultats de l'annotation morphosyntaxique, on rencontre ainsi une ligne vouée à informer l'utilisateur du « pourcentage de pronoms personnels à la première personne par rapport aux pronoms possessifs » observés par *Cordial* dans le texte analysé. Les pronoms personnels forment, certes, une classe hétérogène d'éléments aux fonctionnements sémantiques dichotomiques, à la morphologie complexe. Néanmoins, les pronoms personnels ne peuvent en aucun cas être comparés sur la base de pronoms possessifs, dont la morphologie et l'expression sémantique sont incontestablement différentes. Selon nous, une partie considérable des variables morphosyntaxiques proposées dans ce type de fichier présente une

pertinence extrêmement faible. Nous restons en effet sceptique quant au caractère discriminant de variables à modalité du possible telles que le *pourcentage d'adverbes utilisables avec un substantif animé*, le *pourcentage de verbes parfois impersonnels parmi l'ensemble des verbes*, le *pourcentage de verbes à COD pouvant être humain ou humanoïde*, etc.

Enfin, de façon encore plus problématique, le fichier de type « sta » constitué par *Cordial* est une sorte de « boîte noire » dans lequel sont livrés des résultats d'annotation sans qu'il soit à aucun moment possible de retourner à cette annotation et de la contrôler. En raison de l'opacité totale du processus d'annotation, présentant en outre l'inconvénient de ne donner accès qu'à des sorties statistiques inadaptées à la méthodologie de l'analyse factorielle des correspondances, nous avons choisi de ne pas travailler sur les résultats de type « sta » de *Cordial*.

Nous avons donc employé le deuxième type de fichier d'étiquetage automatique proposé par *Cordial*, qui permet d'une part, de personnaliser les paramètres d'annotation de *Cordial*, ou plus précisément de ne sélectionner que les données qui nous semblent pertinentes et fiables *a priori*, et d'autre part, de visualiser et de vérifier l'annotation assignée à chaque élément du texte.

5.2.2. Remarques sur l'étiquetage effectué par *Cordial*

Considérons à présent la qualité de l'étiquetage opéré de façon automatique par *Cordial*. Les instruments d'annotation automatique, comme le rappelle B. Habert (2005), sont des *instruments imparfaits*. Le logiciel *Cordial* ne déroge pas à ce constat. Si l'on évalue le taux de réussite entre 95 et 97% de mots correctement annotés pour le français écrit standard selon ce même auteur (Habert, 2005 : 40), un certain nombre d'erreurs d'étiquetage récurrentes, notamment liées à la segmentation opérée, s'avèrent presque rédhibitoires. D'autant plus que le taux de réussite de l'annotation varie en fonction de la nature des données analysées.

Les études sur corpus littéraires annotés par *Cordial* ont notamment été l’occasion de montrer les défauts d’identification des temps verbaux : comme l’observent D. Malrieu et F. Rastier (2001 : 556), le plus-que-parfait de l’indicatif est particulièrement mal reconnu par *Cordial*, de même que le passé simple et le subjonctif présent sont sous-estimés au profit du présent de l’indicatif lorsque les formes des verbes ne comportent pas de variation selon la conjugaison. Dans notre corpus de presse, les cas de plus-que-parfait sont extrêmement rares ; de même, ce type de discours recourt dans une moindre mesure que le discours littéraire au passé simple ; néanmoins, il y a de nombreux cas de subjonctifs dans notre corpus. Leur annotation est systématiquement erronée dès lors que les formes verbales du présent de l’indicatif coïncident avec celle du subjonctif.

Il faut que le public donne à l'État le temps de faire les améliorations nécessaires.			
N° mot	Dialogue	lemme	Typegram
1	Il	il	PPER3S
2	faut	falloir	VINDP3S
3	que	que	SUB
4	le	le	DETDMS
5	public	public	NCMS
6	donne	donner	VINDP3S
7	à	à	PREP
8	l'	le	DETDMS
9	Etat	état	NCMS
10	le	le	DETDMS
11	temps	temps	NCMIN
12	de	de	PREP
13	faire	faire	VINF
14	les	le	DETDPIG
15	améliorati	améliorati	NCFP
16	nécessaires	nécessaire	ADJPIG
17	.	.	PCTFORTE

Figure V-1 : Erreur d’étiquetage du présent du subjonctif présent.

La plupart des erreurs que nous avons répertoriées sont imputables à une identification erronée de la catégorie morphosyntaxique de termes homographes polycatégoriels, c'est-à-dire des formes susceptibles de recouvrir plusieurs catégories morphosyntaxiques différentes : par exemple, la forme *ferme* peut être un substantif, un verbe ou un adjectif. Confronté à une forme virtuellement ambiguë (Fuchs, 1996 : 35), *Cordial* fournit des résultats souvent très peu satisfaisants. Force est de constater que l'analyse du contexte syntaxique mise en œuvre par le logiciel ne suffit pas à lever l'ambiguïté de la forme. Cas d'école des plus classiques, la forme graphique « fait », dont l'utilisation suivante correspond à un participe passé, est étiquetée en tant que nom commun :

N'est-ce pas déjà un grand pas de fait dans la voie que nous préconisons?			
N° mot	Dialogue	lemme	Typegram
1	N'	ne	ADV
2	est	être	VINDP3S
3	ce	ce	PDS
4	pas	pas	ADV
5	déjà	déjà	ADV
6	un	un	DETIMS
7	grand	grand	ADJMS
8	pas	pas	NCMIN
9	de	de	PREP
10	fait	fait	NCMS
11	dans	dans	PREP
12	la	le	DETDFS
13	voie	voie	NCFS

Figure V-2 : Erreur d'annotation de la forme *fait*.

Dans certains cas, une forme ambiguë fautive ment identifiée compromet la validité de l'analyse morphosyntaxique d'une phrase entière. Dans l'exemple suivant, l'adjectif *séculier* ayant été identifié à tort comme un nom commun, *Cordial* considère par suite le verbe *prodigue* comme un adjectif. La non-reprise

du sujet après la coordination « et » compromet pour sa part l'identification du verbe *rêve*, qui est annoté comme un nom commun.

Sous prétexte de nationalisme, le clergé régulier ou séculier prodigue au sabre les avancées et rêve d'un nouveau César.			
N° mot	Dialogue	lemme	Typegram
1	Sous	sous	PREP
2	prétexte	prétexte	NCMS
3	de	de	PREP
4	nationalisme	nationalisme	NCMS
5	,	,	PCTFAIB
6	le	le	DETDMS
7	clergé	clergé	NCMS
8	régulier	régulier	ADJMS
9	ou	ou	COO
10	séculier	séculier	NCMS
11	prodigue	prodigue	ADJSIG
12	au	au	DETDMS
13	sabre	sabre	NCMS
14	les	le	DETDPIG
15	avancées	avancée	NCFP
16	et	et	COO
17	rêve	rêve	NCMS
18	d'	de	PREP
19	un	un	DETIMS
20	nouveau	nouveau	ADJMS
21	César	César	NPMS
22	.	.	PCTFORTE

Figure V-3 : Exemple d'erreur d'identification provoquant la non-reconnaissance du cotexte immédiat.

Les exemples de telles « contaminations » sont très fréquents dans notre corpus. Entre autres exemples de suites ambiguës mal identifiées par *Cordial*, on notera le cas des titres de journaux dans lesquels l'adjectif *Petit* précède un substantif, systématiquement analysé par *Cordial* comme la succession d'un adjectif et d'un nom commun, comme en témoigne la figure V-4.

L'Administration et la Rédaction du Petit Comtois croient n'avoir rien négligé pour assurer au lecteur un service d'informations intéressantes, complètes et rapides.			
N° mot	Dialogue	lemme	Typegram
1	L'	le	DETDFS
2	Administration	administration	NCFS
3	et	et	COO
4	la	le	DETDFS
5	Rédaction	rédaction	NCFS
6	du	du	DETDMS
7	Petit	petit	NCMS
8	Comtois	comtois	ADJMIN
9	croient	croire	VINDP3P
10	n'	ne	ADV
11	avoir	avoir	VINF
12	rien	rien	PII
13	négligé	négliger	VPARPMS

Figure V-4 : Erreur d'identification des formes *petit* et *comtois*.

Dans ce qui semble constituer la même classe d'erreurs, on mentionnera encore les difficultés du logiciel à reconnaître les verbes du troisième groupe conjugués à la troisième personne du présent succédant à un nom commun, comme dans l'exemple suivant, où la forme conjuguée de *courir* est assimilée à un adjectif :

Le bruit court, non sans apparence de raison, que c'est l'abandon par le gouvernement de l'article			
N° mot	Dialogue	lemme	Typegram
1	Le	le	DETDMS
2	bruit	bruit	NCHMS
3	court	court	ADJMS
4	,	,	PCTFAIB
5	non	non	ADV
6	sans	sans	PREP
7	apparence	apparence	NCFS
8	de	de	PREP
9	raison	raison	NCFS
10	,	,	PCTFAIB
11	que	que	SUB
12	c'	ce	PDS
13	est	être	VINDP3S
14	l'	le	DETDMS
15	abandon	abandon	NCMS

Figure V-5 : Erreur d'identification de la forme conjuguée *courir*.

Autre exemple de contamination : la non-reconnaissance du pronom indéfini « d'autres » provoque l'identification de la forme conjuguée à la troisième personne du pluriel du verbe *écrémer* à l'indicatif présent ; comportant d'ailleurs une faute d'accentuation, comme un nom commun dont il n'existe, à notre connaissance, aucune attestation.

d'autres écrément beaucoup trop le lait ;			
N° mot	Dialogue	lemme	Typegram
1	d'autres	d'autres	ADJIND
2	écrément	écrément	NCI
3	beaucoup	beaucoup	ADV
4	trop	trop	ADV
5	le	le	DETDMS
6	lait	lait	NCMS
7	;	;	PCTFORTE

Figure V-6 : Erreur d'étiquetage d'un pronom indéfini et de la forme conjuguée *écrémer*.

Il est encore des cas où l'identification de *Cordial* est plus invraisemblable. Ainsi, dans l'exemple suivant, alors qu'il est pourtant suivi d'un pronom relatif et que sa terminaison ne correspond nullement à celle d'une forme infinitive, le substantif féminin « ouvrée » est identifié comme un forme infinitive. Très certainement, il convient là de voir l'effet de l'aspect lacunaire des dictionnaires et grammaires internes de *Cordial*. On peut en effet légitimement supposer que le terme d' « ouvrée », – employé dans l'Est de la France pour désigner l'unité de mesure de la surface de vigne pouvant être bêchée par un paysan en une journée –, est inconnu aux dictionnaires implémentés au logiciel. Ce qui n'explique cependant en aucun cas pourquoi cette forme puisse être analysée en tant que verbe à l'infinitif.

Il faut abandonner les dénominations de journal, de mesure, de carte, d'ouvrée, qui ne correspondent plus ne à rien et dont la valeur varie en chaque endroit.			
N° mot	Dialogue	lemme	Typegram
1	il	il	PPER3S
2	faut	falloir	VINDP3S
3	abandonner	abandonner	VINF
4	les	le	DETDPIG
5	dénominations	dénomination	NCFP
6	de	de	PREP
7	journal	journal	NCMS
8	,	,	PCTFAIB
9	de	de	PREP
10	mesure	mesure	NCFS
11	,	,	PCTFAIB
12	de	de	PREP
13	carte	carte	NCFS
14	,	,	PCTFAIB
15	d'	de	PREP
16	ouvrée	ouvrier	VINF
17	,	,	PCTFAIB
18	qui	qui	PRI
19	ne	ne	ADV

Figure V-7 : Erreur d'identification d'un terme inconnu.

Dans l'exemple suivant, la préposition *outré* est analysée comme la forme conjuguée au présent du subjonctif à la troisième personne du singulier, alors que ce mot se place en position initiale de la phrase, ce qui nous laisse perplexe quant aux règles qui ont présidé à ce choix d'étiquetage :

Outre une collecte faite lundi dernier au repas de noce des époux Garnier-miette, de la Prise-d'Eau, au bénéfice des écoles laïques de Dole,			
N° mot	Dialogue	lemme	Typegram
1	Outre	outrer	VSUBP3S
2	une	un	DETIFS
3	collecte	collecte	NCFS
4	faite	faire	VPARPFS
5	lundi	lundi	NCMS
6	dernier	dernier	ADJMS
7	au	au	DETDMS
8	repas	repas	NCMIN
9	de	de	PREP
10	noce	noce	NCFS
11	des	de	DETDPIG
12	époux	époux	NCMIN
13	Garnier-Mie	Garnier-Mie	NPI
14	,	,	PCTFAIB
15	de	de	PREP
16	la	le	DETDFS
17	Prise-d'Eau	Prise-D'eau	NPI
18	,	,	PCTFAIB
19	au	au	DETDMS
20	bénéfice	bénéfice	NCMS
21	des	de	DETDPIG
22	écoles	école	NCFP
23	laïques	laïque	ADJPIG

Figure V-8 : Erreur d'étiquetage de *outré*.

Un certain nombre d'erreurs récurrentes ont attiré notre attention : elles surviennent dans un contexte univoque, c'est-à-dire lorsque les formes à identifier ne disposent d'aucun homographe polycatégoriel et que la structure syntaxique de la phrase est pour sa part régulière. Les pronoms indéfinis se

révèlent ainsi mal reconnus lorsqu'ils sont précédés d'une coordination. Dans l'exemple suivant, le pronom indéfini *personne* est ainsi analysé par le logiciel comme un nom commun :

Mais personne ne s'étant présenté décidément de la part de M. de Lareinty			
N° mot	Dialogue	lemme	Typegram
1	Mais	mais	COO
2	personne	personne	NCFS
3	ne	ne	ADV
4	s'	se	PPERS3S
5	étant	être	VPARPRES
6	présenté	présenter	VPARPMS
7	décidément	décidément	ADV
8	de	de	PREP
9	la	le	DETFDS
10	part	part	NCFS
11	de	de	PREP
12	M.	m.	NCMIN
13	de	de	PREP
14	Lareinty	Lareinty	NPI

Figure V-9 : Erreur d'étiquetage de la forme *personne*.

Autre catégorie dont les erreurs d'étiquetage ne peuvent être excusées en raison d'une quelconque ambiguïté virtuelle : les mots en plusieurs formes. Très fréquents en langue française, ces derniers sont faiblement reconnus comme unité par *Cordial* : cette classe d'erreurs pointe les lacunes des dictionnaires de *Cordial*.

Faute de reconnaissance de l'unité de la locution adverbiale « En outre », massivement utilisée dans notre corpus, celle-ci est décomposée par *Cordial* comme la succession d'une préposition et d'un nom commun :

En outre, le jeune garçon sera chargé de tenir ce qu'on peut appeler le Livre de culture.			
N° mot	Dialogue	lemme	Typegram
1	En	en	PREP
2	outre	outre	NCFS
3	,	,	PCTFAIB
4	le	le	DETDMS
5	jeune	jeune	ADJSIG
6	garçon	garçon	NCMS
7	sera	être	VINDF3S
8	chargé	charger	VPARPMS
9	de	de	PREP
10	tenir	tenir	VINF
11	ce	ce	PDS
12	qu'	qu	PRI
13	on	on	PPER3S
14	peut	pouvoir	VINDP3S
15	appeler	appeler	VINF
16	le	le	DETDMS
17	Livre	livre	NCMS
18	de	de	PREP
19	culture	culture	NCFS
20	.	.	PCTFORTE

Figure V-10 : Erreur de reconnaissance de l'expression adverbiale *en outre*.

Ne connaissant pas l'expression verbale *porter atteinte*, *Cordial* analyse de façon très étonnante cette collocation comme la succession d'un verbe à l'infinitif et d'un participe passé, qui ne nous évoque aucune règle distributionnelle :

Napoléon se garde bien de porter atteinte à la loi Falloux.			
N°Mot	Dialogue	lemme	Typegram
1	Napoléon	Napoléon	NPMS
2	se	se	PPER3S
3	garde	garder	VINDP3S
4	bien	bien	ADV
5	de	de	PREP
6	porter	porter	VINF
7	atteinte	atteindre	VPARPFS
8	à	à	PREP
9	la	le	NCFS
10	loi	loi	NCFS
11	Falloux	Falloux	NPMS
12	.		PCTFORTE

Figure V-11 : Erreur d'étiquetage de la forme *porter atteinte*.

Enfin, parce que ne disposant d'aucun outil sémantique lui permettant d'établir une différence entre les deux catégoriels, *Cordial* se révèle incapable de distinguer les déterminants indéfinis des numéraux :

M. le ministre de la guerre vient d'envoyer pour être données comme prix aux élèves qui ont fait le plus de progrès dans les cours de gymnastique et d'exercices militaires:			
Une médaille de vermeil et trois mentions honorables au lycée de Besançon			
N°mot	Dialogue	lemme	Typegram
1	Une	un	DETIFS
2	médaille	médaille	NCSFS
3	de	de	PREP
4	vermeil	vermeil	NCMS
5	et	et	COO
6	trois	trois	ADJNUM
7	mentions	mention	NCFP
8	honorables	honorable	ADJPIG
9	au	au	DETDMS
10	lycée	lycée	NCMS
11	de	de	PREP
12	Besançon	Besançon	NPHSIG
13	;		PCTFORTE

Figure V-12 : Exemple de confusion entre un déterminant indéfini et un numéral.

Une autre défaillance de la qualité de l'étiquetage morphosyntaxique dont il reste à mesurer les effets sur notre analyse est justement relative à l'identification des chiffres dont regorge le discours de presse. Un adjectif numéral ne sera pas reconnu comme tel si l'entité qu'il quantifie (heures, minutes, francs, centimes) est mise en ellipse, comme on peut l'observer ci-dessous.

Les membres participants, qui doivent avoir figuré sur les contrôles de l'armée à un titre quelconque, et ne pas avoir dépassé l'âge de quarante-cinq ans, versent 1 fr. 50 par mois et 1 fr. au décès d'un sociétaire participant.			
N° mot	Dialogue	lemme	Typegram
[...]	[...]	[...]	[...]
31	versent	verser	VINDP3P
32	1	1	ADJNUM
33	fr.	fr.	ADJMS
34	50	50	NCMIN
35	par	par	PREP
36	mois	mois	NCMIN
37	et	et	COO
38	1	1	NCMIN
39	fr.	fr.	ADJMS
40	au	au	DETDMS
41	décès	décès	NCMIN

Figure V-13 : Erreur d'étiquetage des nombres.

Ceci nous amène par suite à remettre en cause l'étiquetage des abréviations par *Cordial*. Comme on peut le constater sur la figure ci-dessous, l'abréviation *fr* pour franc est ainsi systématiquement analysée comme un adjectif masculin singulier.

Un autre cas particulier nous amène à remettre en cause l'absence des abréviations dans les dictionnaires du logiciel, catégorie qui se révèle extrêmement utilisée dans le discours de presse : celui de la représentation graphique tronquée des appellatifs tels que *MM.* pour *messieurs*, etc. Ainsi, sur

la figure suivante, on observera que l'abréviation de l'appellatif *messieurs* est analysée comme un nom propre.

MM.HELM et DURAND, conseillers à la cour			
N° mot	Dialogue	lemme	Typegram
1	MM.	mm.	NPMS
2	HELM	Helm	NPI
3	et	et	COO
4	DURAND	Durand	NPMS
5	,	,	PCTFAIB
6	conseillers	conseiller	NCMP
7	à	à	PREP
8	la	la	DETDFS
9	cour	cour	NCFS

Figure V-14 : Etiquetage erroné de la forme *mm*.

Si dans les grammaires traditionnelles, les appellatifs tels que *Monsieur*, *Madame*, *Messieurs* sont faussement considérés comme des noms communs, on peut s'interroger sur la pertinence d'une telle catégorisation. La création d'une catégorie particulière pour ces appellatifs contribuerait notamment à faciliter l'identification des noms propres qui leur succèdent généralement. De surcroît, les noms propres, lorsqu'ils sont virtuellement ambigus en raison de la non-prise en compte de la saillance graphique de la majuscule, sont analysés comme des noms communs.

Le bureau de recensement de Saint-hippolyte a proclamé M. Borne, élu par 850 voix, contre 845 à M. Brocard.			
N° mot	Dialogue	lemme	Typegram
1	Le	le	DETDMS
2	bureau	bureau	NCMS
3	de	de	PREP
4	recensement	recensement	NCMS
5	de	de	PREP
6	Saint-Hippolyte	Saint-Hippolyte	NPSIG
7	a	avoir	VINDP3S
8	proclamé	proclamer	VPARPMS
9	M.	m.	NCMIN
10	Borne	borne	NCFS
11	,	,	PCTFAIB
12	élu	élire	VPARPMS
13	par	par	PREP
14	850	850	ADJNUM
15	voix	voix	NCFIN
16	,	,	PCTFAIB
17	contre	contre	PREP
18	845	845	ADJNUM
19	à	à	PREP
20	M.	m.	NCMIN
21	Brocard	brocard	NCMS
22	.	.	PCTFORTE

Figure V-15 : erreur d'étiquetage des noms propres

Enfin, il convient de remarquer que les scories liées à l'océrisation mais aussi les caractéristiques typographiques des textes du journal génèrent un certain nombre d'erreurs d'annotation. Comme on peut le constater sur les fichiers images du *Petit Comtois*, les voyelles majuscules ne sont pas accentuées dans le quotidien. Cet état de fait engendre des résultats problématiques quant à la reconnaissance d'une préposition telle que « A (À) [+groupe nominal] », qui est annotée comme forme du verbe *avoir* à la troisième personne de l'indicatif présent. L'exemple suivant illustre ce cas d'erreur :

A treize ans, le jeune garçon rentre dans sa famille			
N° mot	Dialogue	lemme	Typegram
1	A	avoir	VINDP3S
2	treize	treize	ADJNUM
3	ans	an	NCMP
4	,	,	PCTFAIB
5	le	le	DETDMS
6	jeune	jeune	ADJSIG
7	garçon	garçon	NCMS
8	rentre	rentrer	VINDP3S
9	dans	dans	PREP
10	sa	son	DETPOSS
11	famille	famille	NCFS

Figure V-16 : Erreur d'étiquetage de la préposition A.

Au terme de cette rapide revue des erreurs les plus fréquentes d'annotation morphosyntaxique de *Cordial* observées dans notre corpus, il nous semble légitime de nous interroger sur les effets de ces erreurs sur les résultats d'une typologie des formes textuelles du discours de presse. Il est coutume de postuler que le taux d'erreurs d'étiquetage de *Cordial* ne saurait suffire à perturber les résultats d'une analyse statistique⁵, dès lors que celle-ci s'attache à faire émerger les tendances lourdes d'un ensemble de données. Les résultats des études inductives menées à partir de cet étiquetage qui parviennent à opposer des types de discours et des genres tendent à confirmer cette logique. Il reste néanmoins à évaluer les conséquences d'un taux d'erreurs estimé à près de 5%, fluctuant selon les types de discours étudiés, quant aux résultats de la caractérisation d'un très grand volume de textes d'un même discours, analysés dans une perspective contrastive et inductive. Ce taux d'erreurs pourrait être fortement atténué par la mise en place d'une interface permettant de contrôler et de modifier systématiquement certaines annotations, tout en offrant la possibilité d'élargir les dictionnaires de l'étiqueteur et de les adapter

⁵ Notamment, F. Rastier et D. Malrieu (2001 : 556), jugent que les taux d'erreurs de *Cordial* « ne sont pas assez forts pour remettre en cause les variations constatées ».

aux particularités du type de discours observé, en intégrant les formes inconnues rencontrées du corpus.

5.2.3. Méthodologie

Afin de les rendre exploitables dans le cadre de notre étude, nous avons extrait les informations relatives aux types grammaticaux, contenues dans la colonne *Typegram*. Ces codes ont ensuite été transformés à partir d'une nomenclature visant à créer de nouvelles catégories plus synthétiques. Les codes grammaticaux complexes constitués d'une succession d'informations relatives à la catégorie grammaticale d'appartenance des items, à leur genre et nombre, ou à la personne, au temps grammatical (exemple : NCFS, VINDPS2P) ont ainsi été analysés et décomposés à partir d'une nomenclature. Les nouvelles classes morphosyntaxiques ont alors été codées comme suit :

Code	Légende
1	Marque de première personne du singulier
2	Marque de deuxième personne du singulier
3	Marque de troisième personne du singulier
4	Marque de première personne du pluriel
5	Marque de deuxième personne du pluriel
6	Marque de troisième personne du pluriel
ADJ	Adjectif
ADJIND	Adjectif indéfini
ADJNUM	Adjectif numéral
ADJORD	Adjectif ordinal
ADV	Adverbe
COO	Coordination
DETD	Déterminant défini
DETDEM	Déterminant démonstratif
DETI	Déterminant indéfini
DETPOSS	Déterminant possessif

F	Féminin
INT	Interjection
M	Masculin
NC	Nom commun
NP	Nom propre
P	Pluriel
PCTFAIB	Ponctuation faible
PCTFORTE	Ponctuation forte
PD	Pronom démonstratif
PHRA 1	Entre 0 et 8 mots
PHRA 2	Entre 8 mots et 16 mots
PHRA 3	Entre 16 mots et 24 mots
PHRA 4	Entre 24 mots et 32 mots
PHRA 5	Plus de 32 mots
PI	Pronom interrogatif
PPER1	Pronom personnel de première personne du singulier
PPER2	Pronom personnel de deuxième personne du singulier
PPER3	Pronom personnel de troisième personne du singulier
PPER4	Pronom personnel de première personne du pluriel
PPER5	Pronom personnel de deuxième personne du pluriel
PPER6	Pronom personnel de troisième personne du pluriel
PR	Pronom relatif
PREP	Préposition
S	Singulier
SUB	Subordination
UEUPH	Chaîne euphonique
V	Verbe
VCON	Verbe au conjonctif
VFUT	Verbe au futur
VIMP	Verbe à l'impératif
VIMPFT	Verbe à l'imparfait
VINF	Verbe à l'infinitif
VPPR	Verbe au participe présent
VPPS	Verbe au participe passé
VPRES	Verbe au présent
VPS	Verbe au passé simple
VSUBP	Verbe au subjonctif présent

Pour chaque article, les occurrences de chaque trait morphosyntaxique ont ainsi été comptabilisées. Ces données ont ensuite été intégrées dans une matrice contenant en lignes les 29 676 articles de notre corpus et en colonnes les variables morphosyntaxiques. À ces dernières, il nous a semblé pertinent d'ajouter deux variables supplémentaires que nous jugions *a priori* discriminantes, relatives à la longueur de la phrase (nombre de phrases répondant aux seuils de 8, 16, 24 et 32 mots) et au nombre de mots comptés dans l'article. Pour identifier les caractéristiques discursives de chaque article et permettre ensuite des regroupements, la rubrique d'appartenance et la date de parution de chaque article ont été intégrées à cette gigantesque matrice. Celle-ci, ne pouvant être soumise en l'état aux procédures de l'analyse factorielle des correspondances en raison de sa taille et de sa dispersion, est conçue pour construire de nouvelles matrices, plus réduites et plus concentrées, répondant aux objectifs des analyses envisagées successivement ci-dessous.

Nombreuses sont les analyses de typologie textuelle qui recourent à la Classification Hiérarchique Ascendante (CAH) (Rastier & Malrieu, 2001, Beauvisage, 2001). Cette méthodologie présente l'intérêt de permettre une visualisation très nette des classes automatiquement construites par la statistique. Cette méthodologie, dite aussi classification par agrégation, permet d'obtenir, à partir d'un ensemble d'éléments décrits par des variables, une hiérarchie de classes partiellement « emboîtées » les unes dans les autres. Le principe de cette analyse est de regrouper des éléments dont la similarité est détectée à partir du calcul de la *distance de Ward* : à chaque étape de la procédure, les éléments les plus proches sont fusionnés en *cluster* et ainsi de suite. Pour plus de détails sur cette méthodologie, il convient de se reporter à L. Lebart et A. Salem (1994) rappelant en ces termes le principe méthodologique de la Classification Hiérarchique :

On part d'un ensemble de n éléments (que l'on appellera ici des éléments de base ou encore *éléments terminaux*) dont chacun possède un poids, et entre lesquels on a calculé des distances (il y a alors $n(n-1)/2$ distances entre les différents couples possibles). On commence par agréger les deux éléments les plus proches. Le couple ainsi agrégé constitue alors un *nouvel élément* dont on

peut recalculer à la fois le poids et les distances à chacun des éléments qu'il reste à classer. À l'issue de cette étape, le problème se trouve ramené à celui de la classification de $n-1$ éléments. On agrège à nouveau les deux éléments les plus proches, et l'on réitère ce processus ($n-1$ fois au total) jusqu'à épuisement de l'ensemble des éléments. L'ultime et ($n-1$)ème opération regroupe l'ensemble des éléments au sein d'une classe unique (Lebart & Salem, 1994 : 112-113).

Pour notre part, nous avons opté pour soumettre les données du *Petit Comtois* aux méthodes d'analyse factorielle des correspondances (Lebart & Salem, 1994). Cette classe d'analyse présente en effet l'avantage précieux de hiérarchiser l'information sur les distributions des traits morphosyntaxiques, sans présenter l'inconvénient de la dichotomiser. En donnant à lire sur un continuum la distribution en diverses variables des traits morphosyntaxiques, l'AFC permet en effet de dépasser une appréhension dichotomique de classes de rubriques partageant (ou non) le même profil distributionnel morphosyntaxique. De fait, l'AFC nous permet d'évaluer les similarités de profils morphosyntaxiques qu'entretiennent rubriques et traits morphosyntaxiques en termes de gradient. Dans cette perspective, le mode de visualisation géodésique présente l'intérêt de pallier les difficultés d'interprétation indéniables des sorties classiques AFC en deux dimensions, trop souvent illisibles en leur centre, et permet d'appréhender de façon optimale la spatialisation des relations entre points, notamment grâce à la possibilité de construction de zooms centrés.

5.3. Analyse de la distribution des catégories grammaticales dans les rubriques du Petit Comtois

Dans un premier temps, nous n'avons sélectionné qu'un nombre restreint de variables, se réduisant aux parties du discours, dont P. Guiraud dès 1954, montrait que la distribution fluctuait fortement en fonction des genres et des époques.

En vue d'examiner leur ventilation dans la structure externe du rubriquage du *Petit Comtois*, nous avons opéré certaines transformations dans la matrice fonctionnelle d'origine. Le nombre des lignes de la matrice a été réduit en rapportant chaque article à sa rubrique d'appartenance. Seules les quarante rubriques les plus fréquentes et durables du *Petit Comtois* ont été conservées. Le nombre de traits morphosyntaxiques a pour sa part été ramené aux seules parties du discours (*Adjectif, Adverbe, conjonctions de coordination et subordination, déterminant défini, nom commun, nom propre, préposition, verbe*). La figure ci-dessous donne une représentation des associations entre lignes et colonnes de cette matrice.

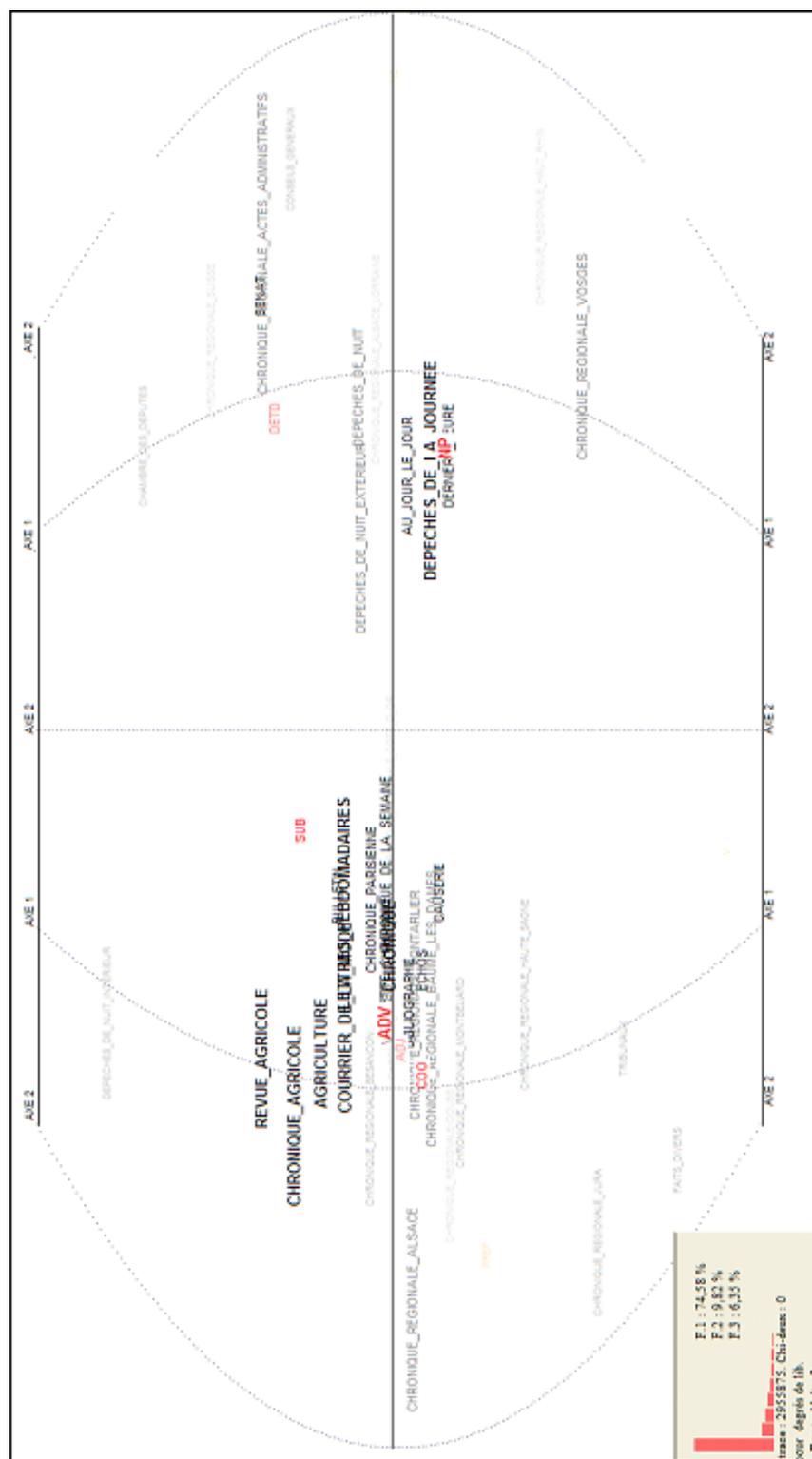


Figure V-17: Visualisation géométrique de l'analyse de la ventilation des parties du discours dans les rubriques du corpus.

On remarquera que le premier facteur exprime près de 75% de la variance totale. Il établit un clivage assez nettement défini entre d'une part, un groupe constitué des dépêches (*Dépêches de la journée, Dernière heure, Dépêches de Nuit-Extérieur, Au jour le jour*) sur la moitié droite de l'axe 1 et d'autre part, les rubriques dérivées de la forme générique de la chronique (*Chronique parisienne, Chronique, Petite chronique*) ainsi que les rubriques basées sur la forme épistolaire (*Courrier de la mode, Lettres hebdomadaires*) à gauche. La grande majorité des chroniques régionales (*Besançon, Pontarlier, Baume-les-Dames, Doubs, Haute-Saône, Alsace, Jura*) relèvent également de cette moitié gauche, bien qu'elles présentent des profils moins saillants⁶.

Une opposition se donne ainsi à lire entre les variables *noms propres* et *déterminants définis* d'une part, et les variables *adverbes, adjectifs et coordinations* d'autre part. Cette configuration des parties du discours est stimulante à plusieurs titres. Tout d'abord, contrairement à ce qu'on constate le plus souvent dans une étude du même type sur corpus clos⁷, la bipolarité observée ne traduit pas une opposition entre le groupe nominal et le groupe verbal. Contre toute attente, les variables *noms communs* et *verbes* revêtent de surcroît des positions peu saillantes. On vérifiera les profils peu dynamiques de ces variables dans le tableau de leur distribution, justifiant l'absence de zone d'influence marquée par le groupe nominal et le groupe verbal :

⁷ Voir M. Kastberg Sjöblom (2002 : 299), E. Brunet (1985 : 155).

	NC (en éc.réd)	V (en éc.réd)
AGRICULTURE	-2,9	-1,3
AU_JOUR_LE_JOUR	-7,3	-1,4
BIBLIOGRAPHIE	-1,4	-5,7
BULLETIN	-6,5	1,6
CAUSERIE	-5,4	-1,5
CHAMBRE_DES_DEPUTES	2,7	-2,2
CHRONIQUE	-7,0	1,1
CHRONIQUE_AGRICOLE	1,5	-0,1
CHRONIQUE_DE_LA_SEMAINE	-7,6	4,6
CHRONIQUE_PARISIENNE	-2,6	-7,0
CHRONIQUE_REGIONALE	-0,2	1,3
CHRONIQUE_REGIONALE_ACTES_ADMINISTRATIFS	3,3	-3,4
CHRONIQUE_REGIONALE_ALSACE	1,2	2,7
CHRONIQUE_REGIONALE_ALSACE_LORRAINE	-0,2	-1,2
CHRONIQUE_REGIONALE_BAUME_LES_DAMES	-2,9	5,7
CHRONIQUE_REGIONALE_BESANCON	8,2	-4,4
CHRONIQUE_REGIONALE_COTE_D_OR	-1,9	1,3
CHRONIQUE_REGIONALE_HAUT_RHIN	5,1	0,8
CHRONIQUE_REGIONALE_Haute_SAONE	-3,0	4,2
CHRONIQUE_REGIONALE_JURA	4,0	2,2
CHRONIQUE_REGIONALE_MONTBELIARD	3,0	-1,0
CHRONIQUE_REGIONALE_PONTARLIER	-2,1	3,9
CHRONIQUE_REGIONALE_SUISSE	-3,6	-5,9
CHRONIQUE_REGIONALE_VOSGES	0,7	2,6
CHRONIQUE_REGIONALE-DOUBS	3,4	3,7
CONSEILS_GENERAUX	4,0	-1,6
COURRIER_DE_LA_MODE	-1,4	-5,9
DEPECHES_DE_LA_JOURNEE	-0,7	0,1
DEPECHES_DE NUIT	10,3	-4,2
DEPECHES_DE NUIT_EXTERIEUR	-17,1	0,7
DEPECHES_DE NUIT_INTERIEUR	-2,4	-1,7
DERNIERE_HEURE	-0,3	1,8
ECHOS	-1,7	4,5
FAITS_DIVERS	-3,0	15,6
LETTRES_HEBDOMADAIRES	-8,3	-0,2
PETITE_CHRONIQUE	-6,8	-1,6
REVUE_AGRICOLE	-0,4	0,2
SENAT	5,0	-3,6
TRIBUNAUX	-2,1	3,7
VARIETES	-6,4	-3,0

Tableau V-1 : Distribution des noms communs et verbes (écarts-réduits).

Si les déterminants définis s'inscrivent dans la partie droite du planisphère, ainsi que les noms propres, le point *adjectifs* s'inscrit quant à lui dans la zone opposée, là même où se concentrent les adverbes, les coordinations et les

adjectifs. Il est d'ailleurs tout à fait intéressant d'observer que les adverbes n'apparaissent pas corrélés aux verbes⁸, comme le confirme le graphique ci-dessous montrant la distribution des deux variables, dont l'incompatibilité est très nette.

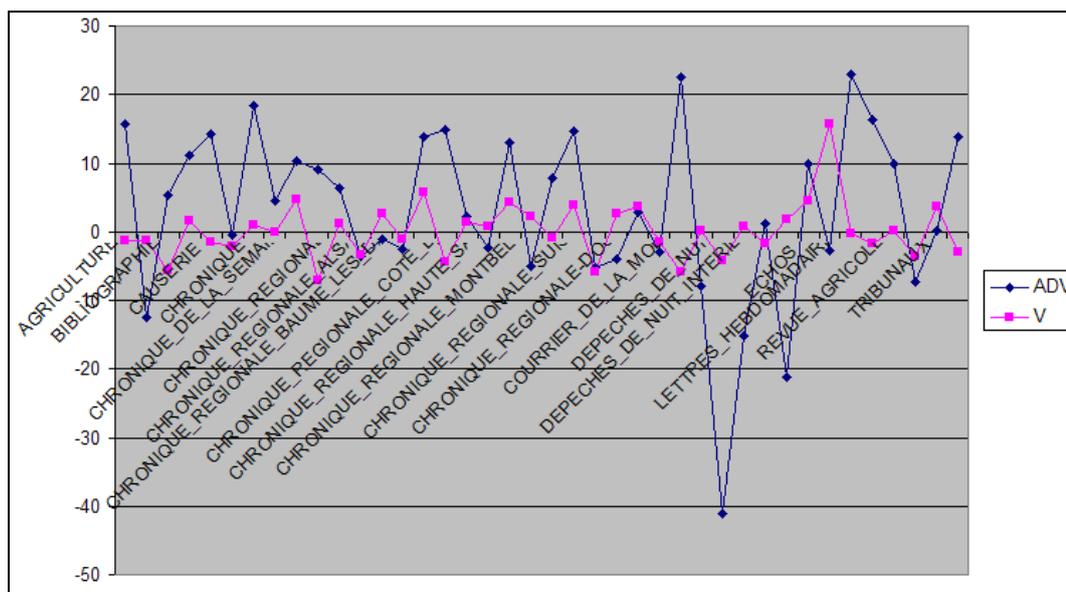


Figure V-18 : Comparaison de la distribution des adverbes et des verbes dans le corpus (écarts-réduits).

Les adverbes s'avèrent en revanche corrélés aux adjectifs, comme la comparaison des courbes distributionnelles de ces variables, au parallélisme évident, présentant de forts écarts-réduits dans les mêmes rubriques.

⁸ Test de corrélation de Pearson : Valeur observée : 0,030 ; p-value bilatérale : 0,856 ; Seuil de signification : 0,05.

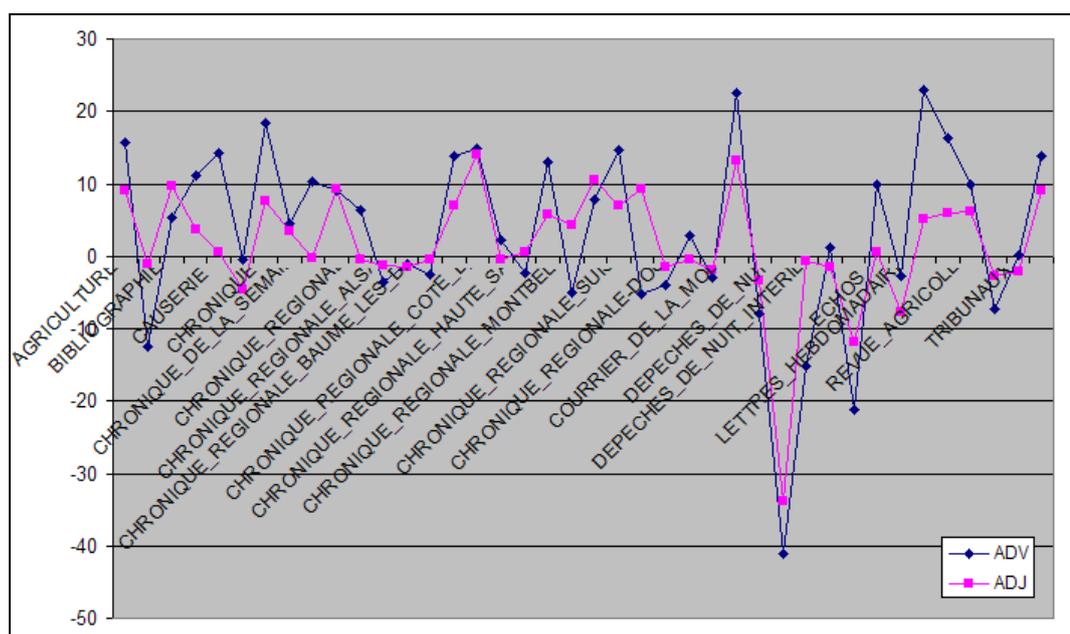


Figure V-19 : Comparaison de la distribution des adverbes et des adjectifs dans le corpus (écarts-réduits).

De fait, il nous semble préférable d'envisager l'opposition mise au jour par le premier facteur comme la confrontation de deux champs, l'un formé autour des noms propres et des déterminants définis⁹, l'autre articulant des variables relatives à un fort degré d'inscription de subjectivité du locuteur et au degré de complexité syntaxique de la phrase, indiqué par les conjonctions de coordination. De façon tout à fait surprenante, l'examen des corrélations des variables des parties du discours nous apprend également que les déterminants définis ne sont pas ici corrélés aux noms communs¹⁰, ce que l'on constate sur le graphique ci-dessous qui représente les profils distributionnels divergents des deux variables. Cette observation nous semble être le premier indice d'une insuffisance de l'étiquetage automatique. Comme nous l'avons signalé ultérieurement, *Cordial* catégorise comme nom commun la forme « M. » devant un nom propre. Cette erreur a pour conséquence d'augmenter

¹⁰ Test de corrélation de Pearson : Valeur observée : 0,039 ; p-value bilatérale : 0,813 ; Seuil Alpha= 0,05.

considérablement la corrélation non-significative *nom commun/nom propre* et de diminuer dans la même mesure la corrélation qu'on pourrait attendre entre *nom commun* et *déterminant*.

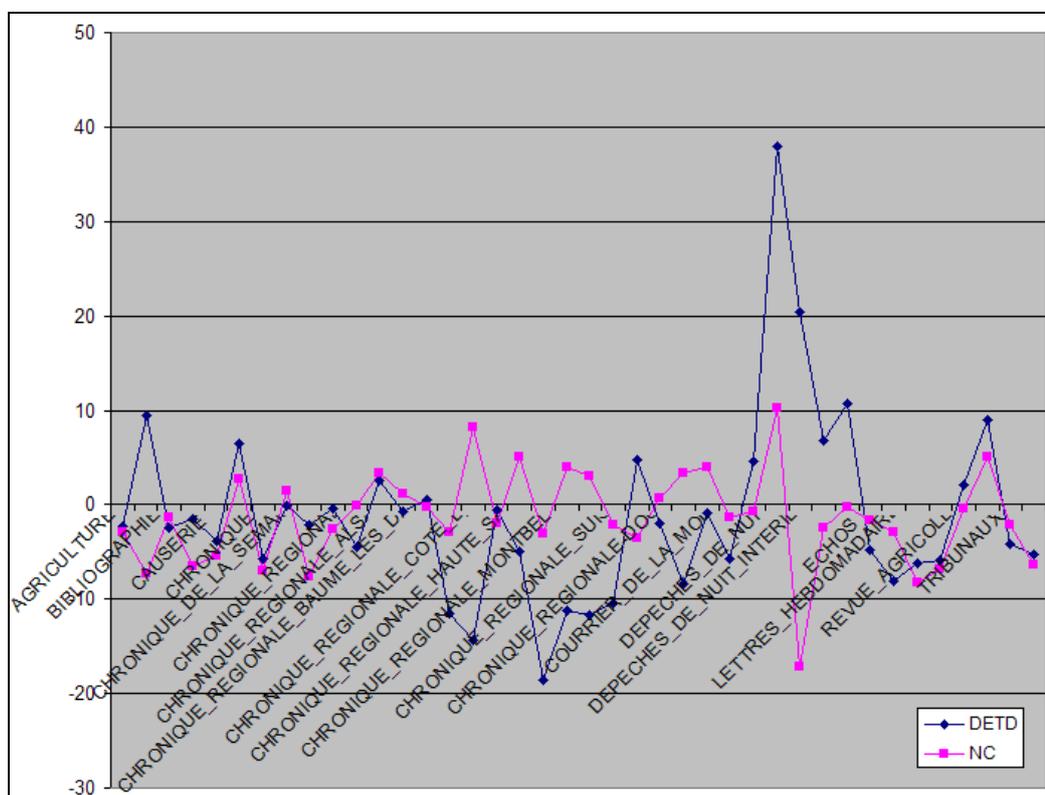


Figure V-20 : Comparaison de la distribution des déterminants définis et des noms communs dans le corpus (en écarts-réduits).

5.4. Analyse de la distribution d'un panel large de variables morphosyntaxiques dans les rubriques du *Petit Comtois*

5.4.1. Analyse de la distribution de 52 variables morphosyntaxiques dans les rubriques du *Petit Comtois*

En vue de préciser le profil de chaque rubrique et d'évaluer la pertinence d'autres regroupements de variables morphosyntaxiques, une matrice des mêmes quarante rubriques en ligne a été construite, intégrant cette fois l'ensemble des traits morphosyntaxiques présentés ci-dessus (52 traits), diminué de quelques colonnes aux effectifs faibles ou redondants¹¹.

La figure V-21, résultant de cette analyse, se montre très cohérente avec les indications livrées par la ventilation des seules parties du discours. L'inertie modérée du facteur 1 (55,34%) engendre un nuage de points plus dispersé, et de fait, une visualisation plus fine de la ventilation des variables. Sur ce facteur, on observe un clivage entre les pronoms du discours et les marques de l'inscription de la subjectivité du locuteur (*adverbes, déterminants possessifs*) d'une part, et les *déterminants définis* et les *noms communs* d'autre part. On retrouve ainsi l'opposition mise au jour par le premier facteur de l'analyse des parties du discours entre d'une part, les rubriques *Petite chronique, chronique, chronique de la semaine, chronique parisienne, Courrier de la mode, Lettres hebdomadaires*, les chroniques régionales et d'autre part, les cases rédactionnelles caractérisées par une désinscription énonciative à savoir la quasi-totalité des *dépêches*.

Le deuxième facteur (12,58%) détermine une gradation entre une dimension historico-narrative forte, reposant sur le couple *imparfait* et *passé-simple* dans le quart « ouest » à gauche du graphique, et une dimension embrayée, marquée par le couple *présent* et *participe passé*, qui réfère au présent de l'énonciation et qui signale le passé composé, accompli du présent, dans le quart est à droite. Dans la moitié gauche du planisphère, ce facteur individualise tout spécialement dans le quart ouest les rubriques *Faits Divers*,

¹¹ Exemple : la macro-catégorie « pronom personnel » a été supprimée car elle était redondante avec l'information apportée par celles des 6 (3x2) sous-classes personnelles.

Tribunaux, *Variétés* et *Causerie*, très fortement marquées par les temps de l'imparfait et du passé-simple.

Si le premier facteur contribue principalement à construire une bipolarité entre les rubriques à forte prise en charge énonciative et les rubriques où les traces de l'énonciation sont gommées, le facteur 3 (10,42%) intervient quant à lui principalement dans la partie gauche pour configurer un gradient entre les rubriques caractérisées par l'emploi de la première personne du singulier (telles les *Lettres hebdomadaires* et la *Petite chronique*) dans la partie nord, et celles marquées par la première personne du pluriel (*Chronique régionale Montbéliard*, *Besançon*, *Baume-les-Dames*, *Pontarlier*), situées dans la partie sud du graphique. Dans la moitié droite de la géode, le facteur 3 ordonne les longueurs de phrases. Il individualise notamment « au nord » les variables indiquant l'emploi de phrases courtes (respectivement moins de 8 mots et de 8 mots à 16 mots).

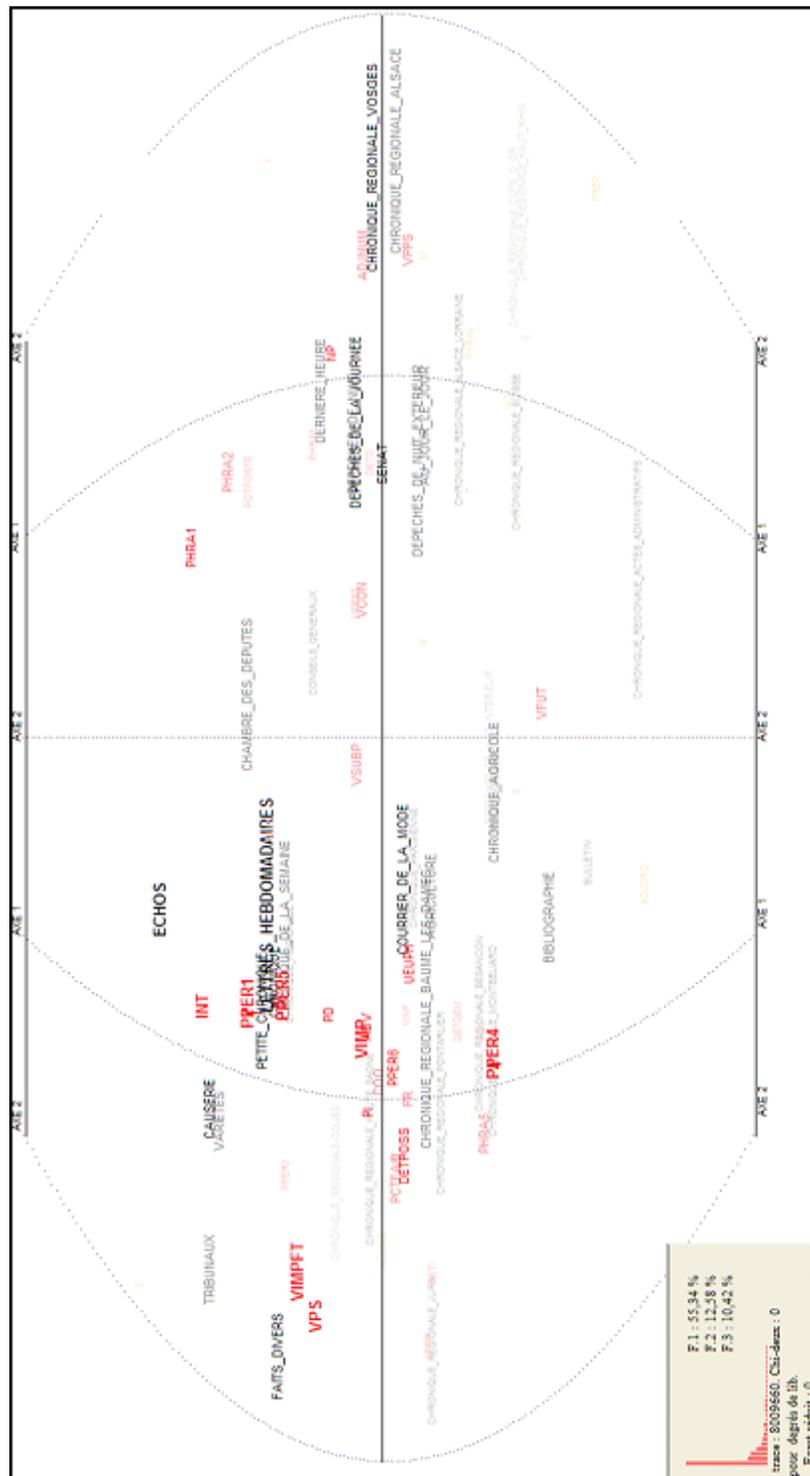


Figure V-21 : Visualisation géodésique de l'analyse de la ventilation des variables morphosyntaxiques dans les rubriques.

Afin de rendre compte du dynamisme très important des contrastes induits par cette analyse, nous procéderons maintenant à une description détaillée des variables les plus discriminantes.

5.4.2. Description des facteurs fortement discriminants

Les personnes

Parce qu'ils contribuent très fortement aux facteurs 1 et 3, nous examinerons tout d'abord les pronoms personnels avant d'envisager les différents marqueurs qui leur sont associés.

Tous les pronoms personnels n'ont pas le même impact discriminant : les pronoms de troisième personne du singulier et du pluriel présentent des profils distributionnels respectivement peu dynamique et non-significatif¹². Seuls les pronoms personnels de première personne du singulier et du pluriel, ainsi que de deuxième personne du pluriel, ont un fort poids sur les trois facteurs. Partant, on peut formuler l'hypothèse que les traces de la subjectivité du locuteur sont plus discriminantes que les indices du mode du récit. L'histogramme suivant (figure V-22) rend compte des profils de distribution respectifs des pronoms personnels de première personne du singulier et de deuxième personne du pluriel, profils tous deux extrêmement dynamiques et fortement congruents.

Sans surprise, les profils distributionnels de ces pronoms établissent un clivage entre les rubriques où les traces de la subjectivité du locuteur sont pleinement assumées et les rubriques où ces dernières sont effacées, où se localisent respectivement les valeurs extrêmement positives et négatives des variables. Ainsi, les pronoms de première personne du singulier et le pronom d'adresse connaissent une utilisation intensive dans les rubriques *Lettres hebdomadaires* (33,53 pour PPER1), *Variétés* (22,20 pour PPER1), *Causerie*, *Petite chronique*.

¹² Et ce, tout particulièrement concernant la troisième personne du pluriel : son profil distributionnel est extrêmement neutre : les 3/5 de ses écarts réduits sont non significatifs puisque 26 de ses écarts-réduits sur 40 sont en effet compris entre -2 et 2.

A *contrario*, les emplois les plus faibles de ces variables sont à trouver du côté des dépêches : *Dépêches de nuit* (-18,40 pour PPER1 ; -21,62 pour PPER5), *Dépêches de nuit-Extérieur* (-20,06 pour PPER1 ; -16,73 pour PPER5), *Dernière heure*, *Dépêches de la journée*. On remarquera qu'à l'intérieur de cette catégorie, la sous-rubrique *Dépêches de nuit-Intérieur* s'individualise en ne présentant pas un profil significatif (-0,20 pour PPER1 ; -1,19 pour PPER5). Remarquons que cette opposition ne saurait à elle seule résumer la signification de ces profils : on lit également sur cet histogramme un gradient entre les articles contenant ou non des phases de discours rapporté : ainsi, les rubriques *Échos* (15,57 pour PPER1) et *Chambre des députés* (5,43 pour PPER1; 11,55 pour PPER5) témoignent de valeurs positives parce qu'elles comportent des dialogues.

De façon plus surprenante, la distribution des pronoms de première personne du singulier et de deuxième du pluriel divise en deux sous-groupes la classe des chroniques régionales : des sous-rubriques comme *Chronique régionale Besançon* (12,89 pour PPER1 ; 13,47 pour PPER5), *Baume-les-Dames* (4,88 pour PPER1; 9,18 pour PPER5), *Doubs*, *Haute-Saône* accusent des emplois assez intenses des pronoms étudiés. À l'opposé, ils sont déficitaires dans les sous-rubriques *Suisse* (-10,39 pour PPER1 : -8,86 pour PPER5) *Vosges* (-3,76 pour PPER1 ; -2,63 pour PPER5), *Haut-Rhin* (-6,64 pour PPER1 ; -1,45 pour PPER5), *Alsace-Lorraine*, *Alsace* et *Côte-d'Or*. Il apparaît donc que les sous-rubriques relevant d'une zone plus éloignée du lieu d'édition du quotidien présentent des profils significativement déficitaires des pronoms de première et de deuxième du pluriel, tandis que les chroniques relatives à la sphère directement rapprochée du lieu d'édition sont pour leur part caractérisées par leur forte utilisation.

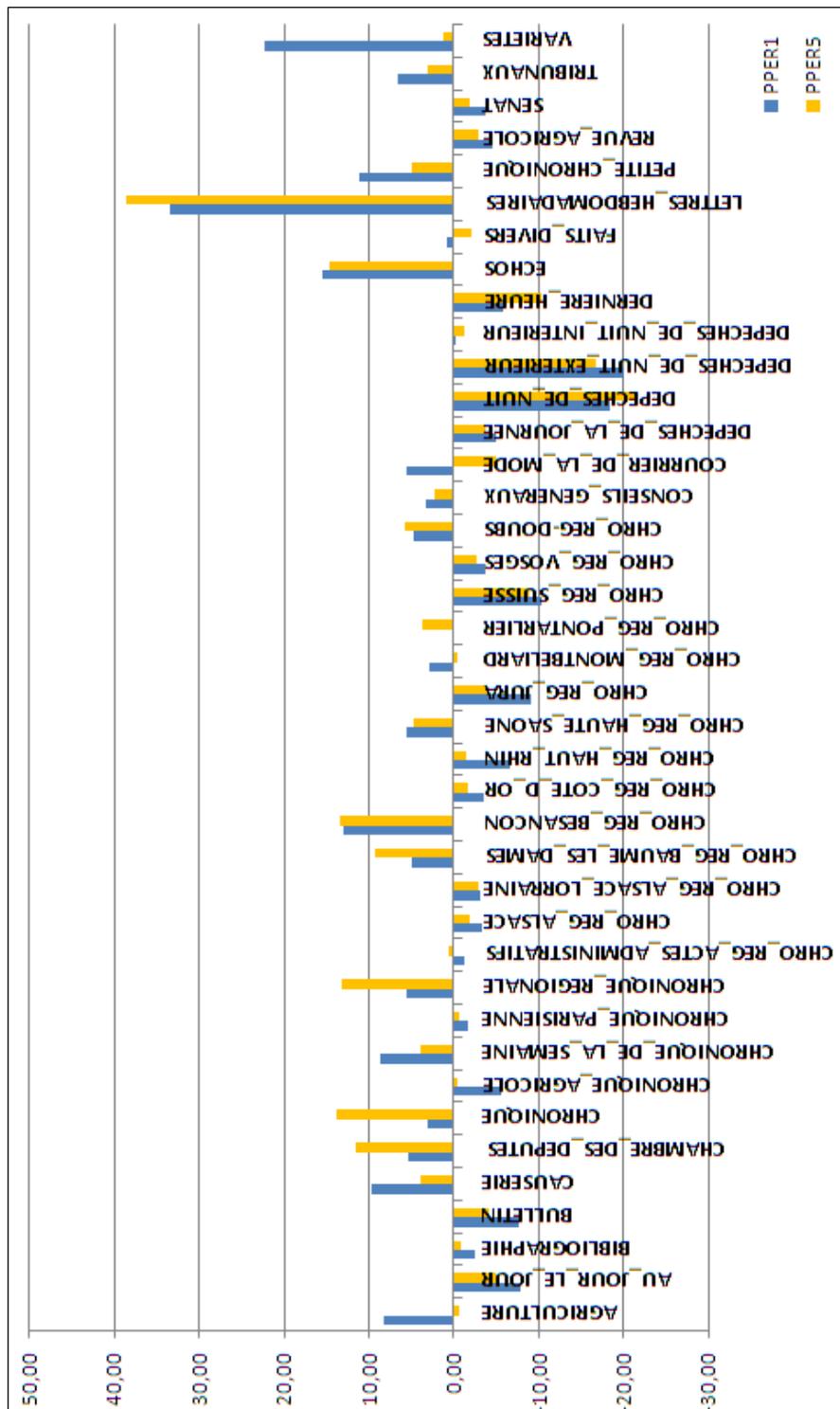


Figure V-22 : Histogramme des profils distributionnels du pronom de 1^{ère} personne du singulier et de 2^{ème} personne du pluriel (en écarts-réduits).

Le pronom « nous » contribue principalement à opposer la classe des dépêches à celle des chroniques régionales : on notera son profil extrêmement dynamique entre sous-emplois pour les *Dépêches de nuit* (-49,40), *Dépêches de nuit-Extérieur* (-21,32), *Dépêches de la journée*, *Dernière Heure* et sur-emplois dans les chroniques régionales : *Chronique régionale-Besançon* (44,39), *Chronique régionale-Baume-Les-Dames* (21,24), etc. Le profil de ventilation de ce pronom est fortement convergent avec celui des pronoms personnels « je » et « vous », puisqu'il met également au jour un clivage selon la zone d'origine des nouvelles des sous-rubriques des chroniques régionales : *Chronique régionale- Haute-Saône*, *Montbéliard*, *Pontarlier*, *Doubs* présentent ainsi de forts écarts-réduits positifs tandis que les rubriques renvoyant loin du siège se caractérisent par des profils distributionnels déficitaires, à l'instar de *Chronique régionale- Alsace* (-3, 18), *Suisse* (-10,70), *Côte d'Or*, *Jura*, *Vosges*. De façon marginale, la rubrique *Chronique régionale-Haut-Rhin* présente un écart-réduit positif significatif sur cette variable (3,32).

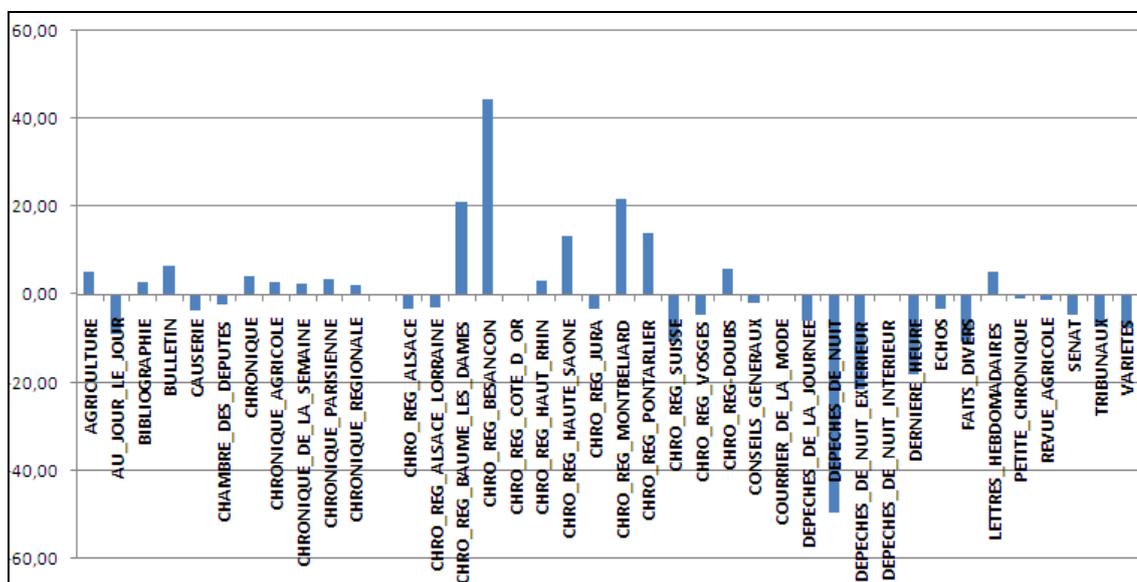


Figure V-23 : Histogramme du profil distributionnel du pronom de 1ère personne du pluriel (en écarts-réduits).

La longueur des phrases

Les variables relatives à la longueur de la phrase ont un poids important dans la structuration du troisième facteur. Les profils de ces variables sont extrêmement dynamiques à l'exception de celui des phrases de longueur comprise entre 24 et 32 mots (PHRA4)¹³. Que la longueur des phrases soit un facteur discriminant est dans une certaine mesure une indication attendue : elle renvoie aux critères d'opposition des genres journalistiques.

Les deux premières classes de phrases courtes (entre 0 et 8 mots, 8 et 16 mots) présentent des profils très dynamiques et similaires. Leurs profils convergent pour mettre au jour une opposition entre les rubriques caractérisées par un emploi intensif des phrases courtes (*Dernière heure, Dépêches de nuit*) et celles privilégiant les phrases plus longues. Les phrases entre 16 et 24 mots confirment également ce clivage, en présentant toutefois un profil nettement moins dynamique. À l'inverse, le groupe des dépêches est fortement déficitaire (-17, 87 pour *Dépêches de Nuit*) à l'aune des phrases de taille très longue (plus de 32 mots)¹⁴ tandis que les valeurs positives extrêmes des phrases longues sont à trouver dans le groupe des chroniques régionales : *Chronique régionale-Besançon* (10, 16), *Chronique régionale Jura* (6,73), et dans une moindre mesure, dans les rubriques *Bulletin* (3, 95), *Chronique de la semaine* (3, 39), *Lettres hebdomadaires* (3, 65).

Il est à cet égard très dommageable que le repérage du passif, et plus encore de la structure dite « passif abrégé » (sans complément d'agent) soit absent des grilles d'analyse de *Cordial*, car ces structures doivent être supposées accompagner la brièveté des phrases dans les diverses rubriques journalistiques. Il s'agit d'un point aveugle des grammaires traditionnelles, qui n'est que partiellement comblé à ce jour dans les plus novatrices adoptant un point de vue discursif (Riegel *et al*, 1994 : 437-440).

¹³ 31 valeurs d'écart réduit sur 40 ne sont pas significatives, tandis que la plus forte valeur accusée par cette variable est de (+3,62).

¹⁴ Seule la rubrique *Dépêches de nuit- Intérieur* réalise un écart réduit positif de 3, 82 à l'intérieur de cette classe

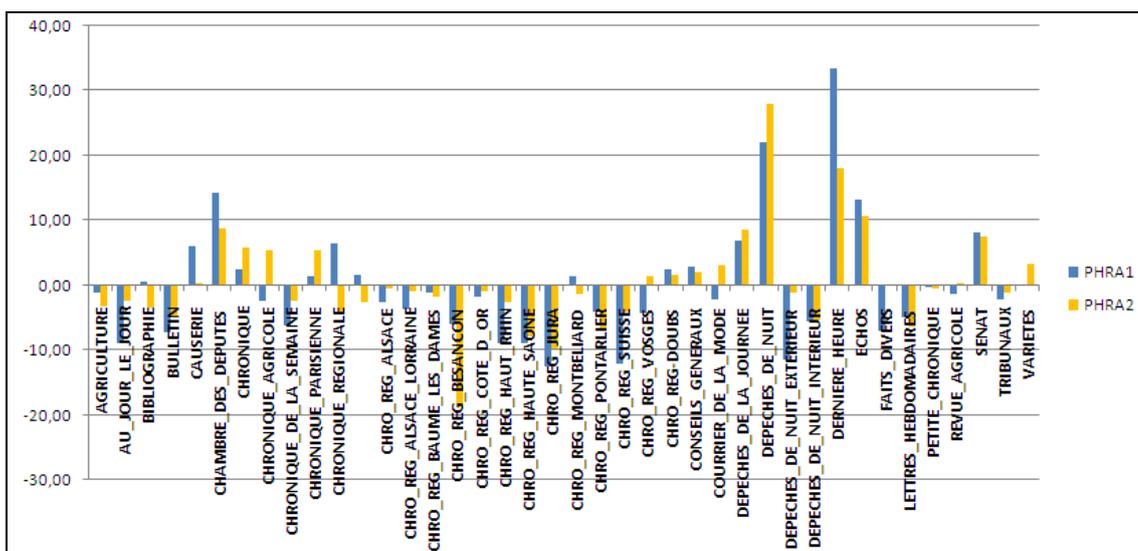


Figure V-24 : Histogramme de la distribution des classes de phrases courtes (PHRA1 et PHRA2) (en écarts-réduits).

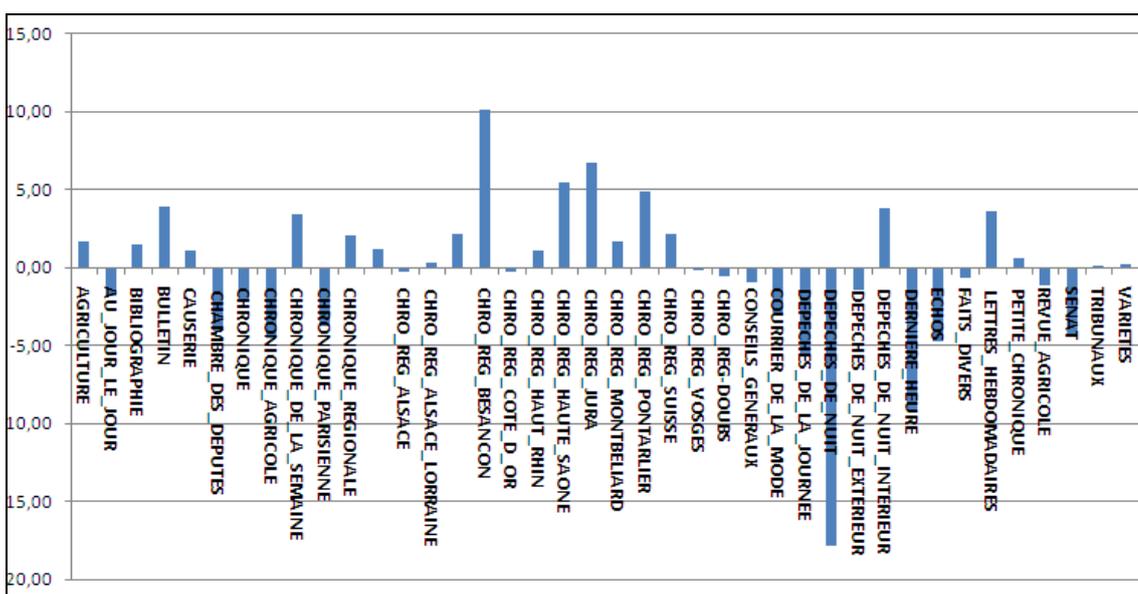


Figure V-25 : Histogramme du profil distributionnel de la classe PHRA5 (en écarts-réduits).

La ponctuation

Les variables de ponctuation forte et faible prennent un fort poids dans les trois premiers facteurs de l'analyse de la ventilation des variables morphosyntaxiques.

La ponctuation faible s'avère fortement caractéristique du groupe des chroniques régionales, sans recouper le clivage constaté par les histogrammes précédents entre les sous-rubriques éloignées ou non du lieu d'édition du quotidien.

Quant à la ponctuation forte, elle connaît un profil distributionnel largement déficitaire dans le groupe des chroniques régionales ainsi que dans les *Lettres Hebdomadaires*, pour se révéler fortement utilisée dans les dépêches¹⁵, et tout spécialement dans *Dépêches de nuit*, *Dépêches de la journée*, *Dernière heure*. On observera la forte corrélation entre les écarts-réduits positifs significatifs de la ponctuation forte (figure V-26) et ceux des phrases courtes (cf. figure V-24). Cette corrélation est évidemment triviale.

¹⁵ *Dépêches de nuit-Intérieur* et *Dépêches de nuit-Extérieur* se distinguent de cette tendance générale en présentant des profils négatifs ou non significatifs (respectivement -6,70 pour *Dépêches de nuit-Intérieur* ; -1,21 pour *Dépêches de nuit-Extérieur*).

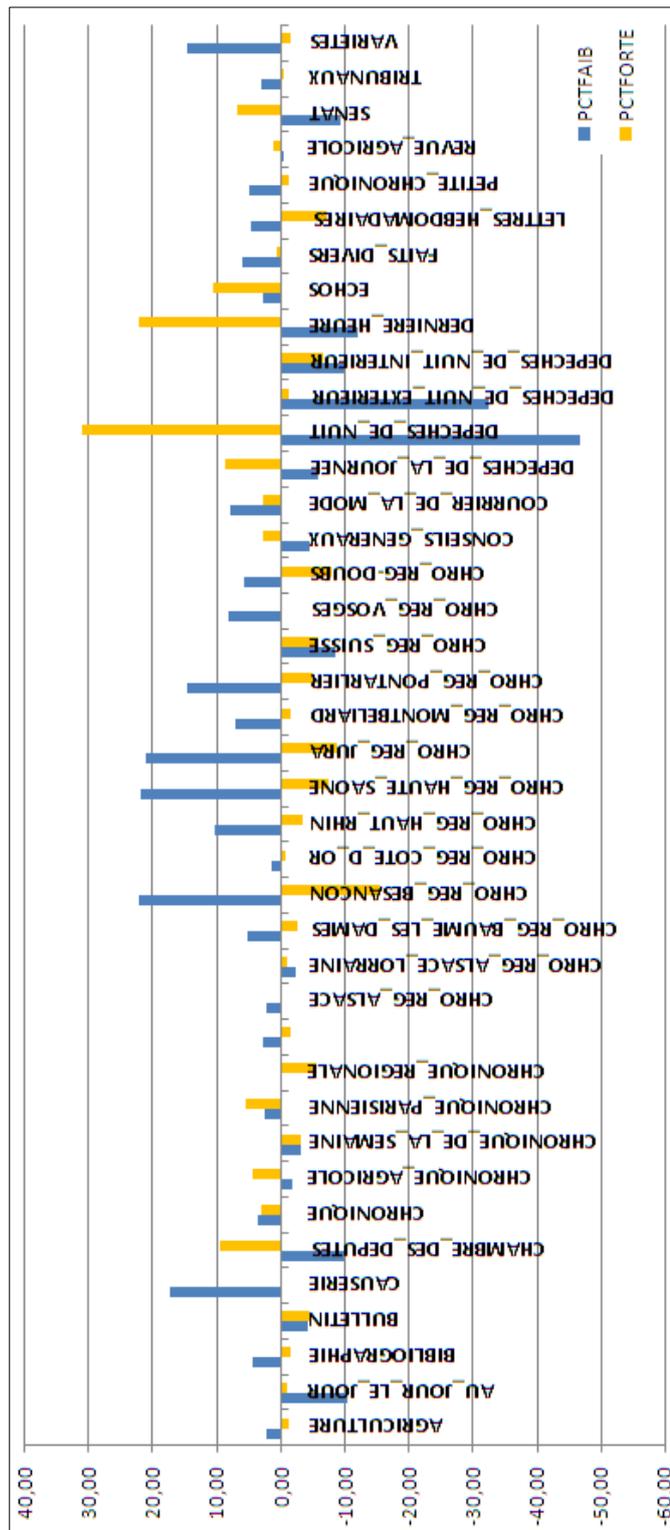


Figure V-26 : Histogramme de la ventilation de la ponctuation faible et forte (en écarts-réduits).

Les adverbes

Cette variable est sujette à caution, en ceci que la classe correspondante est floue en linguistique et en grammaire et que *Cordial* épouse cette indécision. Ayant un fort poids sur les facteurs 1 et 2, la variable *adverbes* présente un profil distributionnel marqué. Les valeurs positives extrêmes de cette variable sont à trouver dans *Lettres hebdomadaires* (22,71), *Petite Chronique* (16,17), de *Chronique parisienne*, *Chronique de la Semaine*, *Bulletin*, ainsi que du côté des *Causeries* et autres *Variétés* ; les plus faibles dans le groupe des dépêches. À l'intérieur du groupe des chroniques régionales, la variable *adverbes* établit une distinction convergente avec la variable des pronoms personnels de première personne du singulier et de deuxième personne du pluriel. Cet indice concerne plutôt les critères de repérage des « adverbes » par *Cordial*, que des propriétés de notre corpus déjà bien établies par l'observation, plus fiable, des pronoms. En effet, le profil distributionnel des adverbes distingue nettement les sous-rubriques « locales », proches du lieu d'édition du quotidien (*Baume les dames* (11,6), *Besançon*, *Montbéliard* (7,34), *Haute-Saône*, *Pontarlier* (14)), des sous-rubriques *Haut-Rhin* (-3,53) et *Jura* (-5,78) notamment.

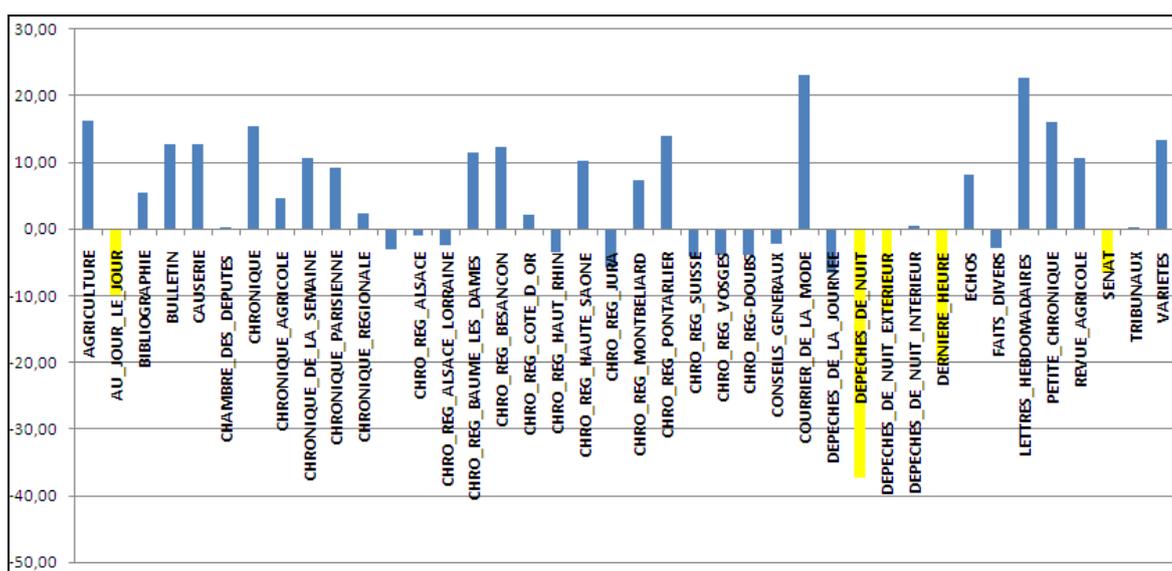


Figure V-27 : Histogramme de la distribution des adverbes (en écarts-réduits).

Les noms

À l'intérieur de la catégorie des noms, les noms propres présentent un profil distributionnel nettement plus marqué que celui des noms communs. Nous émettons de grandes réserves quant à la fiabilité de ces indications, en raison des erreurs massives de reconnaissance de cette catégorie par *Cordial*, que nous avons déjà signalées en début de chapitre. Comme l'indique l'histogramme ci-dessous, les noms propres, tout comme les noms communs, participent à individualiser le groupe des dépêches (*Dépêches de nuit* (52,26), *Dépêches de nuit-Extérieur* (30,61)¹⁶ et la rubrique *Au jour le jour* (23,28) des autres rubriques, qui présentent des profils très largement déficitaires.

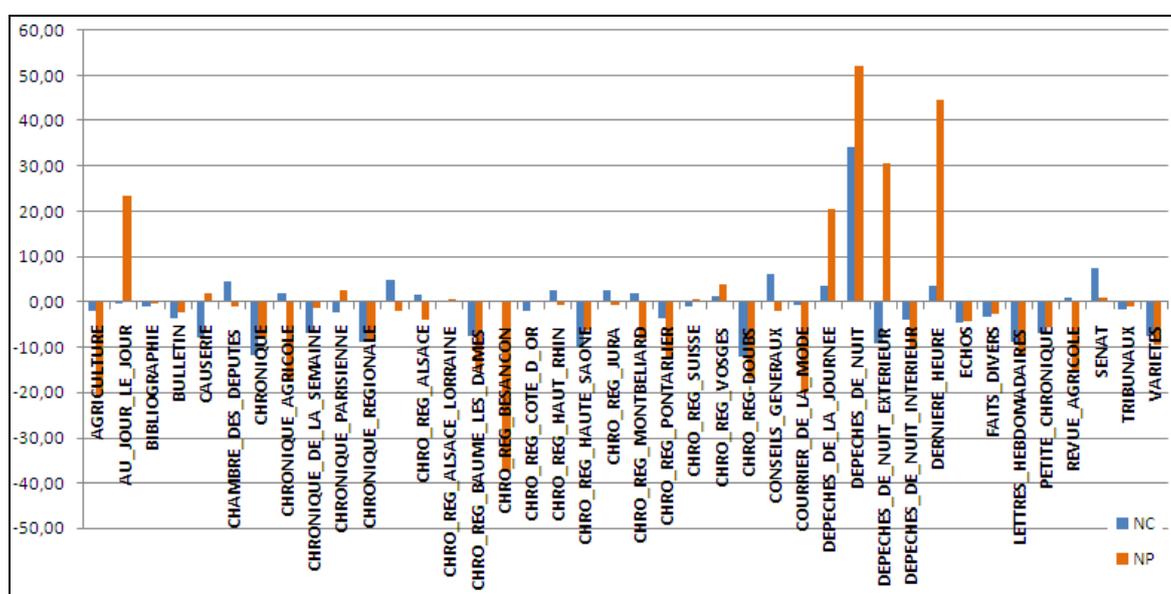


Figure V-28 : Histogramme de la distribution des noms communs et noms propres (en écarts-réduits).

Les déterminants

Deux classes de déterminants sont particulièrement discriminantes : les déterminants possessifs et les déterminants définis.

¹⁶ On remarquera que la rubrique *Dépêches de nuit-Intérieur*, qui présente un sous-emploi des noms propres (-10,44), s'individualise de cette tendance générale.

Les déterminants possessifs se révèlent principalement utilisés dans les rubriques où les marques de première personne du singulier et du pluriel étaient particulièrement sollicitées, ce qui était fort prévisible. Ils établissent le même clivage entre le groupe des dépêches et le groupe des chroniques régionales, tout en opposant encore une fois les sous-rubriques des pages régionales selon que celles-ci concernent un lieu proche ou éloigné du lieu d'édition du *Petit Comtois*. Les emplois les plus importants des déterminants possessifs se concentrent ainsi dans les sous-rubriques « locales » comme *Chronique régionale Besançon* (21,98), *Chronique régionale Montbéliard* (17,68). Les rubriques *Variétés* (14,25), *Petite Chronique* (8,19) et *Chronique* (5,7) se caractérisent également par un profil distributionnel positif des déterminants possessifs.

La variable *déterminants définis* est très nettement caractéristique du groupe des dépêches : son plus fort écart réduit est à trouver dans la case *Dépêches de nuit* (59,81), tandis que les valeurs sont également fortement positives à l'intérieur de ce même groupe de rubriques (*Dépêches de nuit-Extérieur* : 27,83 ; *Dernière heure* : 13,96 ; *Dépêches de la journée* : 7,84 ; *Au jour le jour* : 15,46) et *Sénat* : 11). À l'opposé, les chroniques régionales sont très largement caractérisées par un sous-emploi de ce type de déterminant, tout comme les rubriques *Causerie* (-5,75) et *Lettres hebdomadaires* (-6,54). Comment interpréter l'impact discriminant du déterminant défini ?

Si l'on se rapporte aux études d'E. Brunet et aux propositions des grammairiens tournés vers les linguistiques énonciatives (Riegel *et al*, 1994 : 151 ssqq), E. Brunet (1988b : 193) fait observer que « l'abstraction intéresse plus l'article défini que l'indéfini (on dit moins souvent une justice que la justice) » ; d'où il ressortirait que les textes caractérisés par un sur-emploi des déterminants définis seraient l'indice d'une tendance à l'abstraction. Compte tenu des défauts de reconnaissance de *Cordial*, et du temps dont nous avons disposé pour approfondir ce point très stimulant, il convient de rester prudent, dans la mesure où cette interprétation pourrait se révéler sensible aux genres explorés en

corpus. Par contrecoup, la discussion et l'exploitation de ce point pourraient revêtir un intérêt tout particulier dans un programme de recherche sur ce corpus.

Le sur-emploi des déterminants définis dans certaines rubriques de notre corpus telles que les *Dépêches de la journée*, ou la *Dernière Heure*, nous semble plutôt refléter les modalités d'une écriture journalistique qui présuppose la connaissance d'un élément référentiel et investit de fait l'article défini d'une valeur anaphorique (en un sens étendu, à savoir un renvoi au discours antérieur par-dessus les frontières d'articles, de rubriques et de dates).

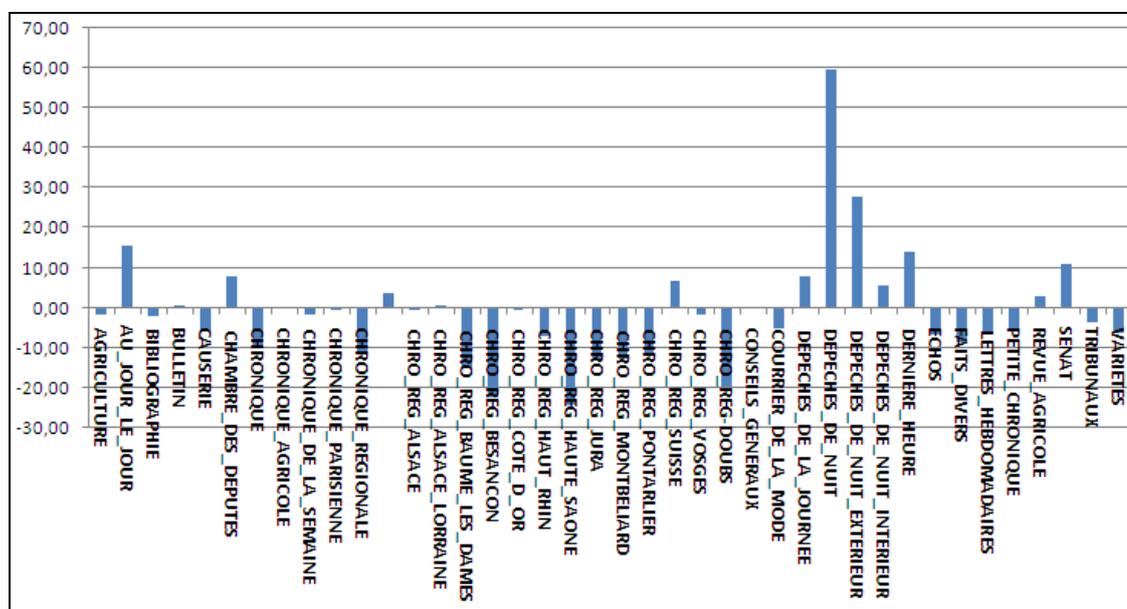


Figure V-29 : Histogramme de la distribution des déterminants définis (en écarts-réduits).

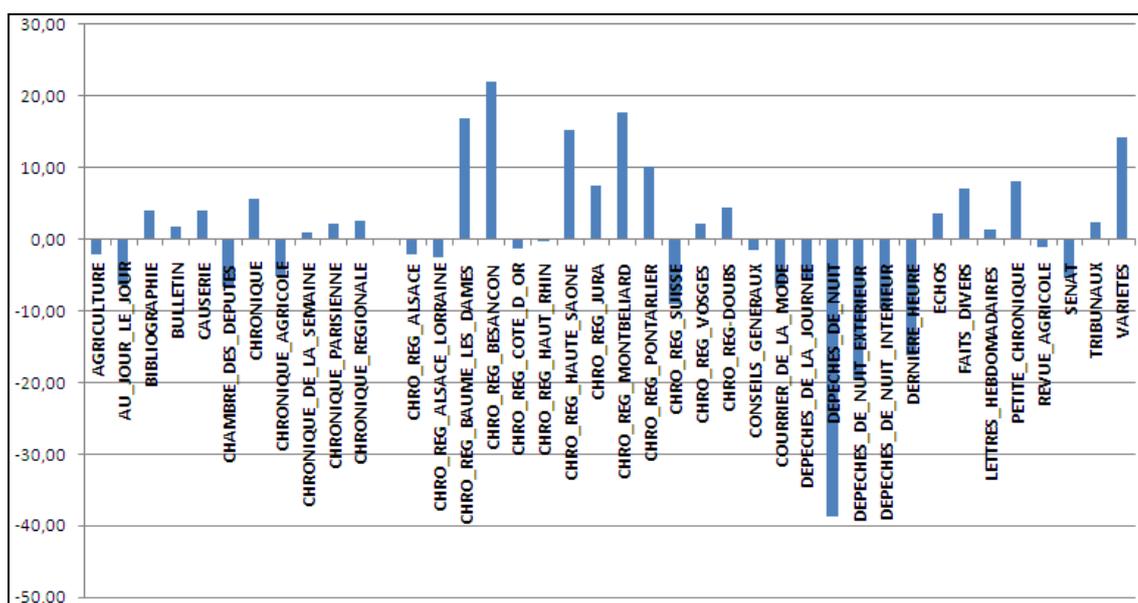


Figure V-30 : Histogramme de la distribution des déterminants possessifs (en écarts-réduits).

Adjectifs et adjectifs numéraux

Les histogrammes parallèles de la distribution des adjectifs et adjectifs numéraux, qui prennent respectivement une part importante dans la structuration du premier et du deuxième facteur, et du premier et du troisième facteur, présentent la curiosité d'avoir des profils distributionnels symétriques. Les adjectifs sont ainsi fortement sur-employés dans les rubriques *Bibliographie* (9,77), *Chroniques*, *Lettres hebdomadaires* (4,8), *Petite chronique* (5,8), *Revue agricole* (6,8), *Variétés* (8,4) et *Courrier de la mode* (13,70), tandis que leur profil distributionnel dans les dépêches est très déficitaire (-27 pour *Dépêches de nuit*). À l'inverse, les adjectifs numéraux enregistrent de forts écarts-réduits positifs là même où les adjectifs qualificatifs connaissent leurs valeurs négatives les plus extrêmes, et plus spécialement dans *Dernière heure* (23,62) et *Dépêches de nuit* (18,17). *Dépêches de Nuit Extérieur* et *Intérieur* se distinguent de ces dernières rubriques en présentant des emplois déficients des adjectifs numéraux. La rubrique *Dépêches de Nuit Extérieur* se caractérise d'ailleurs par un sur-emploi relativement important des adverbes (3,8).

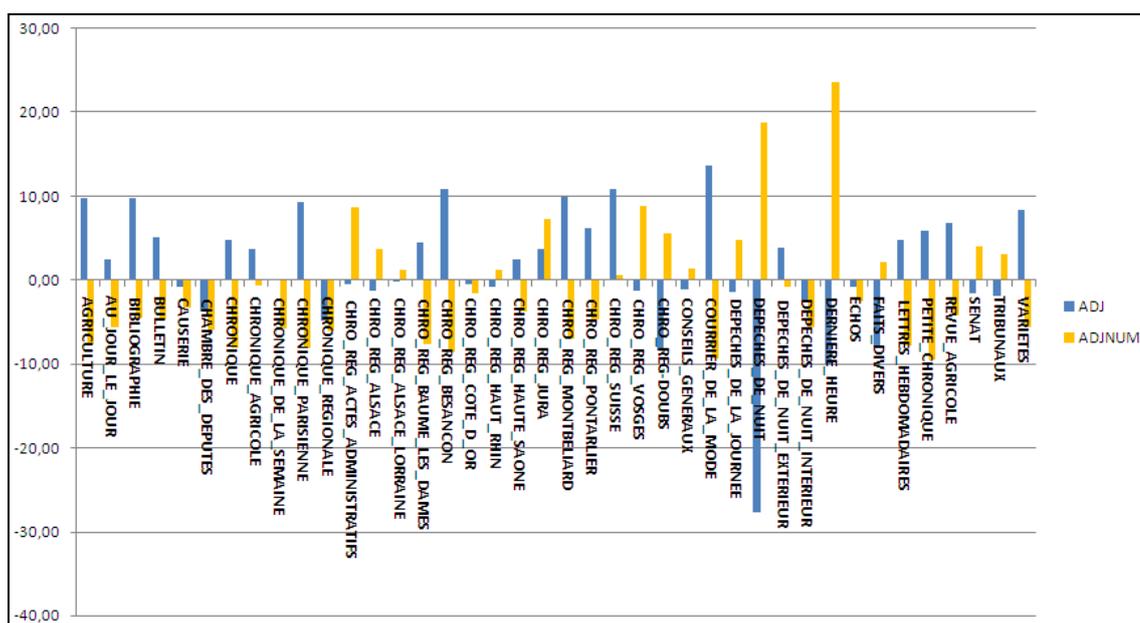


Figure V-31 : Histogramme de la distribution des adjectifs et adjectifs numéraux (en écarts-réduits).

Les pronoms (autres que personnels)

Les pronoms relatifs et démonstratifs contribuent à la structuration des deux premiers facteurs par des profils très dynamiques.

Les valeurs positives maximales de la répartition des pronoms démonstratifs sont à trouver dans les rubriques dont les intitulés sont basés sur une forme générique. On citera à titre d'exemple les écarts-réduits positifs importants des *Lettres hebdomadaires* (14,11), *Échos* (10,25), *Petite chronique* (8,62), *Courrier de la mode*, *Chronique parisienne*, *Chronique agricole*, *Chronique de la semaine*, *Chronique*, *Causerie*, *Variétés* (5,17), *Bulletin*, *Agriculture*, etc. À la lumière de la distribution des pronoms démonstratifs, les chroniques régionales se divisent selon le même clivage déjà observé : les chroniques régionales proches du lieu d'édition du quotidien sont caractérisées par un emploi intensif des pronoms démonstratifs (*Chronique régionale*, *Chronique régionale Baume les Dames*, *Chronique régionale Besançon*, *Chronique régionale Haute-Saône*, *Chronique régionale- Montbéliard*, *Chronique régionale- Pontarlier*), tandis que les chroniques régionales relatives à des zones plus éloignées de Besançon

(*Alsace, Haut-Rhin, Jura, Vosges*) témoignent d'un sous-emploi de ce type de pronoms. La convergence, prévisible, de ce critère avec les déterminants définis ne fait que conforter notre interprétation. Il est d'autre part malaisé, surtout dans l'optique d'une annotation automatique comme celle de *Cordial*, de faire la part entre les « pronoms démonstratifs » et les structures relatives (« celui/celle/ceux qui/à qui/dont... »).

La distribution des pronoms relatifs rend sensiblement compte de la même dynamique générale que celle établie par la ventilation des pronoms démonstratifs. Il est tout d'abord du groupe des dépêches qu'elle concourt à opposer aux *Lettres hebdomadaires, Variétés, Causerie, Bulletin, Bibliographie, Agriculture, Faits Divers, Chronique de la semaine* et aux chroniques régionales. Le profil des pronoms relatifs confirme ensuite le clivage précédemment observé dans le groupe des chroniques régionales : les chroniques régionales proches du lieu d'édition du quotidien présentent des valeurs positives tandis que les sous-rubriques relatives à des zones plus éloignées de Besançon ne semblent pas recourir aux pronoms relatifs (*Chronique Suisse* (-3,49)).

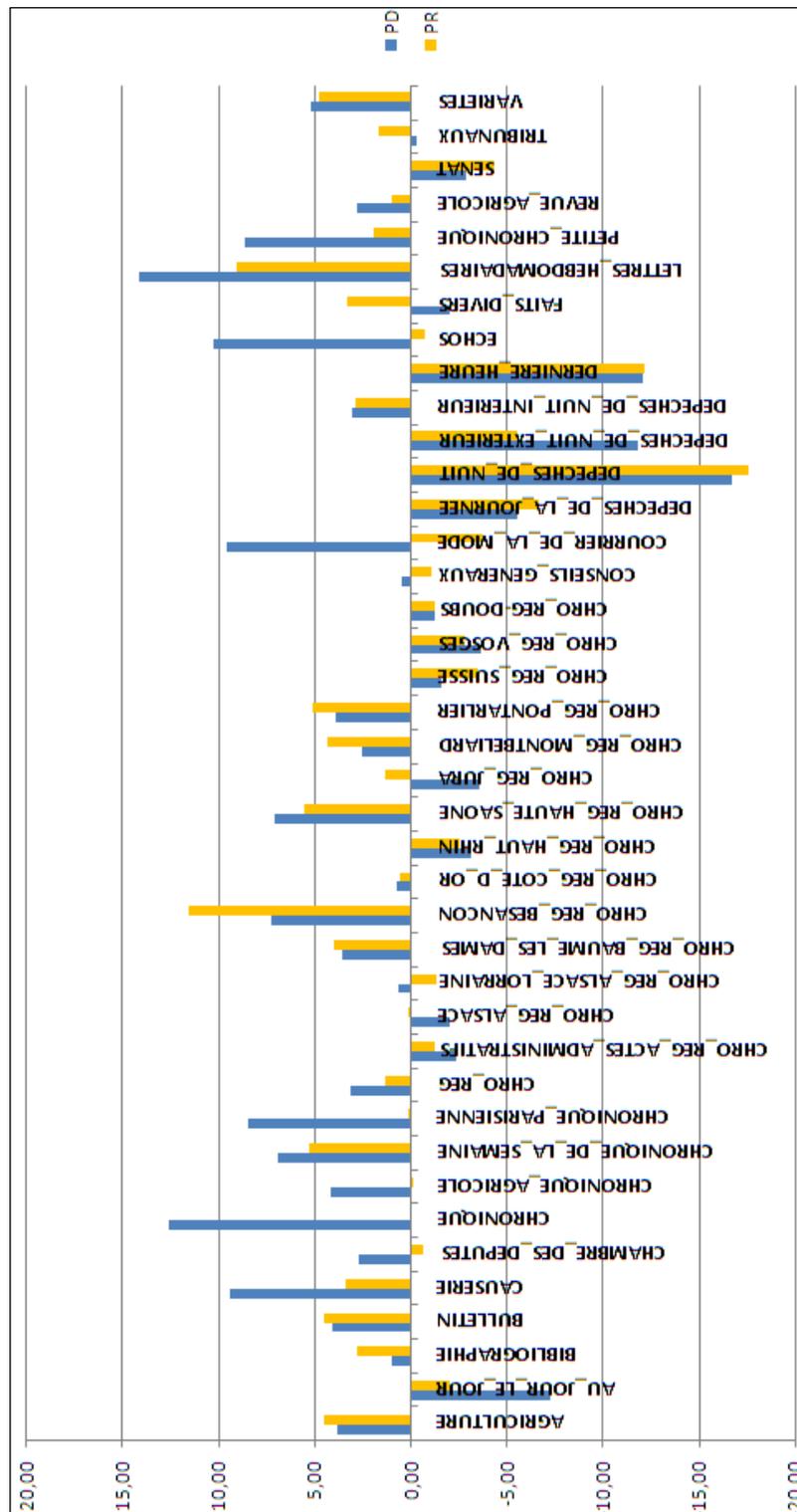


Figure V-32 : Histogramme parallèle de la ventilation des *pronoms démonstratifs* et des *pronoms relatifs* (en écarts-réduits).

On notera, en toute logique, que la variable pronom relatif connaît un profil distributionnel symétrique avec celui des phrases courtes (entre 8 et 16 mots), comme le représente la figure ci-dessous :

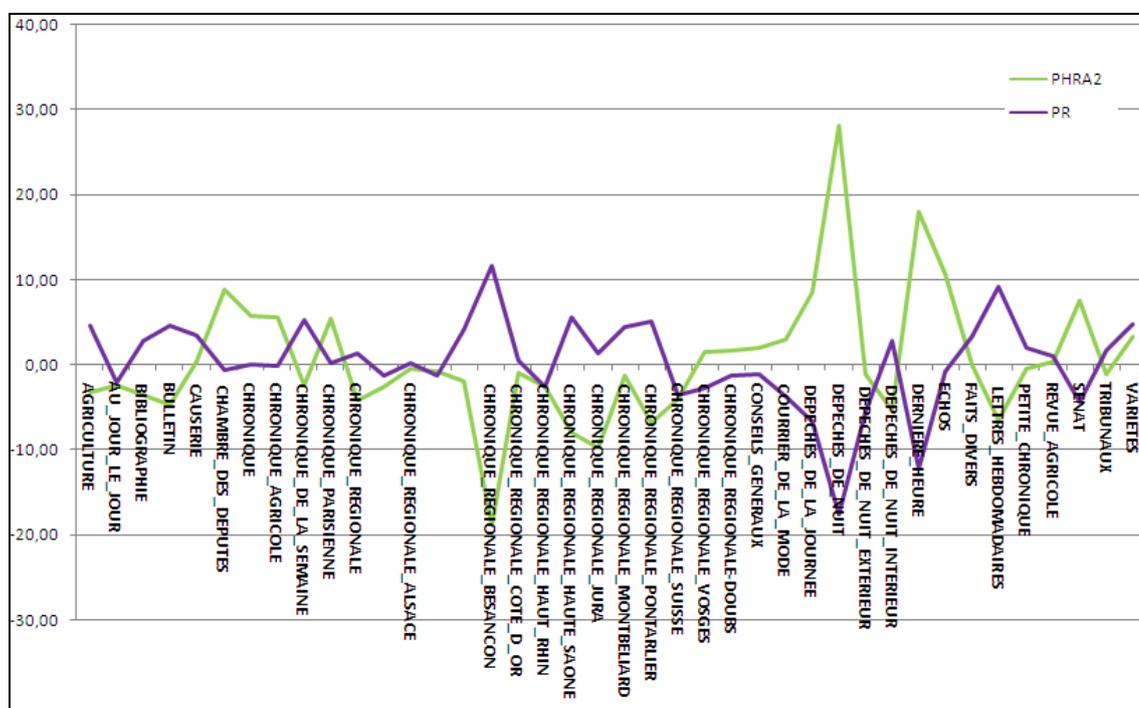


Figure V-33 : Profils distributionnels des phrases courtes et les pronoms relatifs (en écarts-réduits).

Les coordinations

L'examen de la distribution des coordinations, variable au poids conséquent sur le facteur 1, laisse apparaître les mêmes dynamiques que l'étude des pronoms relatifs et démonstratifs, mais surtout avec les phrases longues, et plus particulièrement les phrases de plus de 32 mots (figure V-34).

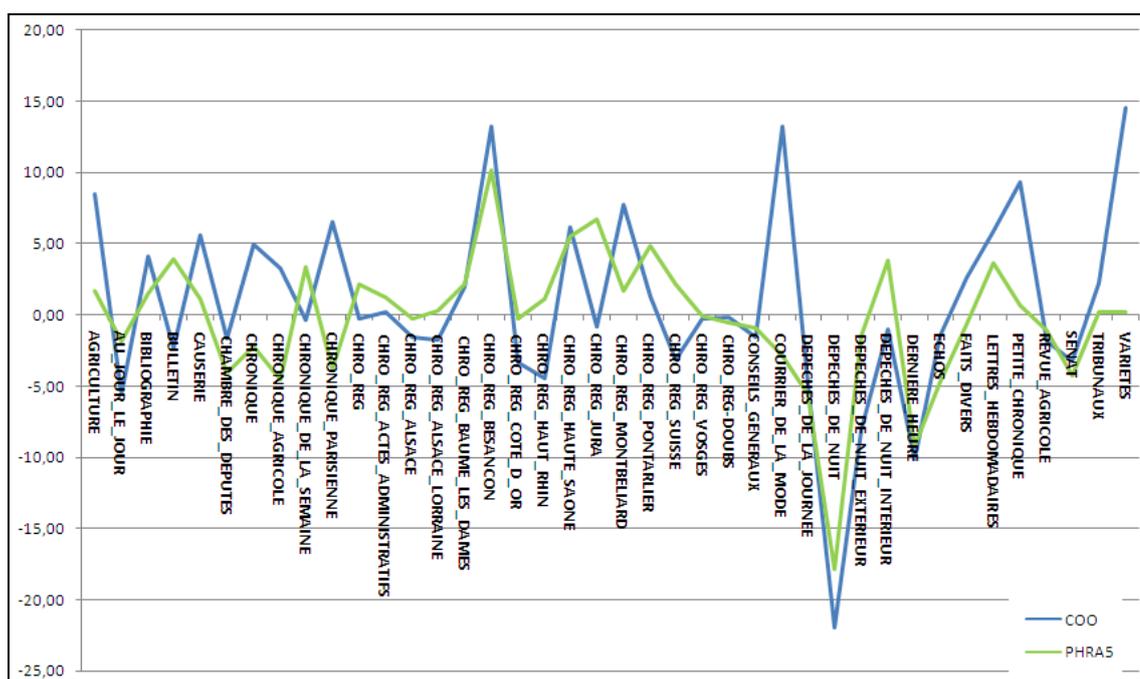


Figure V-34 : Corrélation entre les coordinations et les phrases de plus de 32 mots (en écarts-réduits).

On note ainsi les valeurs positives maximales de la variable *coordinations* dans les rubriques *Variétés*, *Petite chronique*, *Lettres hebdomadaires*, *Courrier de la mode* (13,26), *Chronique*, *Chronique parisienne*, *Causerie*, *Bibliographie*, *Agriculture* tandis que les valeurs négatives maximales sont à trouver dans la macro-rubrique des dépêches, qui présente un profil déficitaire indifféremment des différences de rubriques et sous-rubriques. Selon la ventilation de cette variable, les chroniques régionales sont divisées selon la désormais bien repérée opposition entre chroniques régionales « locales » et chroniques régionales des zones plus éloignées. Les chroniques régionales « locales » sont ainsi caractérisées par un excédent tandis que les autres chroniques régionales présentent un déficit (*Chronique régionale Haut-Rhin* : -4,46).

Les temps verbaux

Les temps verbaux contribuent pour une part très importante aux trois premiers facteurs, pour l'imparfait et le passé simple aux deux premiers, pour le

participe passé et le futur, au second et troisième ; au second pour le présent et le conditionnel, enfin. On observe un clivage des rubriques suivant les temps verbaux qu'elles privilégient, qui évoque l'opposition entre les systèmes de l'énonciation de discours et de l'énonciation historique (Benveniste, 1966 : 237-250).

Comme on peut l'observer sur l'histogramme suivant, les profils distributionnels de l'imparfait et du passé-simple sont extrêmement semblables et convergent pour indiquer une dimension historico-narrative dans les rubriques des *Faits divers* (32,73), *Variétés* (29,32), *Tribunaux* (14,21), *Causerie* (13,19), mais également dans la *Chronique régionale Jura* (13,83) et les *Chronique régionale Haute Saône*, *Chronique régionale Vosges*.

Ces temps suggèrent un mode de narration-diégétisation non actualisé, que nous qualifierons à la suite de J.-M. Adam de « diégétisation autonome » ou « énonciation historique » (Adam, 2005a : 198).

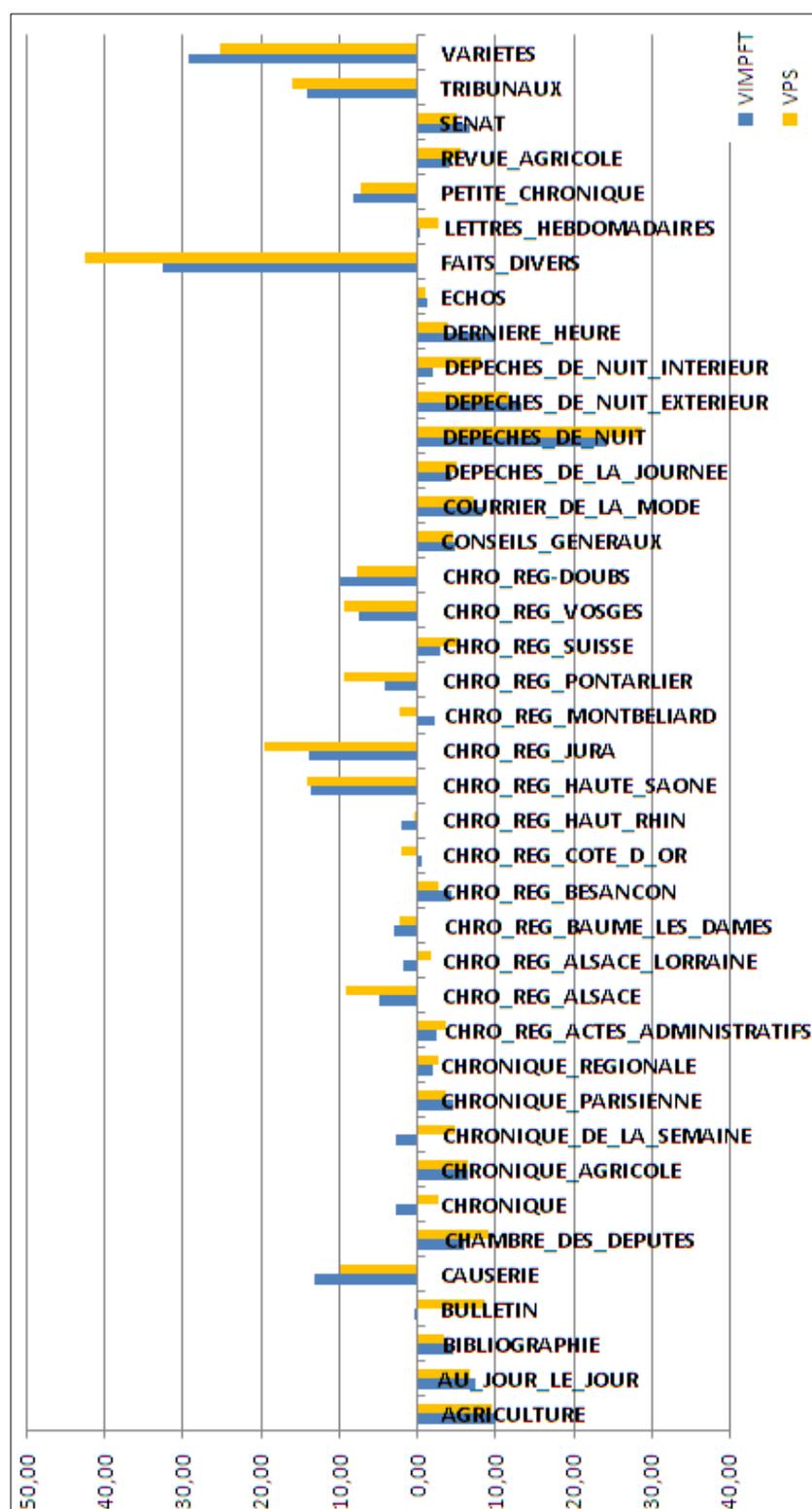


Figure V-35 : Histogramme croisé de la distribution de l'imparfait et du passé-simple (en écarts-réduits).

Le futur, le présent, et le participe passé, que nous envisageons comme l'indice du passé composé¹⁷, connaissent leurs plus forts sur-emplois dans les rubriques présentant des déficits d'emploi de l'imparfait et du passé-simple. Le groupe des dépêches, caractérisé par un profil distributionnel très déficitaire de l'imparfait et du passé-simple (cf. figure V-35), témoigne ainsi de très forts écarts-réduits positifs du présent (*Dépêches de nuit* : 26,98 ; *Dépêches de nuit-Extérieur* : 11,65)¹⁸, du passé-composé (*Dépêches de Nuit* : 26,54, *Dernière heure* : 18,65)¹⁹ ainsi que des temps des mondes projetés, à savoir le futur (*Dépêches de nuit-Intérieur* : 6,78) et le conditionnel (*Dépêches de nuit* : 11,32). Le profil des dépêches suggère donc une représentation discursive ancrée aux paramètres de la situation d'énonciation (Adam, 2005a : 197). À l'inverse, les profils des rubriques *Bulletin* (conditionnel : 11,06, futur : 5), *Courrier de la mode* (futur : 9,87), *Agriculture*, *Revue Agricole*, *Chronique Agricole*, et *Lettres hebdomadaires*, qui recourent à ces mêmes tiroirs verbaux, suggèrent que celles-ci sont régies par une « énonciation de discours » (Adam, 2005a : 198).

¹⁷ *Cordial* ne dispose en effet d'aucune catégorie d'étiquetage du passé-composé.

¹⁸ Seule la sous-rubrique *Dépêches de Nuit-Intérieur* s'individualise de cette tendance d'emploi du présent en présentant un profil distributionnel négatif (-4,97).

¹⁹ Encore une fois, la sous-rubrique *Dépêches de Nuit-Intérieur* se distingue en présentant un sous-emploi du passé-composé (-3,26).

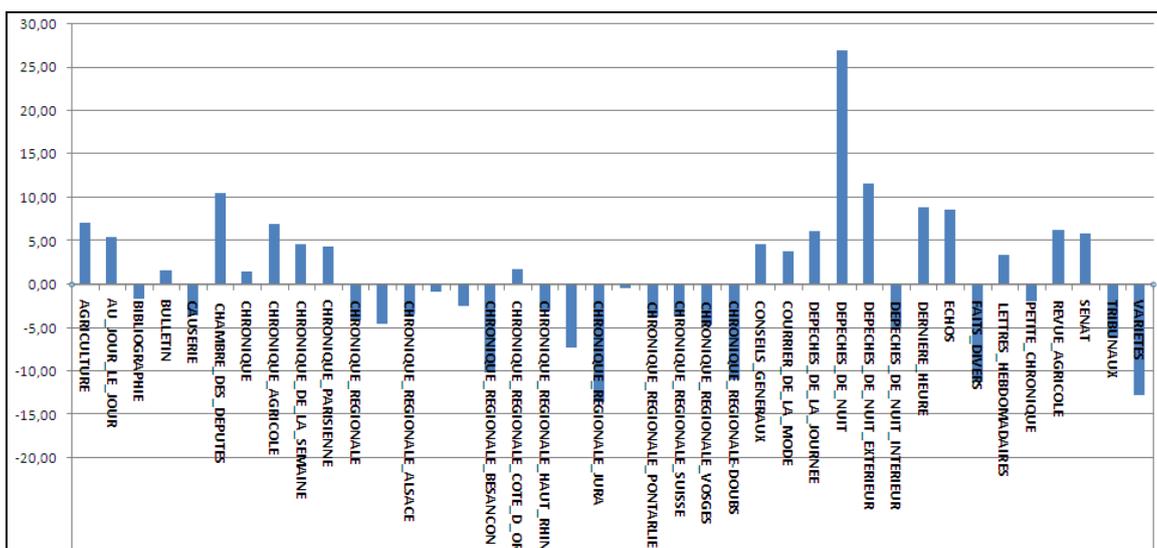


Figure V-36 : Distribution du présent (en écarts-réduits).

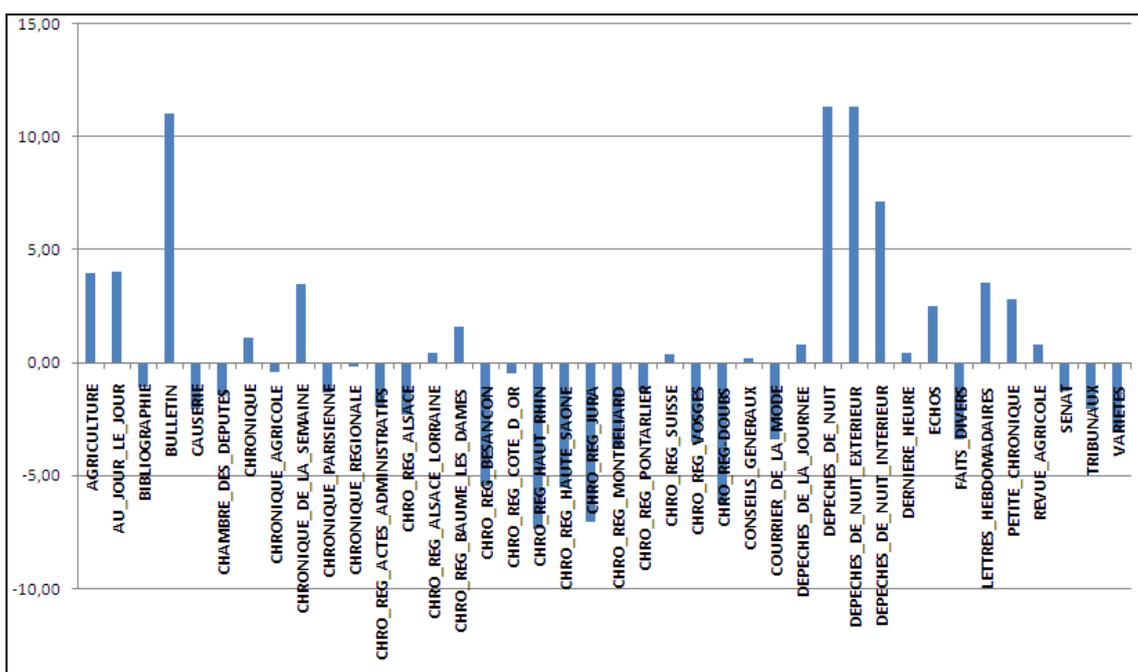


Figure V-37 : Histogramme de la distribution du conditionnel (en écarts-réduits).

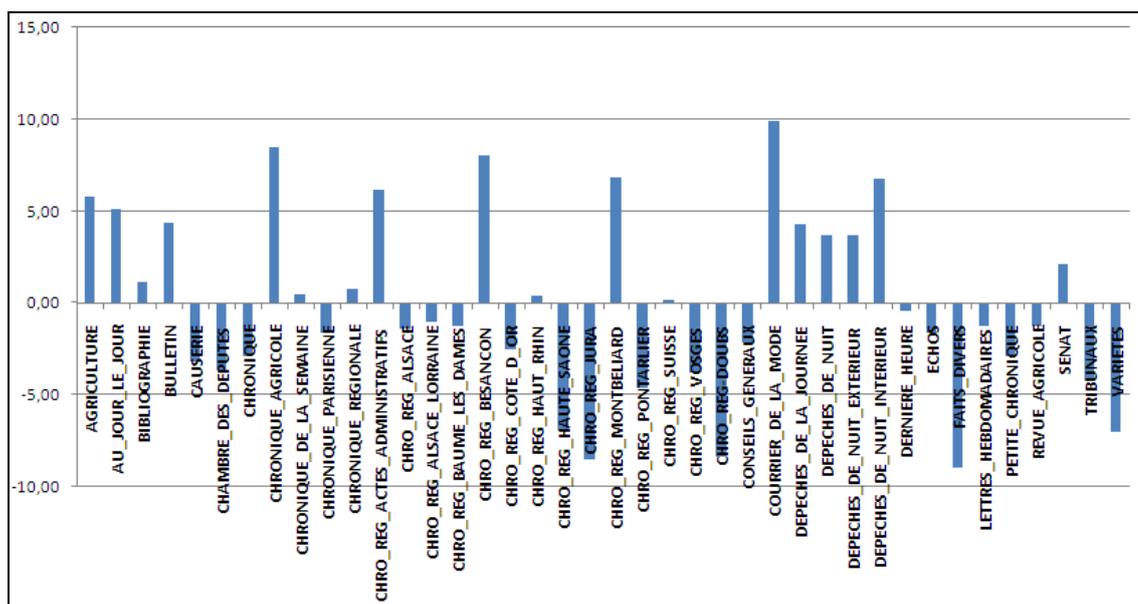


Figure V-38 : Histogramme de la distribution du futur (en écarts-réduits).

Nous nous arrêterons sur le profil distributionnel du passé-composé, forme sous laquelle E. Benveniste s'attachait à distinguer deux valeurs : un accompli du présent et un aoriste de discours (1966 : 249). Aussi l'emploi de ce tiroir verbal se concentre-t-il à la fois dans les rubriques dont les temps verbaux spécifiques suggèrent le plus fortement une énonciation de discours ainsi que dans les rubriques les plus caractérisées par une dimension historico-narrative, comme la rubrique *Faits Divers*.

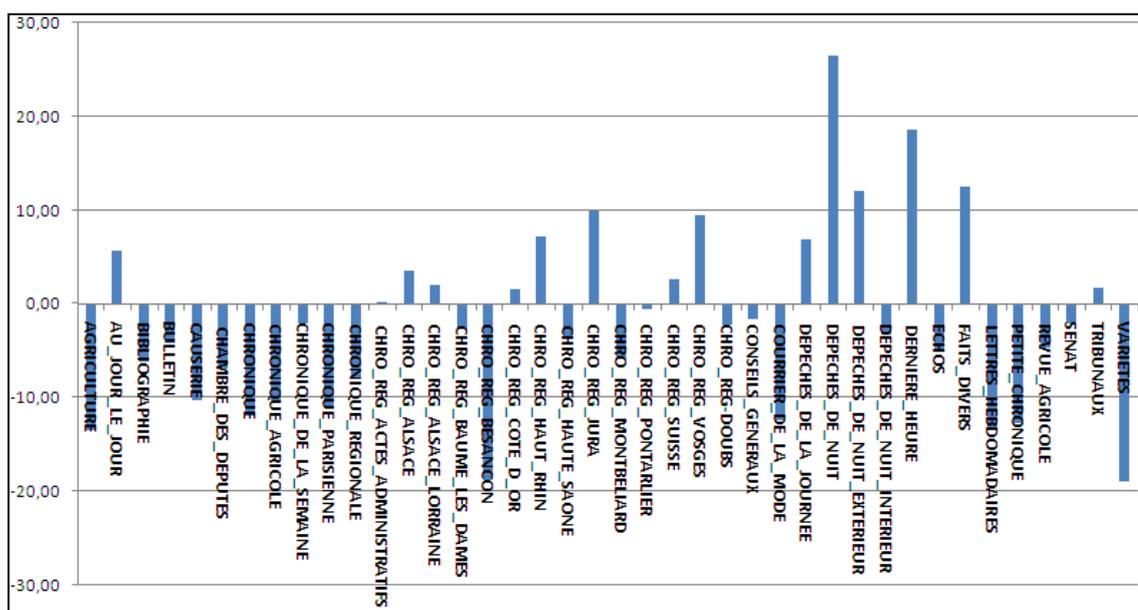


Figure V-39 : Histogramme de la distribution du participe passé, indice du passé-composé (en écarts-réduits).

Le genre

De façon assez surprenante, le genre se révèle un facteur discriminant de l'analyse : le genre féminin, tout particulièrement, participe à la structuration des trois premiers facteurs dans une part importante, tandis que le masculin participe à la structuration des deux derniers facteurs.

L'examen du profil distributionnel de ces deux variables met au jour des informations qui nous semblent devoir être mises en parallèle avec l'utilisation massive des noms communs dans le groupe des dépêches. Une rubrique comme *Dépêches de Nuit* est ainsi fortement caractérisée par le genre du masculin et du féminin, et ce, dans des mesures sensiblement identiques : elle réalise un écart-réduit de 37,04 pour le genre féminin et de 33,30 pour le genre masculin. Néanmoins, on observe sur l'histogramme ci-dessous que certaines rubriques privilégient très nettement l'un des deux genres. Le féminin domine ainsi dans les rubriques *Dépêches de nuit-Extérieur* (12,63), *Revue Agricole et Sénat*, *Au jour le jour*, *Courrier de la Mode*, *Chronique Agricole* et *Chronique parisienne*. Que signifie le sur-emploi du féminin dans ces cases ?

S'il est difficile en langue française d'attribuer une signification à l'opposition masculin/féminin, E. Brunet rappelle que l'abstraction « concerne davantage le féminin que le masculin » (1988b : 193). En d'autres termes, dans les corpus qu'il observe, la proportion féminin/masculin pour les substantifs s'inverse selon que ces substantifs ont une valeur plutôt abstraite ou plutôt concrète. Partant, l'hypothèse selon laquelle le sur-emploi serait l'indice d'un régime d'abstraction élevé des textes de ces rubriques pourrait être avancée. Néanmoins, la fréquentation de notre corpus nous conduit à émettre certaines réserves quant à la validité de cette hypothèse : les *dépêches de nuit Extérieur*, rubrique axée sur le factuel ne nous semble pas portée sur l'abstraction, pas plus que la rubrique du *Courrier de la mode*. L'hypothèse reste néanmoins *a priori* pertinente pour une rubrique telle que *Sénat*, où sont rapportés les débats parlementaires, et à approfondir pour l'ensemble.

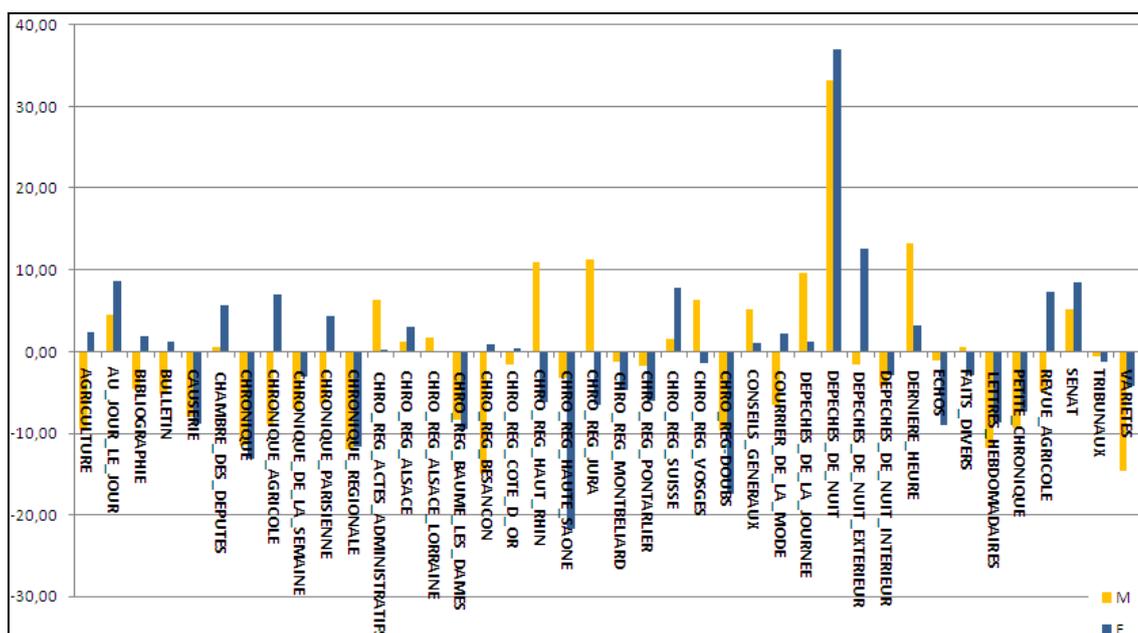


Figure V-40 : Histogramme de la distribution du genre (en écarts-réduits).

5.4.3. Synthèse sur les facteurs discriminants

L'examen des variables les plus discriminantes parmi les 52 traits convoqués pour caractériser les rubriques du *Petit Comtois* ne manque pas de rappeler les critères sous-jacents aux typologies des genres du discours journalistique, tout en les complétant par des traits plus originaux et moins attendus.

- Un faisceau de variables fortement discriminantes converge pour opposer les genres à énonciation impliquée aux genres marqués par un effacement énonciatif :
 - Il en va ainsi, tout d'abord, de l'impact hautement discriminant des pronoms personnels de première personne du singulier et du pluriel, ainsi que de la deuxième personne du pluriel, qui contribuent essentiellement à opposer le groupe des dépêches à l'ensemble des autres rubriques. Si cette indication était somme toute fort attendue, l'analyse de ces pronoms personnels met au jour un clivage plus subtil à l'intérieur du groupe des chroniques régionales, entre les sous-rubriques relatives à la sphère directement proche du siège du *Petit Comtois* et les sous-rubriques relevant d'une zone plus éloignée (*Suisse, Haut-Rhin, Alsace, etc.*).
 - Autres variables fortement discriminantes, les temps verbaux suggèrent pour leur part la pertinence de l'opposition entre les systèmes d'énonciation du discours et de l'histoire (Benveniste, 1966).
 - Selon toute attente, les indices de la subjectivité que sont les adverbes et les adjectifs sont également discriminants.
- Un deuxième ensemble de variables discriminantes nous semble pouvoir être constitué en cela qu'elles concernent le degré de complexité et de longueur de la phrase :

- Parmi les critères fortement structurants, la longueur des phrases marque un clivage évoquant l'opposition communément acceptée entre les genres « courts » tels que la brève et le filet d'une part, et les genres longs, correspondant essentiellement aux genres « de commentaire. »
- Loin de constituer un indice de la subjectivité du locuteur, la ponctuation forte individualise le groupe des dépêches, privilégiant les phrases courtes et simples. Au contraire, la ponctuation faible se révèle caractéristique des chroniques et autres articles à énonciation impliquée (*Causerie, Variétés*, etc.).
- La corrélation entre le profil des pronoms personnels et celui des liens de coordination et de subordination, suggère pour sa part un gommage des liens de causalité, ce qui tend à corroborer le rôle structurant de la tension entre genres *objectivisés* et genres *subjectivisés*.
- Le groupe nominal est également discriminant :
 - les noms communs et les noms propres opposent ainsi principalement le groupe des dépêches aux autres rubriques.
 - C'est principalement au niveau des variables relatives au groupe nominal que se jouent des critères discriminants plus surprenants que ceux précédemment évoqués. On remarquera ainsi que les déterminants définis ont un profil extrêmement dynamique, qui contribue à individualiser le groupe des dépêches. Plutôt que d'interpréter cet indice comme le témoin d'un contenu tourné vers l'abstraction (Brunet, 1988b), nous privilégions l'hypothèse que ce sur-emploi est caractéristique d'une écriture journalistique présupposant la connaissance d'éléments co(n)textuels, convoquant l'article défini pour sa valeur anaphorique large, sans écarter complètement un contenu tourné vers l'abstraction, d'autant moins que le genre grammatical constitue une variable

fortement discriminante dont il reste à définir la signification. Faut-il voir dans les sur-emplois du genre féminin un indice d'un niveau élevé d'abstraction, comme le suggère E. Brunet (1988) ?

5.4.4. Caractérisation morphosyntaxique des rubriques

L'analyse de la distribution des variables morphosyntaxiques met ainsi en relief cinq pôles de rubriques à l'aune de leurs proximités distributionnelles, dont nous décrivons les caractéristiques, tout en soulignant certains garde-fous méthodologiques. En effet, issus d'une analyse statistique où la rubrique a été convoquée comme un seul tenant, les profils morphosyntaxiques dont nous allons faire état correspondent à des tendances lourdes, homogénéisant les variations qui peuvent exister entre les différents textes d'une rubrique d'une part, et entre les parties d'un même texte, d'autre part.

5.4.4.1. Les dépêches

Les différentes rubriques du *Petit Comtois* signalées comme relevant du même matériau, les « dépêches », partagent des profils morphosyntaxiques fortement semblables. Ce groupe de rubriques se distingue tout d'abord par un excédent du groupe nominal, comme en témoignent les sur-emplois observés des déterminants définis, des noms propres, des noms communs qui tendent à les opposer aux autres groupes de rubriques. À l'exception des *Dépêches de Nuit Intérieur* et *Dépêches de Nuit Extérieur*, les adjectifs numéraux sont également sur-employés dans ce groupe. L'ensemble des dépêches se distinguent ensuite par la longueur et la construction de leurs phrases. Celles-ci sont courtes et se caractérisent par un emploi déficitaire des relations de subordination et de coordination. À ce qu'on peut interpréter comme la trace linguistique d'un gommage des liens de causalité, de conséquence et d'opposition, fait écho un effacement des marques de personne du locuteur. Ces indications tendent à suggérer une écriture condensée de l'information, axée sur le factuel, mettant l'accent sur le « qui a fait quoi, où et quand » :

Ch.V (1) « La révolution », Dépêches de nuit, Extérieur, Brésil, (1894-02-06)

Une dépêche de Rio-de-Janeiro du 3 février annonce que les troupes gouvernementales ont été battues à Curitiba ; deux cents hommes ont été tués. L'amiral de Mollo a expédié 1,000 hommes à Laps, au sud-ouest de Curitiba.

Ch.V (2) « La grève de Marseille », Dépêches de nuit, (1901-03-05)

Marseille, 4 mars.

La situation est calme ce matin sur les quais.

Des groupes de grévistes stationnent à l'entrée de divers chantiers où le chômage est complet.

Les hangars des messageries maritimes sont occupés par un escadron de hussards et des gendarmes à cheval.

Les ouvriers charbonniers, au nombre de 400, ont décidé de se mettre en grève. Cette décision aggrave la situation.

Privilégiant le présent et le passé-composé marquant l'aspect accompli, les dépêches sur-emploient aussi le futur, laissant percevoir qu'il n'est pas seulement question dans ces articles des faits passés mais aussi des faits à venir :

Ch.V (3) « Voyages ministériels », Dépêches de nuit, (1884-01-02)

M. Raynal, ministre des travaux publics et M. Méline, ministre de l'agriculture, se rendront cette semaine, à Bordeaux.

M. Raynal assistera à l'inauguration du chemin de fer des Landes, qui aura lieu le 6 janvier et présidera le banquet à Andongé, dernière station de la nouvelle ligne.

M. Méline visitera les grandes exploitations agricoles fondées dans le canton de La Teste par MM. Feray et De Cauville.

Un banquet sera offert à Arcachon.

De Bordeaux, M. Méline ira à Montpellier.

M. Waldeck-Rousseau s'absentera aussi pendant quelques jours.

Le conditionnel est pointé comme une spécificité de ce groupe de rubriques. Un retour au texte confirme qu'il s'agit le plus souvent d'un conditionnel dit « journalistique »²⁰, traduisant une prise de distance de la part de l'énonciateur vis-à-vis de propos dont la véracité reste à établir. Les exemples suivants illustrent l'utilisation de ce conditionnel journalistique où la rédaction du *Petit Comtois* veille à émettre des réserves face aux nombreuses dépêches qu'elle publie :

²⁰ De nombreuses études ont analysé cet emploi particulier du conditionnel dans les pratiques discursives journalistiques sous le terme de « conditionnel de non-prise en charge » (Abouda 1997) ou de « conditionnel de la rumeur » (Togebly 1982), (Grevisse & Goosse 2008).

Ch.V (4) « Agitation bonapartiste », Dépêches de nuit, (1883-08-09)

Le bruit court que l'ex-impératrice serait à Paris, où elle recevrait des visites dans un but politique. Elle voudrait ramener l'accord entre Victoriens et Jérômistes. Mais si l'entente est faite dans la famille, il paraît impossible de la faire dans le parti.

Ch.V (5) « Le moment serait mal choisi », Dépêches de nuit-Intérieur, (1900-03-23)

[...] On nous rapporte – et nous ne donnons l'information que sous les plus expresses réserves – qu'un établissement d'ordre religieux, réservé spécialement aux militaires, ferait en ce moment l'objet d'une enquête de la part de l'autorité militaire pour fait d'immoralité qui lui auraient été dénoncés. Le ministre de la guerre, n'a pas, jusqu'à présent, reçu de renseignements sur ces faits.

Si le profil morphosyntaxique que nous venons de préciser tend à caractériser l'ensemble des dépêches, une rubrique s'en distingue néanmoins très nettement. Il s'agit de *Dépêches de nuit-Intérieur*. L'examen des profils distributionnels des variables discriminantes dans cette case montre que, contrairement au profil général observé des dépêches, les articles de *Dépêches de nuit- Intérieur* sont constitués de phrases longues, caractérisées par un emploi important des coordinations et des subordinations. Cette rubrique privilégie également l'emploi des pronoms relatifs et des pronoms démonstratifs. Quant aux temps verbaux, le présent s'y avère en sous-emploi, au profit du conditionnel et du futur.

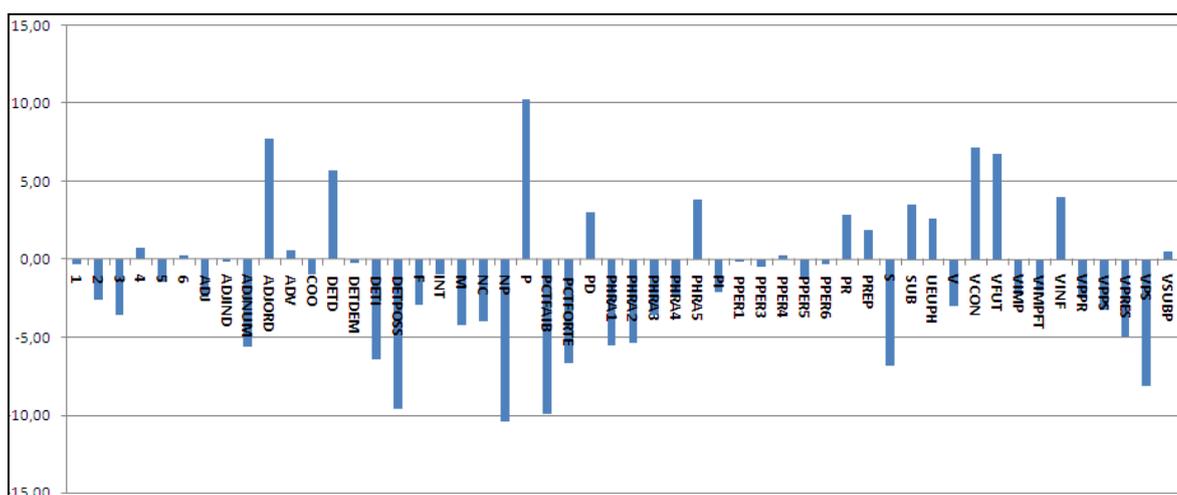


Figure V-41 : Profil distributionnel de « Dépêches de nuit- Intérieur » (en écarts-réduits).

Comment expliquer ce profil singulier ? Un retour aux articles de cette rubrique montre le statut particulier de cette case, qui héberge tout autant des filets ou des brèves que des articles empruntés à d'autres quotidiens, des correspondances relatives à des affaires intérieures ou encore des circulaires, comme en témoignent les exemples suivants :

Ch.V (6) « L'incident Viviani-Henri d'Orléans », Dépêches de nuit-Intérieur, (1898-02-27)

La *Fronde* publie la lettre suivante que M. Viviani vient d'adresser au prince Henri d'Orléans, en réponse à celle que nous avons publiée hier :

Monsieur,

Moi non plus je n'aime pas l'équivoque. Si vous avez entendu près de vous des gens insulter des officiers, et cela sans motif, sans provocation, pour le triste plaisir, le sentiment que vous a inspiré cette attitude est partagé par tous.

Mais je suis trop habitué aux polémiques, pour ne pas douter, sinon de votre bonne foi, au moins de la sûreté de votre appréciation.

Je sais ce que veulent dire d'ordinaire ceux qui se plaignent qu'on insulte l'armée. C'est l'insulter, à les entendre, que de réclamer sa subordination au pouvoir civil, de demander qu'on ne condamne pas pour un geste de pauvres soldats à mort, d'émettre de justes critiques sur le haut commandement.[...]

Ch.V (7) Dépêches de nuit-Intérieur, (1903-03-30)

M. le ministre de l'intérieur vient d'adresser aux préfets la circulaire suivante :

Les débats récents du Parlement ont fait ressortir les procédés odieux dont un instituteur avait été victime. Je suis informé que ce fait ne serait pas isolé : dans

plusieurs départements, des instituteurs se verraient l'objet de vexations, de tracasseries et même de violences ; des manifestations injurieuses seraient publiquement organisées contre eux, avec sinon l'approbation, du moins la complicité tacite des chefs des municipalités.

Il importe que l'autorité supérieure mette un terme à cet état de choses, au moment surtout où, l'application ferme des lois scolaires et de la loi sur les associations est l'occasion d'un redoublement d'hostilité de la part des adversaires de l'enseignement laïque. [...]

On remarquera les profils proches des rubriques *Chambre des députés*, *Sénat* et *Au jour le jour*, rubriques à part entière aux débuts de la parution du quotidien, progressivement intégrées comme sous-rubriques dans la case *Dépêches de nuit*.

La case *Chambre des députés* partage avec les dépêches un sur-emploi marqué des déterminants définis, des noms communs, de la ponctuation forte, des phrases courtes, du présent. Elle se distingue néanmoins par un sur-emploi significatif des subordinations (2,80), du subjonctif (2,73), des pronoms de première personne du singulier (5,43) et de deuxième personne du pluriel (11,55). Ces indices nous semblent à relier au genre auquel sont dédiées les productions hébergées dans cette case : le compte-rendu des séances parlementaires, qui se caractérise par la reproduction d'allocutions des députés, d'où un sur-emploi des pronoms personnels, comme l'illustrent les exemples suivants :

Ch.V (8) Chambre des députés, (1887-01-12)

[...]Après avoir déclaré ouverte la session ordinaire de 1887, M. Pierre Blanc a prononcé l'allocution suivante : « Messieurs les députés, En ouvrant cette session, mon souhait le plus ardent est qu'elle soit plus heureuse que sa devancière et qu'elle inaugure définitivement une politique de progrès, d'économie et de travail. (Très bien.) Je ne doute pas que les bonnes volontés qui siègent dans cette Assemblée ne s'associent à ce souhait et que, pour le réaliser, elles ne s'unissent sur un programme d'action commune [...]

Ch.V (9) Chambres des députés, Dépêches de nuit, (1894-02-14)

[...]On interrompt la discussion sur les blés pour que M. Bouge puisse adresser une question au ministre de l'intérieur.

M. Bouge dit qu'il s'est produit dimanche, autour de la tombe de Vaillant, des scènes scandaleuses. Des bandes anarchistes ont crié : Vive la Commune ! Vaillant, tu seras vengé !

L'orateur demande au gouvernement de ne plus tolérer des manifestations semblables. Allez-vous, dit-il, permettre plus longtemps le déploiement du drapeau rouge ? (Applaudissements au centre.) Le drapeau rouge est celui du vol et de l'émeute. (Applaudissements.) Je vous demande de poursuivre les agitateurs, dussiez-vous trouver parmi eux des députés. (Vifs applaudissements.)

Hébergeant, à l'instar de *Chambre des députés*, des articles dont la forme générique est celle du compte-rendu de séance, la rubrique *Sénat* se distingue de cette dernière par un déficit des marques des premières personnes du singulier et du pluriel. Affiliée au groupe des dépêches en raison de ses emplois intensifs des adjectifs numéraux, des déterminants définis, des noms communs, de la ponctuation forte, des phrases courtes, du futur et du présent, la rubrique *Sénat* ne présente de surcroît, de façon tout à fait étonnante, pas de sur-emploi significatif des noms propres. Ces indices nous semblent suggérer le recours à une forme d'écriture plus narrativisée que celle des comptes-rendus de la chambre des députés, dans laquelle le discours indirect primerait sur le discours direct et les actants seraient désignés par leur statut ou par une périphrase plus que par leur nom propre au fur et mesure de la progression du texte. C'est du moins l'hypothèse qu'il nous semble légitime de formuler à la lecture des exemples suivants :

Ch.V (10) Sénat, (1885-01-14)

[...] M. Carnot déclare la session ouverte et prononce une allocution.

L'orateur envoie, au nom du Sénat, un tribut de reconnaissance à nos soldats de l'extrême Orient, ces vaillants pionniers de la civilisation, qui vont ouvrir de nouveaux débouchés pour les produits du travail national, les marchés de l'Europe étant obstrués par la redoutable concurrence étrangère.

La politique coloniale n'est pas une politique de conquête, car elle doit poursuivre seulement l'établissement de relations pacifiques et emploie les armes à regret.

L'orateur, parlant de la crise industrielle et agricole, dit que la France n'a pas sujet de s'inquiéter plus que les autres nations. [...]

Ch.V (11) Sénat, (1886-01-31)

[...] M. Demole lit un article paru dans *l'Avant-Garde*, journal monarchiste. (Marques unanimes de réprobation à gauche).

L'orateur conclut de cette lecture que la provocation est partie du camp monarchiste. On n'a rien fait contre le journal conservateur, on ne fera rien contre le journal républicain. L'orateur termine en conviant à la modération tous les partis. M. de Ravignan réplique que le *Républicain Landais* était assuré de la complicité du gouvernement. (Vives protestations à gauche.)

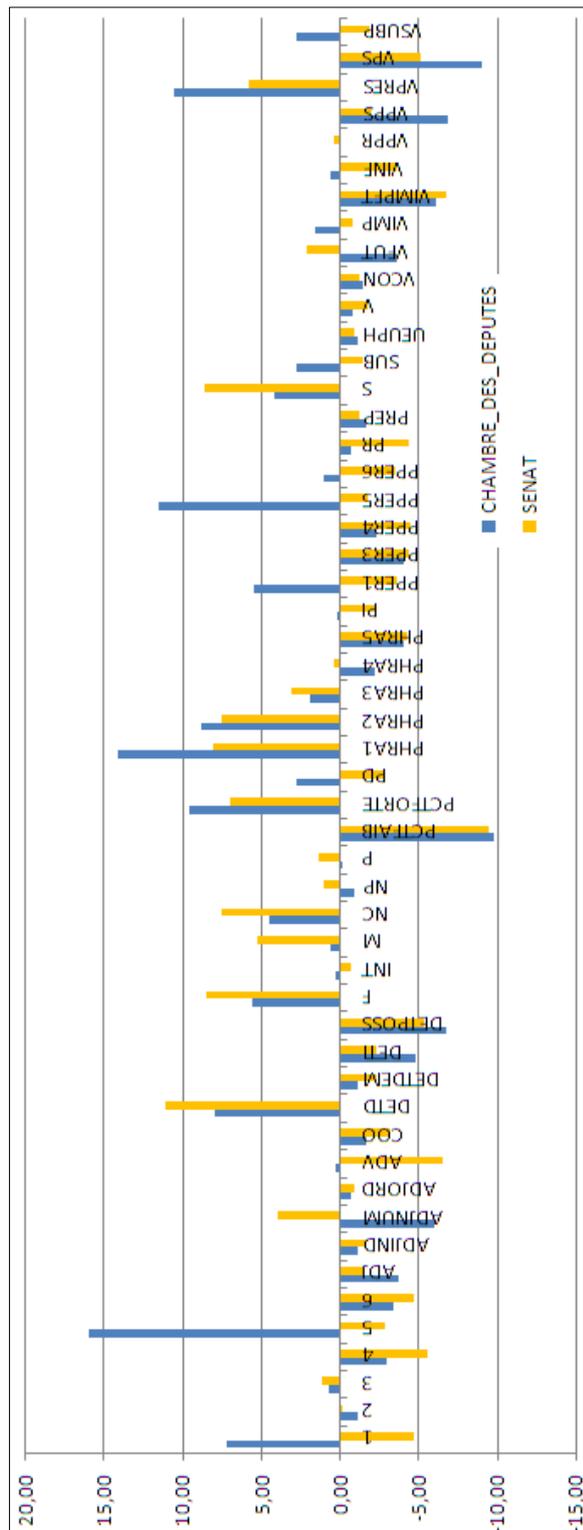


Figure V-42 : Profil distributionnel des rubriques « *Chambre des députés* » et « *Sénat* » (en écarts-réduits).

5.4.4.2. Les chroniques régionales

Formant un groupe unifié à l'aune de leur sur-emploi des phrases longues et des coordinations, les chroniques régionales se divisent en deux principaux sous-groupes, auxquels correspondent deux profils distributionnels distincts. On distinguera ainsi, d'une part, les chroniques régionales relevant de la zone très proche du lieu d'édition du quotidien, à savoir le département du Doubs et sa proche périphérie, et d'autre part, les chroniques régionales provenant de départements plus éloignés.

Les chroniques régionales *Doubs*, *Baume-les-Dames*, *Pontarlier*, *Besançon*, *Montbéliard* et *Haute-Saône* forment le noyau dur du premier sous-groupe. Ce sont les emplois intensifs de la deuxième personne du pluriel, des marques des pronoms démonstratifs, relatifs et interrogatifs, des déterminants possessifs, des adjectifs et des adverbes de la ponctuation faible et de l'impératif qui particularisent ce sous-groupe. À l'exception de *Chronique régionale Pontarlier*, toutes se caractérisent également par un sur-emploi du pronom de première personne du singulier. Le sur-emploi de la première personne du singulier et du pluriel est à relier aux conditions de production de la majorité des textes hébergés dans ces cases, qui se révèlent être des reproductions de courriers adressés au quotidien. Assumant leur responsabilité énonciative, les locuteurs rapportent les événements survenus dans leur ville en recourant tantôt au pronom « nous » renvoyant à la communauté géographique tantôt à un « je » renvoyant au statut singulier de l'énonciateur.

Ch.V (12) Chronique régionale, Haute-Saône, (1883-08-02)

Vesoul. — On nous écrit:

On se prépare aux élections du 12. La lutte dans quelques cantons paraît devoir être très vive. Les quelques succès remportés à Paris et ailleurs par des conservateurs donnent aux monarchistes surtout un aplomb qu'ils semblaient avoir perdu. Aussi en voyons-nous quelques-uns se risquer : M. de Rotalier, M. Desloyes, ancien député et ancien conseiller général, qui se remet sur les rangs après avoir été blackboulé aux précédentes élections.

Nous avons des conseillers à nommer dans quatorze Cantons. Nous donnons dans le tableau suivant les noms de candidats connus à ce jour. Au fur et à mesure

que d'autres se présenteront et qu'arriveront ceux de la dernière heure, je vous les ferai connaître. [...]

À ces pronoms personnels se combine le *nous* de la rédaction, qui introduit dans les sous-rubriques de la périphérie géographique les communications de ses correspondants et/ou lecteurs et qu'on retrouve logiquement dans les articles hébergés dans la sous-rubrique *Besançon*, rédigés intégralement par la rédaction :

Ch.V (13) Chronique régionale, Besançon, (1896-01-25)

Nous rencontrons dans les journaux de Paris une information qui nous paraît de nature à intéresser nos ouvriers de l'arsenal et ceux des équipements militaires. M. le ministre de la guerre aurait, dit-on, le projet d'instituer une commission chargée de régler d'une façon uniforme, en y apportant les améliorations nécessaires, les conditions du travail des ouvriers civils employés dans les établissements du ministère de la guerre. [...] Nous n'avons pas besoin de dire que le député de Besançon, M. Beauquier, a toujours fait partie de ce groupe.

On remarquera que les temps verbaux ne participent pas à l'unification de cette sous-classe de rubriques. Le passé simple est néanmoins néanmoins sur-employé dans toutes ces chroniques régionales à l'exception de *Chronique régionale Besançon*. Quant à lui, l'imparfait est utilisé dans les chroniques régionales *Baume-les-Dames*, *Haute-Saône*, *Pontarlier*. Les chroniques régionales *Besançon* et *Montbéliard* se distinguent par leur sur-emploi significatif du futur.

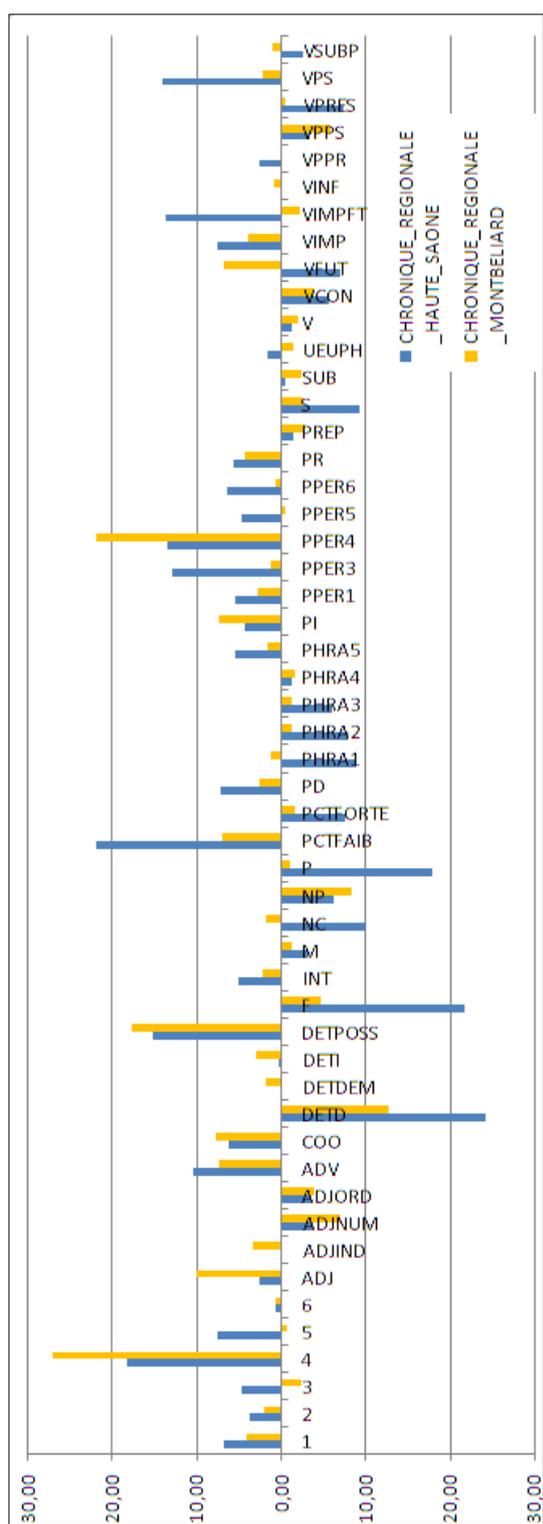


Figure V-43 : Profils distributionnels des rubriques « *Chronique régionale Haute-Saône* » et « *Chronique régionale Montbéliard* » (en écarts-réduits).

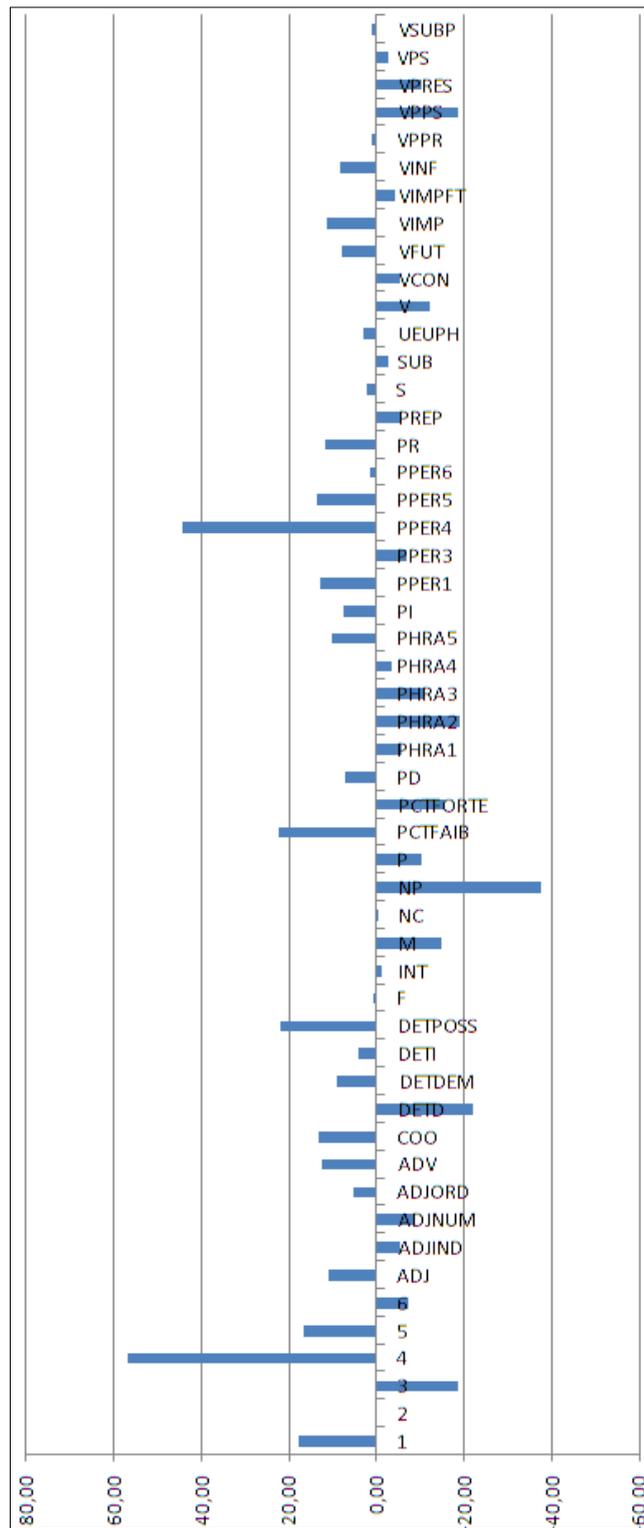


Figure V-44 : Profil distributionnel de « *Chronique régionale Besançon* » (en écarts-réduits).

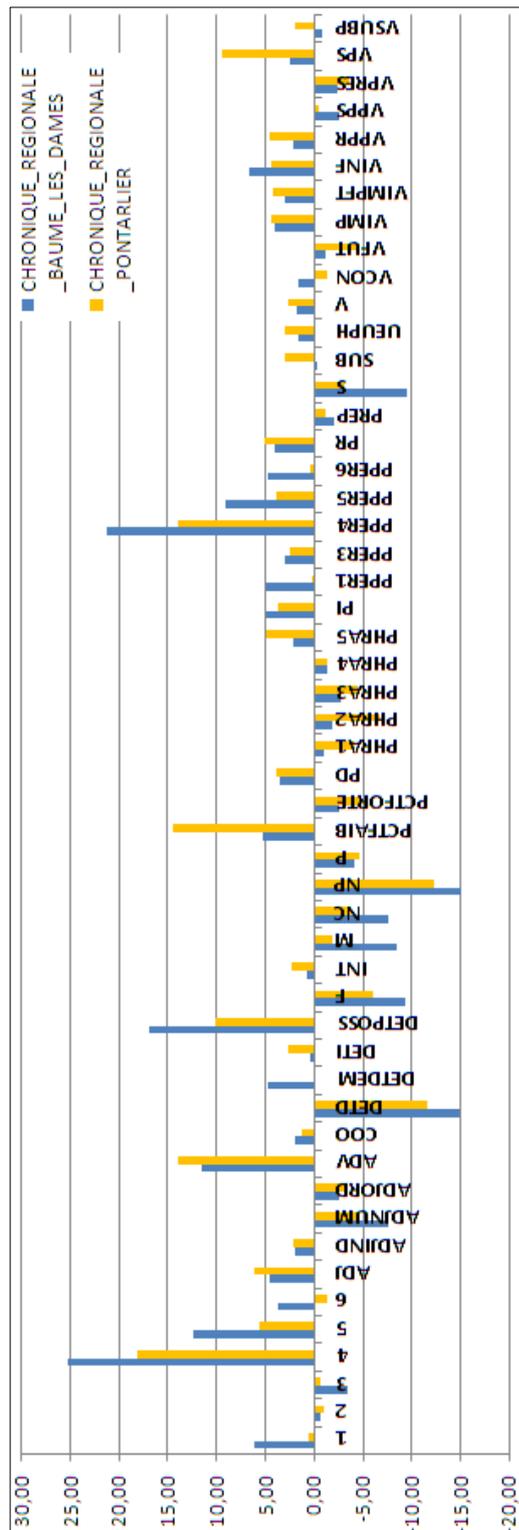


Figure V-45 : Profils distributionnels des rubriques « *Chronique régionale Pontarlier* » et « *Chronique régionale Baume-les-Dames* » (en écarts-réduits).

Les chroniques régionales *Alsace, Vosges, Alsace-Lorraine, Jura, Côte d'Or et Haut-Rhin* forment un sous-groupe au profil distinct des chroniques régionales proches du siège. Elles montrent un sous-emploi des pronoms personnels de première personne du singulier et du pluriel ainsi que du pronom de deuxième personne du pluriel et des marques de subjectivité (adjectifs, adverbes). Très peu fréquemment introduites par une formule du type « on nous écrit », les articles publiés sous ces cases se caractérisent également par un effacement des pronoms personnels de première personne du singulier et du pluriel. Il convient de remarquer ici que la subjectivité du locuteur se donne néanmoins à lire à travers d'autres indices, tels que les déictiques temporels (*dimanche dernier, cette nuit, etc.*) et des commentaires explicites.

Ch.V (14) Chronique régionale Jura, (1883-08-02)

La campagne électorale commence à se dessiner dans l'arrondissement. Dimanche dernier, une réunion publique, composée de nombreux électeurs et délégués du canton de Montmirey, a eu lieu dans la salle de la justice de paix de Montmirey-le-Château. M. Bourgeois, conseiller général du Jura, a été élu président et a ouvert la séance par une allocution applaudie, dans laquelle il a fait ressortir l'urgence qu'il y a à éliminer de nos assemblées politiques les derniers réactionnaires qui y sont encore cantonnés.

M. Henri Picot d'Aligny, conseiller général sortant, n'assistait pas à la réunion. Il était représenté par un ou deux de ces familiers.

L'assistance a désigné, à l'unanimité moins deux voix, M. Emile Maitre-robot, d'Offlanges, comme le candidat républicain que l'on opposera au représentant de la légitimité.

Bien que l'influence de celui-ci soit très grande dans le canton, les républicains comptent sur le succès.

Ch.V (15) Chronique régionale, Haut-Rhin (1886-01-04)

Frais, 2 janvier. Un triste accident vient d'avoir lieu cette nuit à Frais. Il est encore, malheureusement, dans les villages de notre territoire, l'usage d'annoncer le nouvel an par de nombreuses détonations de coups de feu. Le jeune Antoine Hollander, fils du fermier de M. Villien, était tranquillement assis à côté de son frère, auprès de la table de leur poêle, quand, vers onze heures du soir, vint le jeune Jules Mauvais, accompagné de son camarade Adolphe This, pour tirer le bon an aux membres de la famille Hollander. [...]

Le docteur Desprez, de Foussemagne, appelé de suite, constata que Hollander, atteint de trente-un gros plombs, avait reçu le coup mortel. En effet, ce matin, à dix heures, la malheureuse victime a succombé dans d'horribles souffrances. Toute la localité est consternée de cet accident. Puisse cette triste mort servir de

leçon aux jeunes gens du pays et leur apprendre à ne se servir d'armes que dans le cas de nécessité.

Sur ce point, on remarquera la différence de traitement entre l'exemple (14) et l'exemple (12) extrait de la *Chronique Régionale Haute-Saône* qui sont tous deux relatifs à la campagne électorale. Si les pronoms personnels des premières personnes du pluriel et du singulier abondent en (12), ils sont rares en (14).

C'est aussi le recours significativement positif au participe passé, à interpréter comme l'indice d'un fort emploi du passé-composé, qui singularise ce sous-groupe de chroniques régionales, à l'exception de *Chronique régionale Côte-d'Or*. Les chroniques régionales *Alsace/ Vosges/ Jura* utilisent conjointement le temps de l'imparfait dans de fortes proportions, tout en partageant également un sur-emploi les adjectifs numéraux. Le sur-emploi du passé-simple les rapproche du profil de la *Chronique régionale Côte-d'Or*.

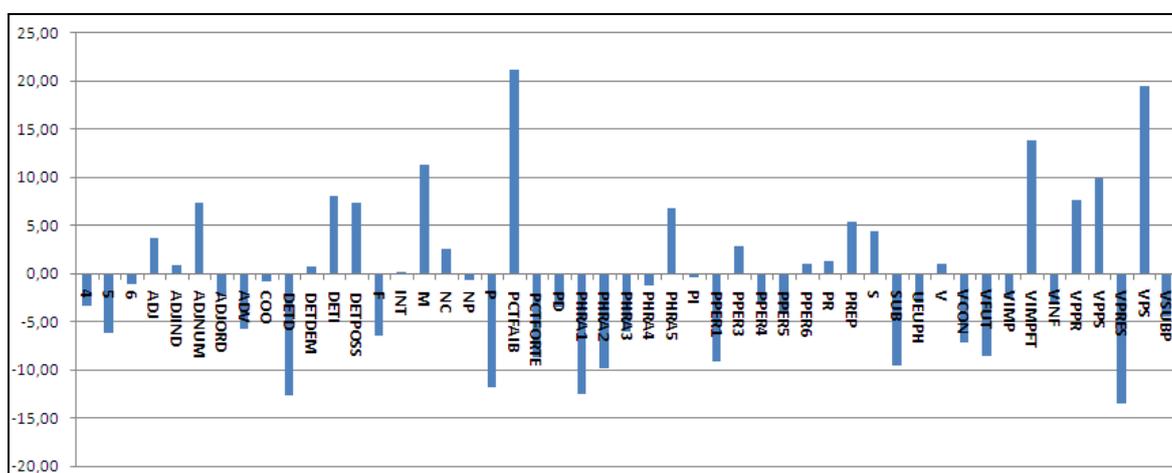


Figure V-46 : Profil distributionnel de « *Chronique régionale Jura* » (en écarts-réduits).

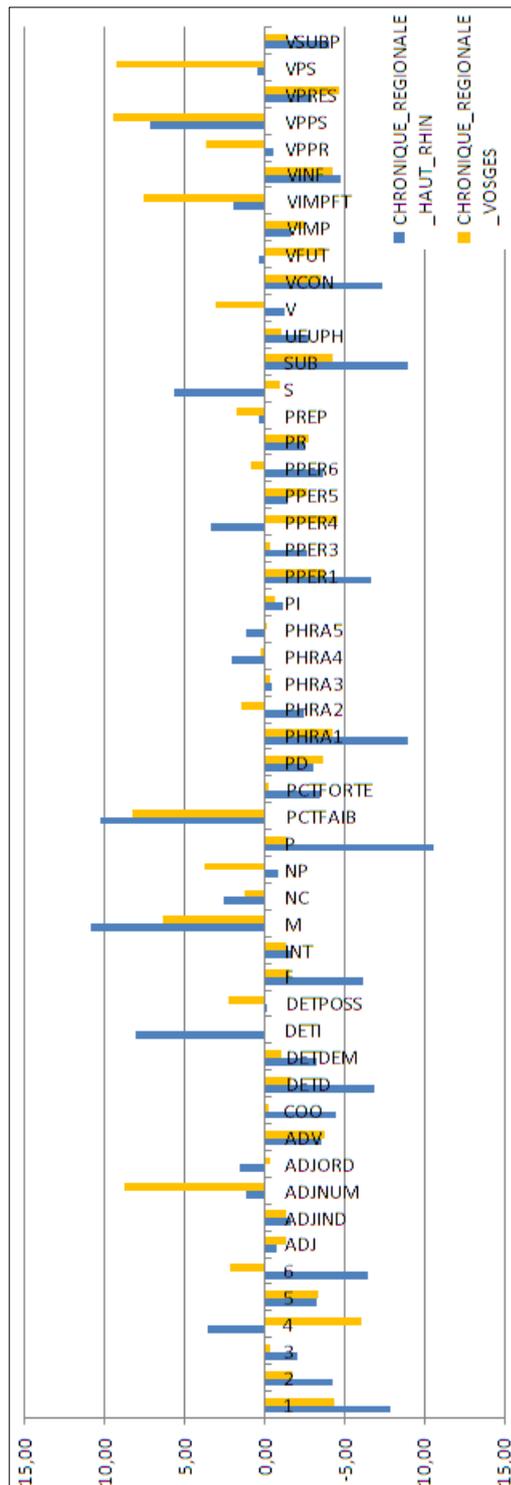


Figure V-47 : Profil distributionnel des rubriques « *Chronique régionale Vosges* » et « *Chronique régionale Haut-Rhin* » (en écarts-réduits)

5.4.4.3. Les rubriques *Petite chronique*, *Chronique*, *Chronique de la semaine*, *Lettres hebdomadaires*

Sur la figure V-21 synthétisant l'analyse de la ventilation des 52 traits morphosyntaxiques dans les rubriques, les rubriques *Petite chronique*, *Chronique*, *Chronique de la semaine*, *Lettres hebdomadaires* se concentrent dans le même hémisphère que les chroniques régionales et autres. Elles ont en commun un emploi intensif du pronom de première personne du singulier et de deuxième personne du pluriel, ainsi qu'une forte tendance à utiliser la liaison syntaxique de subordination, les adverbes, les pronoms démonstratifs ainsi que l'infinitif.

Chronique et Petite chronique

Ayant des intitulés basés sur la même référence générique, *Chronique et Petite Chronique* montrent des profils distributionnels très similaires. Outre les caractéristiques précédemment citées, les deux rubriques se caractérisent par un sur-emploi des adjectifs et des déterminants possessifs. Les liaisons sont assurées par la coordination, dont le sur-emploi est très significatif dans le cadre de phrases également marquées par les signes de ponctuation faible. L'impératif, les interjections sont deux traits communs à ces rubriques, qui recourent également dans des mesures significatives aux pronoms interrogatifs. Les exemples suivants permettent de vérifier la validité de la description morphosyntaxique des rubriques respectives dont ils sont extraits :

Ch.V (16) *Petite Chronique*, (1891-01-28)

Toujours la même histoire, toujours le petit jeune homme qui tue sa maîtresse et qui se laisse vivre après, comme dans la plainte de Geneviève de Brabant.[...] Fatal désintéressement! Mme Dida était bien morte, mais le revolver était vide! Que faire? Wladimiroff ignorait sans doute que, dans ces circonstances délicates, les héros de Ponson du Terrail n'hésitent pas à se passer leur pistolet à travers le corps, c'est pourquoi il se résigna à passer en cour d'assises. Pauvre jeune homme! si jeune et déjà condamné au supplice de vivre! Quel psychologue subtil nous dira les désespérances de cette âme qui n'a eu que le tort d'être trop généreuse? Oh! combien généreuse! Claude Larcher, et toi Bourget, son exécuteur testamentaire, répondez, qui nous les dira?

Et bientôt peut-être, du fond d'un noir cachot, montera la voix mélancolique du jeune Slave frappé par la justice, et cette voix, éveillant les échos de la morne prison, chantera dans la nuit, lugubrement, la si triste ballade de ceux qui souffrent et qui regrettent :

J'ai toujours eu trop de cœur, C'est c'qu'a fait mon malheur!

Ch.V (17) Chronique, (1896-01-01)

Pauvre Arton, voilà que l'Angleterre va nous le livrer pour qu'il déguste cette suprême amertume !

Où donc es-tu, ô popularité? Où donc ! Que j'y coure ! s'écriera un ambitieux. Et bien, non, qu'il ne coure point si ce n'est pour se réchauffer un brin. Brrr ! cette matinée est frisquette. Qu'il ne coure point. La popularité est chose morte. Boulanger en eut une comme on en compte bien peu dans toute l'histoire. Et je deviendrais six fois centenaire, à l'instar de Mathieu Salé, que je ne saurais oublier le spectacle inimaginable dont je fus témoin aux jours de délire populaire. [...]

La rubrique *Chronique*, qui emploie dans une mesure significative les éléments euphoniques (2,38), semble privilégier les phrases brèves (PHRA1, PHRA2) alors qu'aucun profil significatif n'est marqué par la *Petite Chronique* sur ce point.

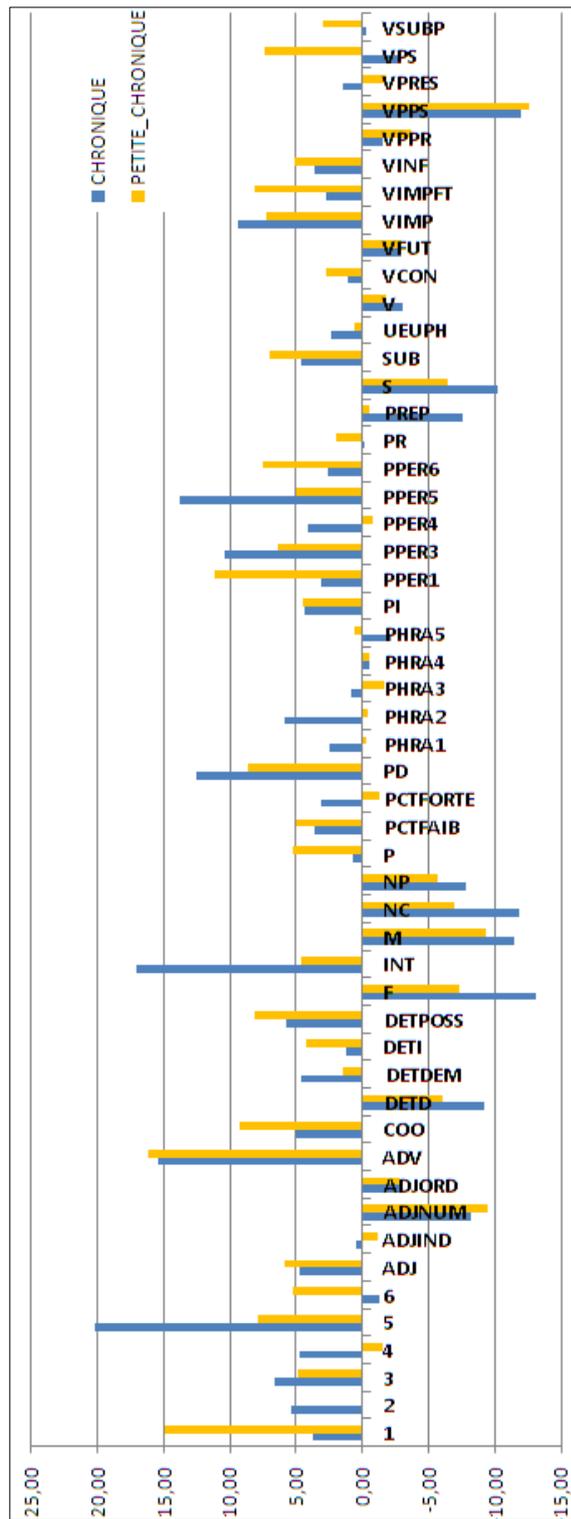


Figure V-48 : Profils distributionnels des rubriques « *Chronique* » et « *Petite Chronique* » (en écarts-réduits).

Chronique de la semaine et Lettres Hebdomadaires

Les rubriques *Chronique de la semaine* et *Lettres Hebdomadaires* se rejoignent dans le sur-emploi des phrases longues (PHRA5), du pronom de première personne du pluriel, de pronoms relatifs et de déterminants démonstratifs, comme on pourra le remarquer à la lecture des extraits suivants :

Ch.V (18) Chronique de la semaine, (1883-08-12)

Les deux empereurs d'Allemagne et d'Autriche ayant éprouvé le besoin de se donner la poignée de mains annuelle à Ischl et de mâcher patriotiquement un peu d'allemand, les journalistes se sont empressés de crier à l'alliance des deux empires. On a averti la Russie d'avoir à se méfier de cette alliance, et quelques malins ont poussé le flair politique jusqu'à prévoir la reconstitution du royaume de Pologne, qu'on nous prédit régulièrement tous les huit jours depuis la mort de Marie-Thérèse d'Autriche. [...] Il ne faut pas nous effrayer outre mesure des pronostics défavorables à la paix. Depuis que je suis au monde pour lire des journaux, je vois les Russes masser des troupes. C'est le métier des Russes de masser des troupes, comme celui des Auvergnats est de fondre du bois, M. Routier excepté.

Ch.V (19) Lettres hebdomadaires, (1902-03-09)

Sans blâmer aucunement, et même en approuvant expressément la motion relative aux survivants du siège de Belfort, je crois que M. Fachard n'était pas qualifié politiquement pour réclamer cette réparation politique, puisqu'il appartient au parti qui a volontairement commis l'injuste oubli dont il a entretenu la France à la tribune. Mais lorsque l'intérêt de son parti l'exige, l'honorable M. Fachard paraît se soucier assez peu de la logique de ses attitudes, ainsi que rétablissait un de nos collaborateurs cette semaine même, en nous le montrant conservateur acharné dans son patrimoine privé, des biens nationaux civilement acquis par ses ancêtres, au mépris des imprécations persistantes de l'Eglise! Je considère, au surplus, que votre campagne pour la prétendue défense patriotique de cette armée que nous n'attaquons pas, vient de se terminer d'une façon absolument lamentable pour vous. [...]

Organisés autour de phrases complexes, les articles hébergés par ces deux rubriques utilisent intensivement l'impératif et le conditionnel, le présent et le subjonctif présent, à relier à leur forte dimension subjective et argumentative. La différence des dispositifs génériques adoptés dans chacune des rubriques, l'une focalisée sur la chronique, l'autre se fondant sur la scénographie de l'épistolaire, se donne à lire à travers deux éléments : les marques de la deuxième personne du pluriel d'une part, qui abondent dans les *Lettres*

Hebdomadaires pour connaître un emploi moindre dans les *Chroniques de la Semaine* et, d'autre part, le temps de l'imparfait caractéristique des chroniques, indice d'une dimension historico-narrative absente des *Lettres Hebdomadaires*.

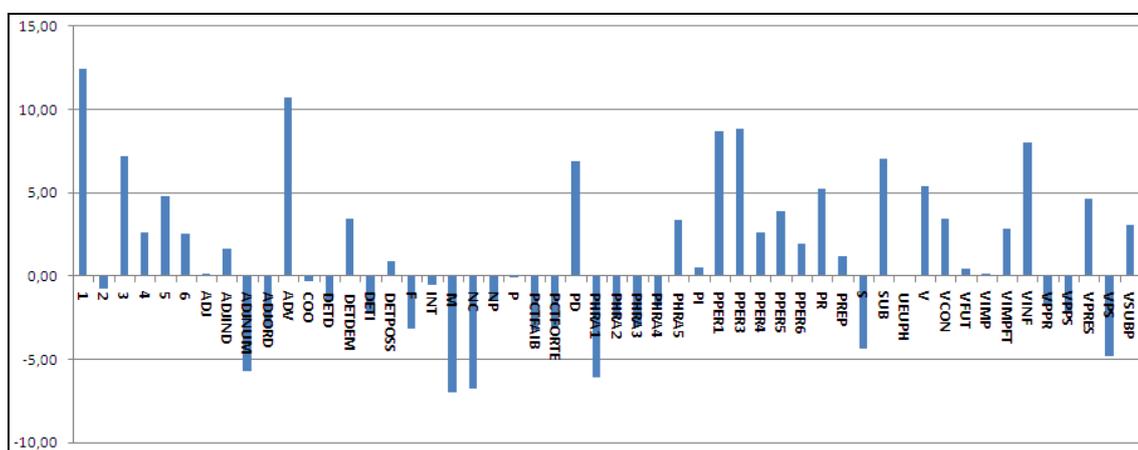


Figure V-49 : Profil distributionnel de « *Chronique de la semaine* » (en écarts-réduits).

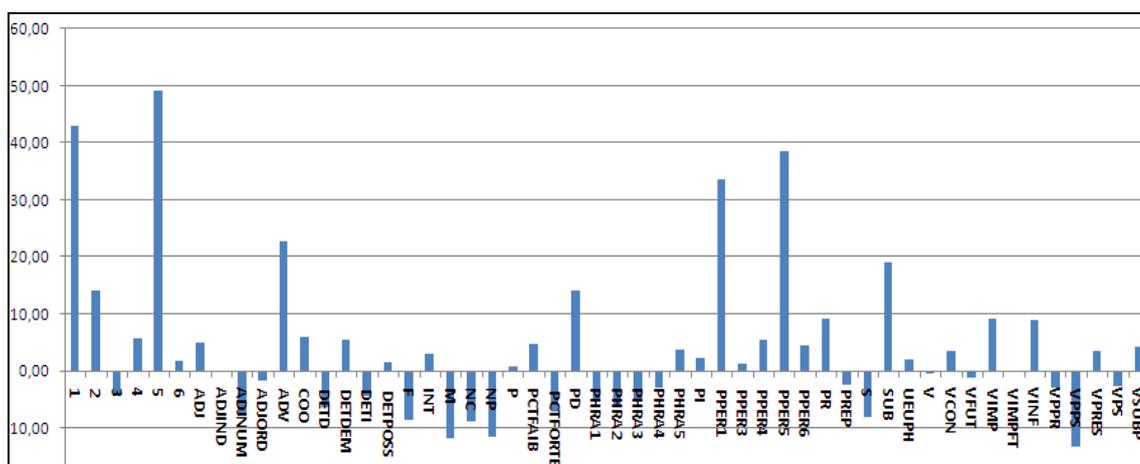


Figure V-50 : Profil distributionnel des « *Lettres hebdomadaires* » (en écarts-réduits).

5.4.4.4. Les rubriques *Faits Divers*, *Tribunaux*, *Variétés*, *Causerie*

Situées dans le quart est de la moitié droite du planisphère, les rubriques *Faits Divers*, *Tribunaux*, *Variétés*, *Causerie* ont en commun de présenter, outre un sur-emploi des temps du récit, celui des coordinations et les déterminants

possessifs.

Si l'on détaille les caractéristiques de chacune de ces rubriques, on notera que les *Faits Divers* sur-emploient les phrases de longueur moyenne (PHRA3). Cette case se caractérise par des emplois significatifs des coordinations et les déterminants possessifs, des adjectifs indéfinis et numéraux, des pronoms relatifs ainsi que des prépositions. La dimension historico-narrative des textes est indiquée par faisceau convergent de traits équivoques parmi lesquels le sur-emploi de l'imparfait, du passé simple et du participe passé, indice du passé composé, ainsi que de celui des marques de la troisième personne du singulier et du pluriel :

Ch.V (20) « Prisonniers étouffés », Faits Divers, (1884-01-01)

Dans la nuit de samedi à dimanche, on entendit, dans la prison d'Aarwagen, des cris terribles; par les soupiraux entrouverts, les prisonniers criaient au secours. Les voisins s'émurent de ce tapage et se rendirent enfin à la maison de détention. En faisant des recherches, on constata qu'une assez grande quantité de bois que le gardien avait mise dans le fourneau pour le faire sécher, s'était allumé; la clef du fourneau était fermée; il en résulta qu'une épaisse fumée, remplit bientôt la chambre, d'où elle se répandit dans toutes les cellules parties bouches à chaleur. Comme il faisait nuit, les prisonniers ne pouvaient se rendre compte des causes de leurs malaise; ils se sentaient peu à peu étouffer, ce qui augmentait leur terreur. Avant toute autre chose, les voisins cherchèrent les clefs. Malheureusement, le gardien était absent et sa femme ne savait pas où elles se trouvaient. Enfin, on finit par les trouver et les pauvres asphyxiés furent enfin délivrés. Il était grand temps, car en ouvrant les cellules, plusieurs prisonniers à moitié morts tombèrent dans les bras de leurs sauveurs. La fumée était si épaisse qu'une grande catastrophe serait arrivée si les portes avaient été ouvertes quelques minutes plus tard : les vingt détenus expiraient inévitablement.

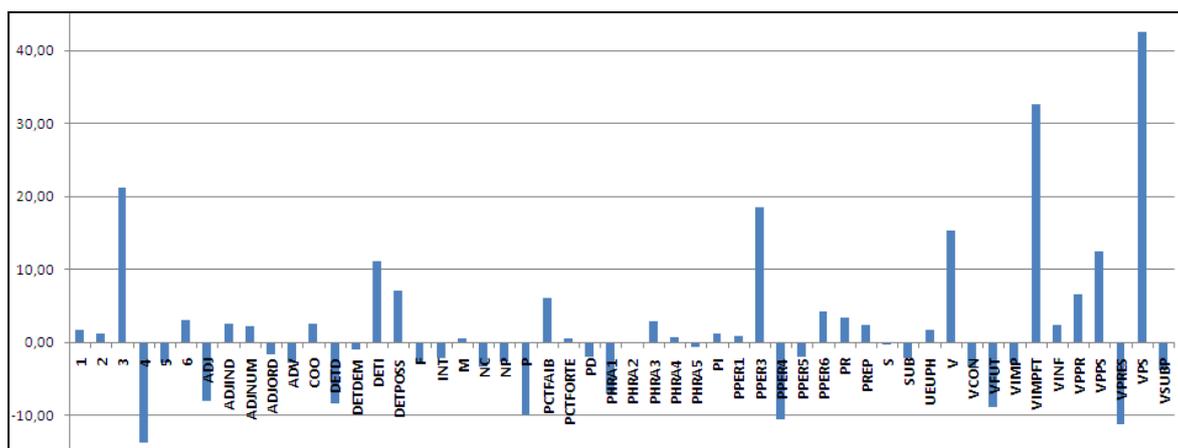


Figure V-51 : Profil distributionnel de « Faits Divers » (en écarts-réduits).

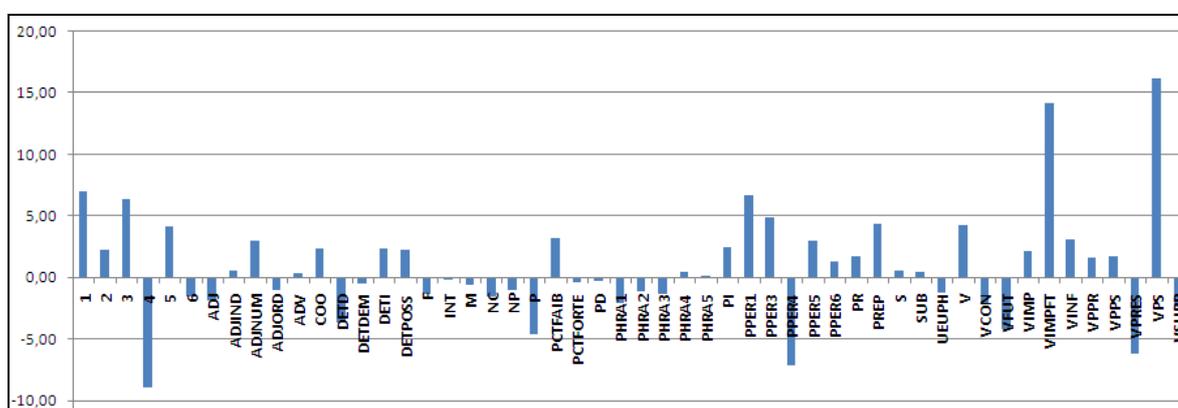


Figure V-52 : Profil distributionnel de « Tribunaux » (en écarts-réduits).

La rubrique *Tribunaux* partage, de façon attendue, un faisceau de traits avec les *Faits Divers* parmi lesquels les sur-emplois du pronom de troisième personne du singulier, des coordinations, des déterminants indéfinis et possessifs, de la ponctuation faible, des pronoms relatifs, des prépositions, de l'imparfait, du passé simple et de l'infinitif. La rubrique des *Tribunaux* se distingue néanmoins des *Faits Divers* en raison de son sur-emploi des pronoms et des marques de première personne du singulier (PPER1 : 6,69), ainsi que de la deuxième personne du pluriel, et de l'impératif. Ces indications pointent que les articles de cette rubrique, *a contrario* de ceux des *Faits Divers*, intègrent

fréquemment des parties de dialogue, résultant de la reproduction de minutes de procès et/ou de correspondances, comme en témoigne l'exemple suivant :

Ch.V (21) « Les amours d'un vicaire Tribunaux », Tribunaux, (1895-02-27)

On lit dans l'*Echo du Rhône* :

Moins tragique que celle de Boudes et de Bruneau, l'histoire de Thorre, le vicaire de Bergonne, dans le Puy-de-Dôme, vaut d'être contée. Ses amours avec Mme Dhumes, une de ses paroissiennes, viennent de défrayer la chronique judiciaire des journaux auvergnats, à l'occasion d'un procès en restitution, que plaidait contre ce don Juan ensoutané son ancienne maîtresse. [...]

Dans sa plaidoirie, le jeune avocat a lu toute une correspondance des plus suggestives, une série de lettres signées du vicaire, qui sont d'un style essentiellement profane. En voici quelques échantillons : [...]

JÉSUS-MARIE-JOSEPH

Ma bonne petite Marie, Ma chère enfant, je viens de lire votre billet et j'en suis navré ; navré de voir que vous méconnaissiez mes sentiments et de voir qu'à cause de cela vous souffrez terriblement. J'avais dit que je ne reviendrais pas à la charge, mais j'y suis obligé, car je ne puis vivre sans vous. Je viens donc, pauvre enfant, faire appel à votre cœur, vous dire que je vous aime et que je veux vous aimer chaque jour davantage. Oh ! de grâce, chère enfant, croyez-moi, pour votre bonheur et pour ma consolation. Puissiez-vous, en lisant ces lignes, être éclairée, sentir s'évanouir vos doutes et dire sincèrement : oui, il m'aime beaucoup je le crois bien ; j'oublie tout et nous allons être heureux. Répondez-moi alors pour me consoler. Oh! chérie, croyez votre père, croyez à son amour : oui, il vous aime ; ne vous torturez pas de l'idée contraire. Je souffre de vos douleurs et, de cette idée que vous avez que je ne vous aime plus..... [...]

Causeries et Variétés

Présentant une manifeste dimension historico-narrative perceptible au travers des temps qu'ils privilégient, les articles de *Variétés* et *Causeries* sur-emploient adjectifs, adverbes, et pronoms démonstratifs, interrogatifs et relatifs, ainsi que les pronoms et des marques de première personne du singulier et l'impératif, tous indices d'une forte prise en charge énonciative, que l'on illustre par la lecture d'un extrait de ces variétés :

Ch.V (22) Variété, (1883-01-07)

Par une belle matinée d'automne, je sonnai à cette porte, et je pénétrai dans un grand antichambre des plus simples ; au fond de la pièce, près d'une porte vitrée qui donne sur le jardin, on aperçoit sur un socle de buste de mon maître Perraud, l'ami de la maison.

Une domestique me fit monter au premier étage, vaste laboratoire. Sur les murs, l'ont voit des fleurs peintes par le fils de M. Pasteur. Encore un escalier à gravir, et j'arrivai dans une chambre sans autre ornement qu'un

rayonnage bondé de livres ; en entrant j'aperçus, derrière une grande table chargée de papiers, le grand savant coiffé d'une petite toque, comme l'a dessiné le peintre Billot. Il était en train de travailler ; il m'accueillit avec son beau sourire et une bonne poignée de main [...]

À la différence de *Variétés*, les phrases de la rubrique *Causerie* sont néanmoins plus courtes, puisque l'on notera le sur-emploi des phrases de moins de huit mots (5,98). La chaîne euphonique revêt enfin une valeur significative dans cette case (3,18).

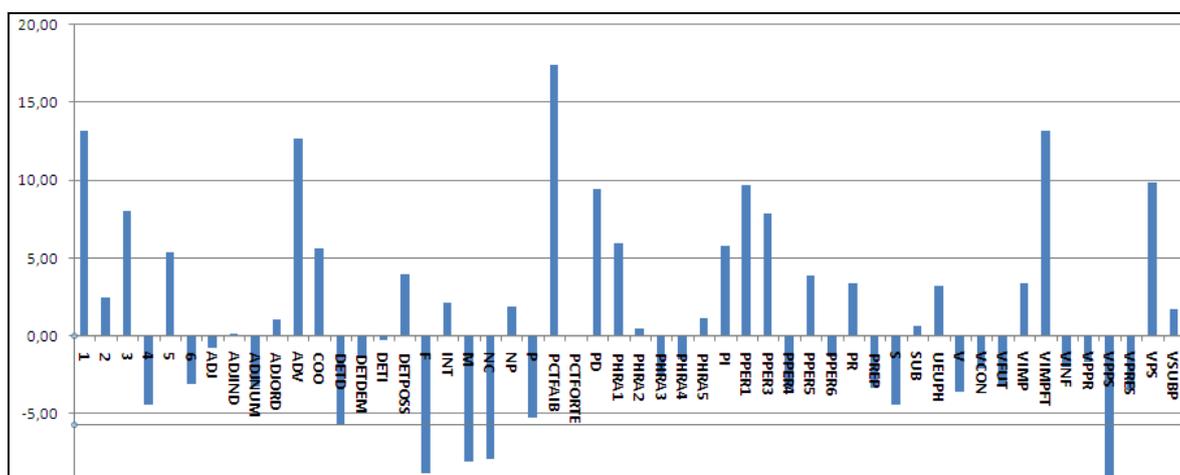


Figure V-53 : Profil distributionnel de la rubrique « Causerie » (en écarts-réduits).

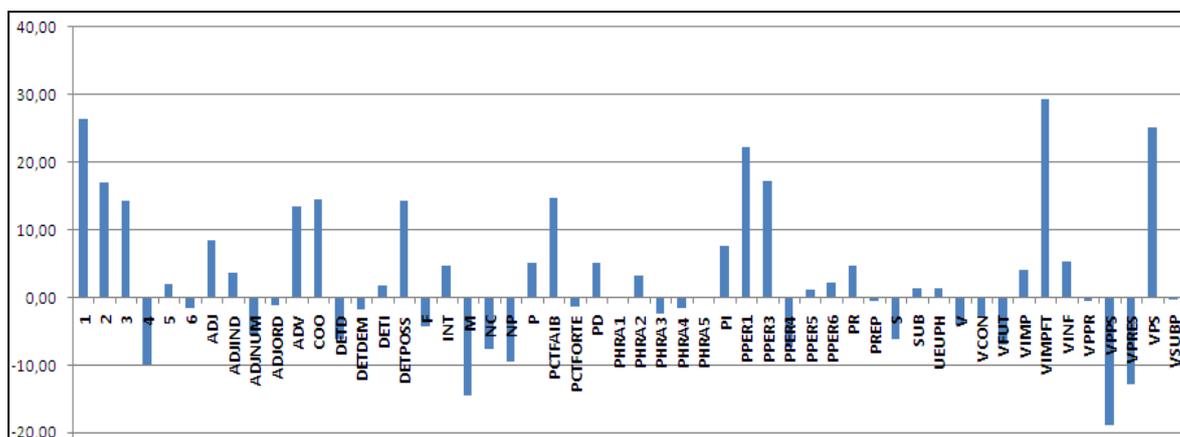


Figure V-54 : Profil distributionnel de la rubrique « Variétés » (en écarts-réduits).

5.4.4.5. *Courrier de la mode, Chronique parisienne, Revue agricole, Agriculture, Bibliographie, Bulletin*

Réparties dans le quart est de la moitié droite du planisphère, *Courrier de la mode, Chronique parisienne, Agriculture, Revue agricole, Chronique Agricole, Bibliographie, Bulletin* présentent des profils distributionnels relativement similaires concernant leur emploi intensif des adjectifs et des adverbes. Le regroupement de ces rubriques s'explique également par leurs utilisations extrêmement discriminantes du genre féminin et du nombre pluriel. On notera par ailleurs que cinq de ces rubriques (*Courrier de la mode, Agriculture, Revue agricole, Chronique agricole, Bulletin*) présentent un sur-emploi extrêmement discriminant de la chaîne euphonique, autrement dit de l'ajout de consonnes dites euphoniques pour éviter l'hiatus, dont on relèvera plusieurs exemples dans l'extrait ci-dessous :

Ch.V (23) Chronique agricole, (1893-07-01)

Les lecteurs du *Petit Comtois* nous permettront bien de consacrer notre article d'aujourd'hui au jardin. [...] Les cultures que l'on rencontre dans un jardin de ferme ne sont pas nombreuses. [...] L'asperge se reproduit par graines que l'on sème en pépinière afin d'obtenir de petites asperges enracinées que l'on connaît dans le commerce sous le nom de griffes.

***Courrier de la mode et Chronique parisienne* : deux rubriques pour une même auteure**

Les rubriques *Courrier de la mode* et la *Chronique parisienne* partagent des profils remarquablement similaires, dont les traits principaux sont le sur-emploi des marques de la première personne du singulier, de celui de phrases de taille moyenne marquées par une ponctuation faible, traduisant une forte propension à l'énumération de groupes nominaux juxtaposés comme on pourra l'observer à la lecture de l'extrait (24). Outre le présent de l'indicatif, sont sur-employés dans ces articles : le subjonctif, indice d'une tonalité prescriptive, le futur de l'indicatif et l'impératif, mode des conseils et des consignes. Les rubriques *Courrier de la mode* et la *Chronique parisienne* partagent des profils

remarquablement similaires, à une notable exception : *Courrier de la mode* est marquée par les première et deuxième personnes du singulier tandis que *Chronique parisienne* ne recourt dans des mesures significatives qu'à la première personne du pluriel (3,4). Au-delà de cette différence, liée à la scénographie épistolaire adoptée dans le *Courrier de la Mode*, il est très intéressant de remarquer que ces deux rubriques ont une seule et même auteure : Christine Carolet-Chonion.

Ch.V (24) Courrier de la mode, (1883-03-31)

Ah ! la mode ne nous met pas la bride sur le cou : plumes, fleurs, ailes, motifs de jais, de passementerie, choux en velours, en mousseline, en tulle, feuillages, plissés, rubans, nous pouvons nous laisser aller à notre préférence pour orner une paille de riz, une paille d'Italie, une forme en grosse paille mélangée de plusieurs tons ou formant des dessins, ou bien une paille ajourée, mélangée de froncés de mousseline, de tulle, de guipure tendre alternant avec des biais de taffetas ou de velours.[...]

Employons les fleurs surtout pour faire des toques et pour orner les capelines de toutes formes. Divisons les capelines en capelines avec fond et capelines plates. Sur ces dernières, on dispose des guirlandes entremêlées de feuillages et de rubans, ou de feuillages et de dentelle, des ailes posées à plat et mêlées de ruban ou de draperies de taffetas; enfin ce qui plus habillé que tout, une amazone de ton uni ou se dégradant du foncé au clair et retenue sur le fond de la capeline par une agrafe, motif de bijouterie, pouvant être un véritable bijou.

Chronique agricole, Agriculture et Revue agricole

Tout comme les rubriques *Chronique de la mode* et *Chronique parisienne*, les rubriques *Chronique agricole, Agriculture et Revue agricole* témoignent d'un intensif emploi des adjectifs et des adverbes, ainsi que du temps du présent de l'indicatif, de la chaîne euphonique, du genre féminin et du pluriel. La coordination est une caractéristique commune à ces rubriques, bien que non partagée par la *Revue Agricole*.

Si elle est marquée par un emploi significativement positif des pronoms de troisième personne du singulier et du pluriel, la rubrique *Agriculture* tend à s'individualiser par une prise en charge énonciative explicite. Contrairement à la rubrique *Revue agricole* et de façon plus nette que *Chronique Agricole*, la rubrique *Agriculture* est en effet écrite à la première personne du singulier

(PPER1 : 8,16) et recourt au pronom de première personne du pluriel ; elle témoigne enfin d'un usage prononcé de l'impératif.

Ch.V (25) « Curages et assainissements », Agriculture (1894-02-07)

[...]Comme je l'ai annoncé précédemment, avant d'aborder la question des irrigations, je voudrais dire quelques mots des syndicats agricoles que je recommande avec instance pour l'exécution des travaux de drainage et de curage ; mais cette question des syndicats ne peut être abordée sans entraîner à quelques réflexions sur la crise agricole proprement dits et c'est là une affaire bien délicate. [...]

À ces indices d'une forte subjectivité, s'ajoutent les traces de phrases relativement complexes, tels que le suggèrent les sur-emplois des coordinations, des subordinations et des pronoms.

Témoignant également d'une prise en charge énonciative forte à travers l'emploi significativement positif du pronom de première personne du pluriel, la *Chronique Agricole* se distingue sur le plan morphosyntaxique de la rubrique *Agriculture* par son utilisation massive d'adjectifs indéfinis et ordinaux. Il faut y voir l'indice d'une case tournée vers une approche plus technique de l'agriculture, comme le confirme la lecture de l'extrait suivant :

Ch.V (26) « Culture de la betterave fourragère », Chronique agricole, (1897-02-20)

[...] La betterave supporte facilement de fortes fumures, 40 à 50,000 kilogr. à l'hectare ne sont pas de trop. On peut employer cet engrais lors du labour d'hiver ou attendre celui de printemps ; l'essentiel est de ne pas laisser séjourner trop longtemps les fumures sur le sol. Lorsque le temps le permettra en mars, on labourera de nouveau à omgo si on le peut et on pourra procéder à un billonnage ; cette opération, augmente la couche de terre et facilite l'écoulement des eaux ; nous préférons les semis sur billons aux semis à plat, pour cette double raison.

Contrairement aux rubriques précédentes, la *Revue Agricole* ne fait preuve d'aucun profil morphosyntaxique suggérant une prise en charge énonciative explicite. Seuls les pronoms de troisième personne du singulier et du pluriel y sont privilégiés. Si cette rubrique se caractérise par un sur-emploi des adverbes, adjectifs, de la chaîne euphonique, du genre féminin (7,29) et du pluriel (12,01), on remarquera qu'aucune valeur positive relative à la complexité de la phrase n'est notable. Le profil morphosyntaxique suggère donc un

traitement objectivé des thématiques agricoles, différant du traitement descriptif voire technique de la *Chronique agricole* comme le suggère le déficit en adjectifs numéraux.

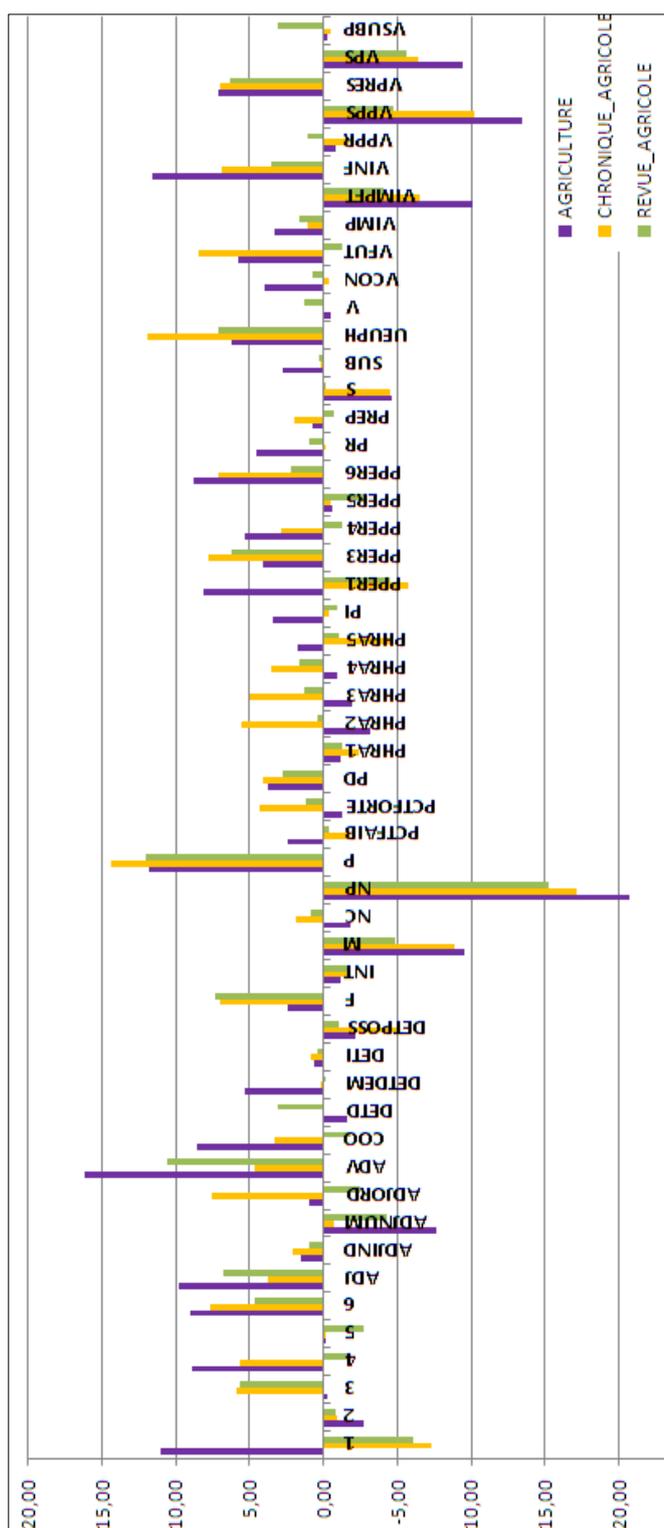


Figure V-55 : Profils distributionnels des rubriques « Agriculture », « Chronique Agricole », « revue Agricole » (en écarts-réduits).

Bibliographie et Bulletin

Enfin, comme les rubriques précédemment évoquées, les profils distributionnels des rubriques *Bibliographie* et *Bulletin* dressent le portrait d'articles rédigés à la première personne du pluriel, marqués par un haut degré de complexité des phrases, comme on peut le supposer d'après le sur-emploi de la subordination et des pronoms relatifs. Ces rubriques se distinguent des cases précédemment observées par leur emploi non-significatif et/ou anecdotique du genre féminin et de la chaîne euphonique.

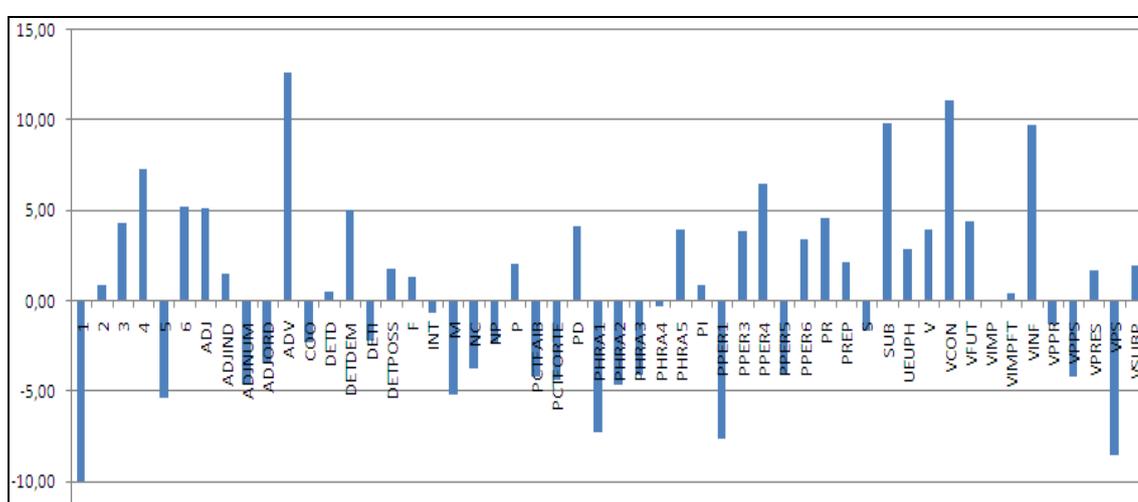


Figure V-56 : Profil distributionnel de « *Bulletin* » (en écarts-réduits).

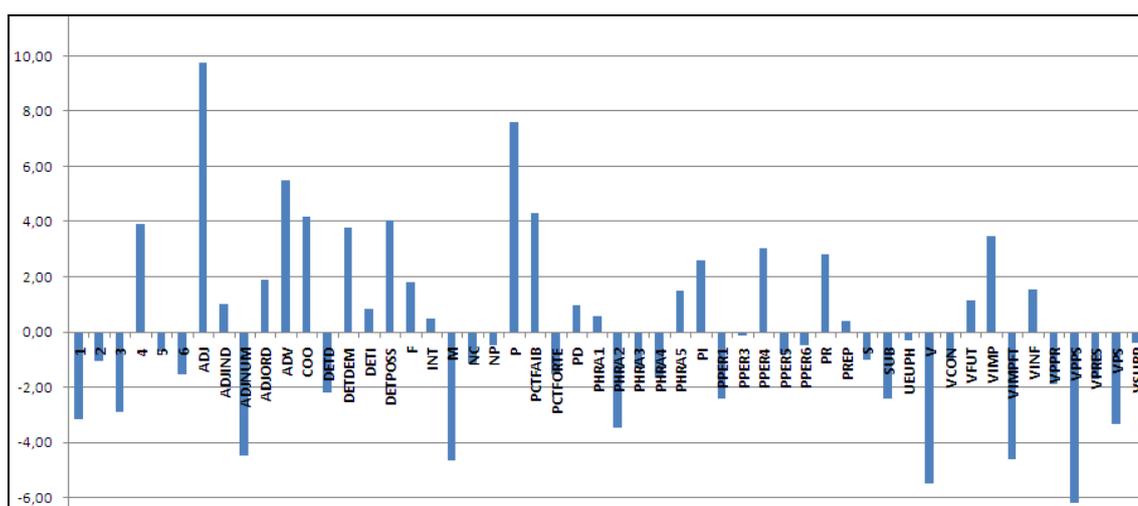


Figure V-57: Profil distributionnel de « *Bibliographie* » (en écarts-réduits).

Conclusion

Au terme de cette classe d'expérimentations, nous soulignerons tout d'abord les résultats prometteurs d'une catégorisation des textes à partir de l'entrée morphosyntaxique.

Le dynamisme des contrastes observés qui se traduit par des écarts-réduits à l'équidistribution systématiquement significatifs, dépassant fréquemment les seuils de +5, de +10 voire de +15, confirme que le plan morphosyntaxique recèle des facteurs plus discriminants encore que celui du vocabulaire. Appréhender l'impact des facteurs discriminants dans toute son ampleur exige de disposer d'un jeu de variables morphosyntaxiques adapté, tant sur le plan quantitatif que sur le plan qualitatif. Des différentes phases d'observations développées au cours de cette analyse, il ressort que si une grande cohérence se donne à lire entre les résultats obtenus à partir d'un panel de variables restreint aux seules parties du discours et à partir d'un jeu d'étiquettes de 52 traits morphosyntaxiques, l'augmentation du nombre de variables présente l'avantage d'affiner la classification textuelle. Cependant, cette augmentation du nombre de variables a pour pendant une sensibilité accrue aux erreurs d'étiquetage commises par *Cordial*. Seule une contre-expérience basée sur un étiquetage fin et contrôlé du texte, – qui n'a malheureusement pas été envisageable ici – permettrait d'évaluer dans quelle mesure le taux d'erreur d'un étiquetage automatique altère les résultats.

Au-delà de ces considérations d'ordre méthodologique, cette première analyse a permis aux variables disponibles de discriminer les diverses rubriques d'un même quotidien. Certaines de ces discriminations étaient plus attendues que d'autres, mais toutes permettent de confirmer le rôle structurant des genres dans les configurations observées. Il en est tout d'abord ainsi de la tension entre texte *objectivé* et texte *subjectivé*, de la longueur des phrases, des temps verbaux, ainsi que du degré de complexité de la phrase, de l'utilisation des adjectifs numériques, mais aussi de façon plus étonnante du

genre grammatical. L'analyse morphosyntaxique des textes de presse écrite, surtout ceux de la presse ancienne et/ou régionale, semble porter en germe une partie notable des bases d'une approche objectivée des typologies de texte journalistiques, habituellement fondées sur des considérations pragmatiques à elles seules contestables (Adam, 1997 ; Charaudeau, 1997, Lorda, 2001). Les déterminations opérées par la forme générique d'un texte étant très loin de se limiter au niveau morphosyntaxique, et s'étendant d'après notre expérience notamment au vocabulaire, l'accès au genre nous semble être conditionné par une approche nécessairement conjointe des niveaux morphosyntaxique et lexical, seuls garants d'un accès aux critères énonciatifs et stylistiques. Le cadre opératoire des deux derniers chapitres de cette recherche démontre le potentiel d'une telle approche, tout en appelant à dépasser le cadre d'une analyse considérant le corpus comme un « sac » de traits linguistiques pour prétendre observer ceux-ci dans la progression du texte : l'observation sur un corpus balisé selon les normes TEI, ouvre dans cette perspective la voie à une approche de la pleine dimension linéaire du texte, tenant compte de sa division en structure logique et/ou en séquences (Adam, 1997)²¹. C'est à cette condition que tient une analyse typologique sur de grands ensembles de données intégrant les caractéristiques compositionnelles des textes. Ces expérimentations et leur développement nous renvoient donc, encore une fois, au rôle décisif du moment d'établissement philologique des données textuelles.

²¹ Dans cette voie, nous avons développé une série d'expérimentations en convoquant le *motif* (Longrée, Mellet, Luong, 2008) comme outil de validation des résultats manifestant une prise en compte, certes localisée et de ressort étroitement linguistique, de la linéarité du texte. Ces expérimentations sont encore insuffisamment probantes (Lethier, 2009).

Conclusion

Au terme de cette thèse, nous voudrions dresser un bilan s'orientant dans deux directions, l'une rétrospective et synthétique, s'attachant à jeter un regard critique sur les réponses apportées par notre travail aux objectifs initiaux ; l'autre, prospective, visant à envisager les suites du processus de recherche dans lequel nous nous sommes engagée.

Nous projetons tout d'abord de contribuer à un programme de dynamisation des lectures de la presse quotidienne régionale de la Troisième République (PQR3R) à partir des voies ouvertes par le numérique. Pour alléger les contraintes liées à son mode de consultation et à son accessibilité, ainsi que celles pesant sur son dépouillement, en explorant les possibilités d'une lecture assistée informatiquement, nous avons mis en œuvre un processus de numérisation de la PQR3R, pionnier par le volume et la nature des données traitées, mais surtout par ses objectifs. Contrairement aux entreprises menées par les bibliothèques municipales et la BNF, se limitant respectivement à une numérisation image de la PQR3R (accès fermé au texte) et à la mise à disposition de données textuelles issues de l'OCR brut (médiocre fiabilité), le projet de constitution de la base PQR3R dans lequel s'inscrivent nos recherches vise en effet à l'acquisition de données plein-texte finement établies. Dans ce dessein, notre travail a amené à développer une méthodologie pour la constitution de grandes bases de données textuelles établies et normalisées pour satisfaire aux besoins des sciences textuelles appelant au partage et à la confrontation des outils et des résultats. De ce point de vue, notre travail représente un essai avancé sur le plan méthodologique : la première campagne de numérisation des fonds du *Petit Comtois* a constitué un laboratoire d'expérimentations.

Il s'est agi, tout d'abord, de développer un protocole d'acquisition de données textuelles à même d'optimiser la qualité des données issues de l'OCR, fortement contrariée par la qualité médiocre des supports-papier, aussi bien que le rythme d'exécution de cette étape première. Notre travail a également porté sur la structuration des données de presse et leur normalisation en XML-TEI, garantes de l'exploitabilité des données et de leur portabilité. Enjeu central pour les sciences textuelles dont les avancées sont conditionnées par le partage et la confrontation des outils et des données, la normalisation des données nous a amenée à recenser et expérimenter diverses propositions permettant d'intégrer aux données des informations ayant pour fonction d'identifier et d'enrichir les données textuelles. En raison du volume des données visées dans la base PQR3R, à savoir près de 200 millions de mots, notre travail s'est attaché à une sélection raisonnée des propriétés pertinentes et à une chaîne de normalisation engendrant, en temps consacré, un investissement lui aussi raisonné. Dans ce cadre, nous avons bénéficié des conseils de l'un des animateurs du consortium TEI, Lou Burnard, avec qui nous avons notamment discuté des divers modes d'enrichissement des entêtes TEI les plus pertinents pour la presse de la Troisième République. Nous avons également coopéré étroitement avec le concepteur de *Sato*, François Daoust, qui nous a fourni une aide fort précieuse pour le versant technique d'encodage de nos données en XML-TEI. Le chantier d'établissement philologique des données du *Petit Comtois* reste aujourd'hui en cours, puisqu'il implique un processus collaboratif interdisciplinaire d'enrichissement des données, notamment au niveau de l'annotation savante des données, et de façon plus prosaïque, un soutien financier continué au projet d'ensemble.

Plus que l'acquisition des ressources numériques des premiers fonds de la base textuelle de PQR3R, le résultat de notre recherche s'inscrit donc dans le programme d'élaboration de protocoles méthodologico-techniques propres aux besoins de la communauté scientifique pour la mise en œuvre de projets de

grande ampleur d'acquisition de données textuelles inscrites initialement sur des documents anciens.

Notre recherche se fixait ensuite pour objectif de fournir un éclairage sur le discours du *Petit Comtois*, analysé au travers d'un dispositif sériel d'observation de quelque 26 000 articles dont la nature et la taille étaient fort hétérogènes, publiés durant la période 1883-1903. Dans ce cadre, nous avons mis en œuvre un parcours d'analyse à visée exploratoire fondé sur l'examen de différentes prises de vues sur la matérialité du discours du *Petit Comtois*. Nous nous sommes appuyée sur différents niveaux d'analyse et différents outils méthodologiques, dont la pluralité se justifie par la nécessité de saisir l'objet textuel dans sa complexité et son hétérogénéité native.

Partant du postulat qu'un texte s'inscrit dans une matérialité scripturale signifiante en cela qu'elle détermine notamment son réseau co-textuel, nous nous sommes tout d'abord tournée vers une analyse de la mise en page du *Petit Comtois*.

D'emblée, nous avons constaté que le terme même de *mise en page* se révélait impropre à décrire ce qui se réduisait à une organisation rudimentaire de la surface du quotidien, où seule la verticalité du plan était exploitée. Conséquence de la densité et de l'austérité de cet espace textuel où les articles de presse se succèdent plus qu'ils ne cohabitent, sans que soit exploitée la fonction incitative des variables typographiques, le parcours de lecture suggéré est fortement contraint.

Dans un deuxième temps, nous avons placé le focus sur le rubriquage, premier niveau de mise en forme, structurant le contenu du quotidien et ritualisant sa production ainsi que sa lecture. Dans ce dessein, nous nous sommes librement inspirée de la méthodologie de l'historien de la presse J. Kayser pour mener un examen quantitatif de la fréquence de parution et de la surface occupée par les rubriques du *Petit Comtois* en diachronie. Celui-ci a permis, d'une part, de

constater la remarquable stabilité de la maquette de 1883 à 1903, et d'autre part, de mettre en relief l'horizon thématique du journal, centré sur les informations de proximité tout en ouvrant largement ses colonnes à la vie politique nationale et internationale, au travers de rubriques telles que *Sénat*, *Chambre des députés* et des *Dépêches de nuit*. Cette dernière rubrique, tant par la surface importante qu'elle occupe que par sa position en Une, constitue la « vitrine » du quotidien, exhibant sa mission d'information tout en exploitant la modernité valorisée de la technologie télégraphique. De plus est ainsi mise en valeur la coïncidence de la temporalité du monde avec celle du journal, rupture essentielle introduite par le quotidien à la fin du XIX^e siècle, également perceptible dans d'autres intitulés tels que *Au jour le jour*, *Dépêches de la journée*, *Dernière Heure*. C'est en considérant les critères sous-jacents aux intitulés de rubriques que nous avons d'ailleurs mis au jour le contraste entre cette modernité affichée, ressort d'une hiérarchisation de la valeur de l'information, et la persistance des liens du discours du *Petit Comtois* avec la matrice littéraire, puisqu'un grand nombre d'intitulés de rubriques réfèrent aux genres discursifs, tandis qu'aucun découpage thématique du monde n'est proposé par le journal. À l'image de la mise en page générale, le rubriquage du *Petit Comtois* est rudimentaire et repose sur des critères sous-jacents que l'on trouve justement honnis dans les manuels de journalisme de la presse de la seconde moitié du XX^e siècle.

L'examen des titres dans *Le Petit Comtois* (3.3.) a abouti au constat que leur mise en œuvre demeure également sporadique dans ce quotidien, du moins dans les dix premières années de notre corpus. L'analyse ciblée des titres de la rubrique des *Dépêches de nuit* nous a permis de préciser les caractéristiques syntaxiques de ces titres, d'un calibrage très court (3,5 mots en moyenne) et reposant sur un syntagme nominal. Loin de restreindre notre analyse à une simple observation des traits formels des titres, nous avons mis en place une grille d'analyse fonctionnaliste, intégrant les dimensions pragmatiques et rhétoriques des titres, étudiés en lien avec la dépêche qu'il surplombe. Cette

analyse nous a permis de mettre au jour les fonctions dominantes des titres, à savoir classifier et commenter l'information plus que la condenser anticipativement. Ce faisant, nous avons montré que le niveau péritextuel offre une prise essentielle sur les ressorts discursifs par lesquels s'articule la mission d'information dont se targue le quotidien, - dont les dépêches sont la clef de voûte -, avec son statut d'organe porte-parole d'une sensibilité politique. En outre, cette étude a été l'occasion de montrer que les titres des dépêches fournissaient un accès privilégié à l'identité du quotidien, au sens où c'est à ce niveau que *Le Petit Comtois* donne sa teinte propre à l'information qui lui est délivrée par ses correspondants et l'agence Havas. Aux aspects factuels de ces dernières, *Le Petit Comtois* oppose, dans ces titres, une parole railleuse ; il joue ici la carte du franc-parler pour entretenir sa connivence avec le lecteur et pour mieux le guider sur ses rails idéologiques.

Nous avons ensuite pris pour entrée du discours du *Petit Comtois* le niveau du vocabulaire, avec pour projet de mettre en œuvre un parcours interprétatif objectivé par les méthodes de l'analyse statistique. Cette étude comporte plusieurs volets d'analyse, envisagés comme autant de focus sur la textualité.

D'une première enquête, traditionnelle en lexicométrie, sur la ventilation du vocabulaire en diachronie, il est tout d'abord ressorti que le discours du *Petit Comtois* était soumis à un clivage important entre les séquences chronologiques 1884-1896 et 1897-1903. En nous appuyant sur les fonctionnalités de représentation géodésique et d'hypertexte expert du logiciel *Astartex*, nous avons examiné le vocabulaire propre à ces séries dans un mouvement de va-et-vient constant avec le texte, de sorte à en construire une interprétation contrôlée. Il en ressort que la première séquence chronologique reflète les troubles des débuts d'une Troisième République, déterminés, d'une part, quant à la politique extérieure sur fond revanchard et colonialiste, et d'autre part, quant à la politique intérieure, par une série d'affaires politico-

financières et par la politique d'apaisement menée par J. Méline, source de divisions, et ayant pour pendant une radicalisation du discours du quotidien perceptible dès 1892. À cette période s'oppose une séquence (1897-1903) marquant le tournant radical de la Troisième République et du *Petit Comtois*, dont l'orientation radicale est revendiquée pourtant depuis 1891. Ce n'est aussi qu'à partir de 1897 que ses colonnes témoignent de l'émergence et de l'intensification de la question sociale, des mutations des forces politiques en présence et plus particulièrement de la fortune du socialisme. On y entrevoit également les débats autour de l'*enseignement* et les *congrégations*, associés à un violent discours anticlérical.

Dans un second mouvement (4.3), nous avons mis en œuvre une analyse des réseaux de co-occurrence du vocabulaire, mode d'accès privilégié à la textualité, dont l'analyse des données textuelles ne découvre aujourd'hui que les prémisses et les potentialités. Il ressort de cette analyse que quatre pôles *isotropiques* structurent le vocabulaire notre corpus, donnant accès aux dominantes lexico-thématiques et rhétorico-stylistiques du *Petit Comtois*. Ces pôles, ayant émergé du corpus lui-même, renvoient respectivement (1) à la place centrale des faits divers et leur traitement en justice, (2) à la prégnance de la vie législative, (3) au vocabulaire du politique à proprement parler, et enfin, (4) indiquent la fonction d'agenda du quotidien régional, chargé de donner le programme des réseaux de sociabilité locaux.

En cherchant à confronter ces résultats à ceux livrés par l'analyse de la ventilation du vocabulaire par année de parution, nous avons ensuite observé l'évolution diachronique de la configuration fine du vocabulaire (4.4). La fonctionnalité de projection géodésique colorisée, en nous donnant accès à une comparaison simultanée des séquences 1883-1893 et 1893-1903, nous a permis de confirmer la remarquable invariance de ces pôles en diachronie, que des manipulations exploratoires, basées sur des analyses successives de la configuration fine du vocabulaire de différentes années de parution, avaient suggérée. Cette stabilité de l'organisation lexicale réticulaire contrarie en partie

nos attentes, fondées sur le postulat selon lequel le discours de presse, soumis aux aléas de l'actualité, devait se caractériser par une évolution des thèmes et thématisations, et contraste avec la dichotomie chronologique observée par l'analyse macro-distributionnelle. Ainsi a-t-elle constitué un résultat tout à fait stimulant. De la première moitié (1883-1893) de notre corpus à la seconde (1883-1893), rares sont les formes telles qu'*ouvriers*, *citoyen*, ou encore *propriété* qui présentent de fortes variations d'emploi : nous avons montré que la migration sur l'espace géodésique de ces formes, renvoyant à des concepts clés du vocabulaire socio-politiques du XIX^e siècle, traduisaient des évolutions sémantiques essentielles. Nous voudrions insister sur les apports de cette méthodologie, permettant d'avoir accès à l'évolution des réseaux co-occurentiels d'une forme, et qui nous semble extrêmement prometteuse pour une observation des concepts de la Troisième République dans un cadre interdisciplinaire. Elle constitue, selon nous, un probant mode d'exploration du niveau sémantique basé sur la mesure de l'évolution diachronique des réseaux contextuels, qui intéressera l'historien pour sa valeur non seulement résultative mais surtout heuristique, qui peut prendre toute son ampleur dans l'hétérogénéité de l'archive.

Enfin, partant de l'hypothèse que la consistance lexico-thématique des pôles isotropiques évoquait très fortement les rubriques du *Petit Comtois*, nous avons mené un volet d'analyse visant à valider ou infirmer cette hypothèse. Nous avons donc consacré le dernier temps de nos explorations du vocabulaire (4.5) à examiner dans quelle mesure le rubriquage structure la configuration fine du vocabulaire, corrélation positive confirmée par l'examen des spécificités lexico-thématiques des rubriques. Plus que d'être soumis à l'imprévisibilité du monde et à l'essence fluctuante de l'actualité, le discours du journal républicain se déploie ainsi sur un fond d'invariant, dont nous formulons l'hypothèse qu'il est l'indice de la formation discursive qui le détermine. La mise en scène du discours du *Petit Comtois* renvoie ainsi aux pratiques de sélection de ce qui constitue l'actualité selon les rails idéologiques de la formation discursive

républicaine de la Troisième République, déterminant les éléments lexicaux, les structures syntaxiques et stylistiques que ces rails imposent pour dire l'information.

Dans la continuité de ces observations sur le niveau du vocabulaire, nous avons interrogé, dans notre dernier chapitre, les apports d'une méthode de typologie textuelle à partir du niveau morphosyntaxique, supposé révéler des critères discriminants plus complexes que le vocabulaire. Sans viser à une typologie textuelle aboutie en elle-même des articles et des rubriques du *Petit Comtois*, ce volet d'analyse a été motivé par le projet de fournir une première étape de consolidation méthodologique d'un protocole pour l'analyse des genres du *Petit Comtois* et de la PQR3R. La première étape consistait à identifier les risques d'un étiquetage automatique de notre corpus, tel que convoqué par la logométrie (Mayaffre, 2000) et la majorité des entreprises de caractérisation textuelle (Malrieu & Rastier, 2001, Beauvisage, 2001). Nous avons rendu compte de façon synthétique des grandes classes d'erreurs commises par *Cordial* appliqué au discours de presse du XIX^e siècle, tandis que les résultats de ce logiciel avaient jusqu'à présent été évalués à partir de corpus littéraires ou de textes scientifiques du XX^e siècle. Nous avons notamment insisté sur les très fréquentes erreurs d'étiquetage des noms composés, des locutions adverbiales, des noms propres, des nombres, s'ajoutant aux erreurs bien connues d'identification des temps et modes verbaux et des polycatégoriels. Ce faisant, cet examen a été l'occasion de revenir sur les limites d'un étiquetage non contrôlé, livré par une « boîte noire », tandis que la pluralité des types de discours ainsi que la finesse du grain de l'étiquetage visée par une analyse textuelle du discours exigent de pouvoir modifier et ajouter à souhait des éléments aux dictionnaires et aux grammaires du système.

Le second objectif de ces expérimentations était de proposer un protocole méthodologique robuste et reproductible d'analyse multidimensionnelle. Il s'est

ainsi agi, d'une part, d'explorer l'emploi de l'AFC et de la représentation géodésique en vue d'une analyse moins dichotomisante que l'ACP habituellement convoquée dans les entreprises typologiques, et d'autre part, de procéder à une série d'expérimentations sur le nombre de traits à prendre en compte pour l'analyse. Nous avons constaté que des analyses respectivement basées sur les seules catégories du discours et sur un nombre plus large de traits (52 en l'occurrence) proposaient des résultats convergents, mais qu'un jeu d'étiquettes plus large permettait néanmoins de les affiner de façon intéressante. Cependant, l'augmentation du nombre de variables morphosyntaxiques issues d'un étiquetage automatique accroît problématiquement les risques liés aux erreurs d'analyse. Nous avons ainsi émis de sérieuses réserves quant aux résultats fournis par l'analyse des variables morphosyntaxiques. Un certain nombre d'indications pour le moins inattendues, parmi lesquelles la non pertinence de l'opposition entre le groupe nominal et le groupe verbal dans la structuration de la typologie, pourtant habituellement observée sur corpus clos, restent à mettre à l'épreuve par une analyse basée sur un étiquetage fin et contrôlé.

Il reste que le dynamisme des contrastes relevés par l'analyse des rubriques en traits morphosyntaxiques confirme la pertinence d'une entreprise de typologisation basée sur cette entrée. Le croisement des résultats livrés par le niveau morphosyntaxique et le niveau du vocabulaire, vers lequel convergent les chapitres IV et V de cette recherche, est une voie méthodologique à approfondir pour l'analyse des genres discursifs.

Si les études que nous avons menées dans les chapitres III, IV et V restent partiellement cloisonnées, c'est le sort d'une entreprise comme la nôtre à cette étape de son déploiement : montrer à la fois la faisabilité, les conditions et l'intérêt d'une telle démarche pionnière. Nous présentons dans cette thèse aussi bien les résultats positifs (probants par positivité), que les incertains, et surtout les négatifs, qui, en matière de recherche à visée scientifique, valent les

positifs, en cela qu'ils permettent d'abandonner des hypothèses erronées et d'ouvrir de nouveaux champs inattendus.

Il était et reste indispensable de dépasser l'analyse lexicale, longtemps privilégiée par la lexicométrie et l'analyse du discours assistée statistiquement, en y associant l'analyse des niveaux plus profonds relevant d'une annotation, donc d'une (pré-)interprétation. Nos recherches sont encore loin d'atteindre le cœur de l'analyse des concepts, mais en esquissent les conditions : l'intérêt des historiens est donc attiré sur la méthodologie proposée et son cadre à vocation interdisciplinaire, plus que sur des résultats historiographiques.

Perspectives

Ces premiers résultats, sur la méthodologie du corpus que sur le discours proprement dit du *Petit Comtois* nous incitent aujourd'hui à prolonger ces travaux dans deux directions : un programme de validation ou d'infirmité de nos résultats, qui implique une comparaison avec un ensemble plus large de données et la confrontation à d'autres outils d'une part, et d'autre part, la mise en œuvre d'analyses ciblées sur des phénomènes discursifs fins.

En premier lieu, nos résultats, portant sur un corpus échantillonné (un mois par année de parution) appellent désormais une étude menée sur un grand corpus issu du collationnement continu de la collection complète du *Petit Comtois* afin de confirmer, d'infirmer ou d'affiner nos résultats selon les cas. Cette phase d'analyse présentera aussi, précisément, l'intérêt d'évaluer la pertinence même d'une telle pratique d'échantillonnage du discours de presse. Ces phases d'expérimentations impliquent la poursuite du développement de la base textuelle *Petit Comtois*, à la direction duquel nous sommes associée. Les axes de ce développement sont tout d'abord relatifs à l'enrichissement des ressources textuelles selon les recommandations TEI. Il s'agit, à terme, d'offrir la possibilité à l'utilisateur d'interroger l'ensemble des articles de la base en vue de la constitution de corpus, interrogation qui pourra se fonder sur différents états du texte. Le projet prévoit en premier lieu de relier chaque article, envisagé comme un fichier individuel XML-TEI, à un fichier image PDF haute

définition correspondant, convocable d'un simple clic. En second lieu, il s'agit de donner accès aux données lemmatisées de la base : le projet PQR3R prévoit ainsi de marquer les lemmes correspondant aux occurrences de fréquence haute et moyenne, et de forte spécificité, par comparaison aux proportions de la base *Frantext* sur la même époque.

Ces objectifs impliquent de développer un environnement d'annotation convivial alternant phases automatiques et phases de décisions expertes, permettant de reconnaître de manière contrôlée les formes simples, mais aussi les mots composés, les formes verbales auxiliées, et d'annoter les phraséologies. Ce sont précisément les directions que se fixe l'équipe qui a encadré notre thèse dans le développement de *Diatag*. Dans le même temps, les développements de ce logiciel visent à améliorer la qualité des données textuelles issues de l'OCR, en proposant une approche nouvelle des modalités de correction. En effet, la correction des sorties d'OCR, dans les traitements de texte ou dans *ABBY FINEREADER 9.0*, se font sur un mode linéaire, document par document, au moyen d'un vérificateur orthographique et grammatical dérivé de ceux des traitements de texte. Ces solutions sont peu ergonomiques et très faiblement rentables, tandis qu'une correction entièrement automatique est incompatible avec une pratique philologique. L'amélioration des conditions de vérification des OCR que visent à opérer les développements de *Diatag* repose sur l'enrichissement des ressources linguistiques et des procédures d'apprentissage au cours d'une lecture assistée du texte, menée par un philologue expert : il s'agit donc d'intégrer ces vérifications dans le continuum de la chaîne de travail philologique, en lien avec la reconnaissance et l'annotation linguistiques de sorte à rendre compatible l'exigence de finesse du grain des données et la faisabilité d'une telle opération sur une très grande base textuelle. Nous entendons affermir notre participation à ce projet aux multiples ramifications interdisciplinaires et internationales.

Ces développements conditionnent directement les contre-épreuves d'ampleur que nous avons évoquées dans le chap.V sur l'étiquetage

(automatique vs assisté) du niveau morphosyntaxique (et lexical, puisqu'ils sont étroitement reliés) pour évaluer l'influence du taux d'erreur d'un étiquetage automatique sur une typologie textuelle et/ou discursive et, partant, consolider le socle méthodologique d'une analyse des genres dans un corpus de presse. Cette optique comparative implique impérativement d'optimiser et d'achever l'étiquetage fin et contrôlé des données de notre corpus. L'annotation linguistique contrôlée s'impose ainsi comme un chantier à dynamiser très rapidement, d'autant plus qu'il doit constituer un véritable état du texte et qu'il conditionne encore d'autres expérimentations : notamment, nous souhaiterions examiner si, au contraire des résultats présentés par les analyses textométriques de première intention, les analyses de la co-occurrence généralisée menée sur l'entrée morphosyntaxique aboutissent à des résultats différents de ceux menés sur le lexique. Ces recherches se poursuivront dans le cadre auquel nous avons étroitement coopéré au cours des quatre années de cette recherche : par cercles concentriques, le développement des logiciels *Diatag* et *Astartex* dans notre équipe de recherche ; la coopération très intégrée avec les concepteurs et responsables d'*Hyperbase* à l'UMR BCL de Nice, avec ceux de SATO à l'UQAM ; la collaboration scientifique et technique avec les partenaires du projet ANR *Textométrie*, et avec le consortium TEI¹.

Dans le cadre de notre travail de thèse, les explorations de notre corpus se sont fondées de façon systématique sur le logiciel *Astartex*. Une fois le corpus entièrement normalisé dans un langage pivot XML-TEI, nous pourrions le soumettre aux fonctionnalités et principes d'autres outils lexicométriques participant aux projets *ATONET* (CRSH du Canada) et *Textométrie* (ANR), pourvu que le volume de ce corpus et ses partitions soient tolérées par ces mêmes logiciels (ce qui en constituera donc un test grandeur nature).

¹Lou Burnard, l'un des animateurs les plus actifs du consortium TEI, utilise notre corpus depuis août 2009 pour les développements des normes TEI.

Nous pourrions dès lors confronter et enrichir nos résultats à l'aune des principes respectifs de ces outils d'autres interfaces d'exploration. Nous pensons notamment à une analyse de notre corpus recourant aux fonctionnalités de *Lexico 3* qui nous permettrait d'enrichir nos observations sur la dimension linéaire du texte ainsi que de confronter les méthodologies permettant de faire émerger du texte les unités phraséologiques. Dans cette même perspective, nous souhaiterions par ailleurs procéder à la confrontation de nos résultats sur la configuration fine du vocabulaire livrés par *Astartex* avec les structures signifiantes que l'outil *Alceste* se propose de dégager: il s'agirait de procéder à une comparaison des pôles d'attractivité structurant le vocabulaire, tels que les deux outils les font ressortir, pour mieux mettre en relief l'impact des choix théoriques et technico-méthodologiques sous-jacents aux deux outils sur l'objet qu'ils permettent respectivement d'appréhender.

Parallèlement à ces compléments et contre-épreuves qui relèvent d'un travail d'équipe très intégré associant les développeurs, les statisticiens et les linguistes, nous entendons prolonger nos recherches en observant l'évolution générale du discours du *Petit Comtois* sur une diachronie plus longue, et à terme sur la durée entière du titre. Il s'agirait tout autant de nous concentrer sur l'évolution de la mise en scène de l'information, des formes journalistiques, que sur l'évolution du sens en diachronie des termes du vocabulaire social et politique que les méthodes lexicométriques mettront au jour dans ce corpus étendu. La construction d'un index et des concordances des termes du vocabulaire politique et social les plus saillants par leur indice de fréquence nous semblerait représenter une intéressante contribution à l'histoire des concepts et à l'affirmation d'une théorie de ces concepts telle que construite et maintenue par l'Analyse de Discours « du côté de l'histoire ». Par ailleurs, il s'agirait d'affiner l'analyse de ces concepts fondamentaux en les observant à la lumière d'un corpus pleinement réflexif, incluant par exemple des discours parlementaires, des textes juridiques, des proclamations électorales, etc.,

permettant d'appréhender la genèse, la circulation et la disparition de concepts dans le vocabulaire politique et social du XIX^e siècle.

Cet affermissement conceptuel et théorique est à la fois un enjeu et une précondition du programme interdisciplinaire auquel nous sommes associée, autour d'une exploitation pleinement historique et des sciences humaines des nouvelles ressources numériques en cours de construction avec notre modeste contribution.

Dans un contexte technologique très évolutif et face à la richesse de la ressource qu'est la presse régionale de la Troisième République, la création d'un espace de dialogue s'impose entre la communauté des sciences humaines (historiens, sociologues, analystes du discours, spécialistes de l'information et de la communication) et la communauté des technologies numériques, dialogue dont le but et les moyens échangent constamment leur place : favoriser la connaissance de cette époque historique et de l'histoire de la presse d'une part, et, d'autre part, la construction des ressources, outils et méthodes d'une science des textes au service de l'analyse du discours.

Bibliographie

- ABOUDA, L.** (2001) "Les Emplois journalistique, polémique et atténuatif du conditionnel." in DENDALE, P. & TASMOWSKI, L. (**Dir.**), *Recherches linguistiques*, n° 25 : "Le Conditionnel en français", Paris : Klincksieck. pp. 277-294.
- ABLALI, D.** (2006) "Écrire en critique : exploration morpho-syntaxique sur corpus." in RASTIER, F. & BALLABRIGA, M. (**Dir.**) *Corpus en Lettres et Sciences sociales. Des documents numériques à l'interprétation*, Paris : Texto! Disponible sur <http://www.revue-texto.net/Parutions/Livres-E/Albi-2006/Actes_ALBI-06.pdf> (Consulté le 12/10/2009).
- ACHARD, P.** (1995) *La Sociologie du langage*. Paris : Presses Universitaires de France.
- ACHARD, P., GRUENAI, M.P. & JAULIN, D. (Dir.)** (1984) *Histoire et linguistique : Actes de la table ronde, Paris, les 28, 29, 30 avril 1983*. Paris : Editions de la Maison des Sciences de l'homme.
- ADAM, J.-M.** (1984) *Le Récit*. Paris : Presses Universitaires de France, Coll. Que sais-je?
- ADAM, J.-M.** (1992) *Les Textes : types et prototypes. Récit, description, argumentation, explication et dialogue*. Paris : Nathan Université.
- ADAM, J.-M.** (1997) "Unités rédactionnelles et genres discursifs : Cadre général pour une approche de la presse écrite." in *Pratiques*, n°9. Metz : CRESEF. pp. 3-18.
- ADAM, J.-M.** (1999) *Linguistique textuelle. Des genres de discours aux textes*. Paris : Nathan.
- ADAM, J.-M.** (2001a) "Entre conseil et consigne : les genres de l'incitation à l'action." in *Pratiques*, n°111/112, Metz : CRESEF. pp. 7-38.
- ADAM, J.-M.** (2001b) "Types de textes ou genres de discours? Comment classer les textes qui disent de et comment faire." in *Pratiques*, n°141. Metz : CRESEF. pp. 10-27.
- ADAM, J.-M.** (2004) "Des genres à la généricité. L'exemple du conte." in *Langages*, n°105. Paris : Larousse. pp. 62-72.
- ADAM, J.-M.** (2005a) *La Linguistique textuelle. Introduction à l'analyse textuelle des discours*. Paris : Armand Colin.

- ADAM, J.-M.** (2005b) "Les sciences de l'établissement des textes et la question de la variation." in ADAM, J.-M. & HEIDMANN, R. (**Dir.**), *Sciences du texte et analyse de discours. Enjeux d'une interdisciplinarité*. Genève : Slatkine érudition.
- ADAM, J.-M.** (2006a) "Texte, contexte et discours en questions." in *Pratiques*, n°129-130. Metz : CRESEF. pp. 21-34.
- ADAM, J.-M.** (2006b) "Autour du concept de texte. Pour un dialogue des disciplines de l'analyse de données textuelles." in *JADT2006*. Disponible sur <http://www.cavi.univparis3.fr/lexicometrica/jadt/JADT2006-PLENIERE/JADT2006_JMA.pdf> (Consulté le 14/03/2009).
- ADAM, J.-M. & LUGRIN, G.** (2000a) "L'hyperstructure : un mode privilégié de présentation des événements scientifiques." in CUSIN-BERCHE, F. (**Dir.**) : *Rencontres discursives entre science et politique. Spécificités linguistiques et constructions sémiotiques*, Carnets du CEDISCOR, n°6. Paris : Presses de la Sorbonne Nouvelle. pp. 133-149.
- ADAM, J.-M. & LUGRIN, G.** (2000b) "Variations des ancrages énonciatifs et fictionnalisation d'une anecdote d'Albert Camus." in *Langue française*, n°128. Paris : Larousse. pp. 96-112.
- ADAM, J.-M & LUGRIN, G.** (2006) "Effacement énonciatif et diffraction co-textuelle de la prise en charge des énoncés dans les hyperstructures journalistiques." in *Semen*, n°22 : "Énonciation et responsabilité dans les médias". Besançon : Presses Universitaires de Franche-Comté. pp. 131-148.
- ADAM, J.-M., ALI BOUACHA, M., GRIZE, J.-B., et al. (Dir.)** (2004) *Texte et discours : catégories pour l'analyse*. Dijon : Editions Universitaires de Dijon.
- ADAM, J.-M. & HEIDMANN, U. (Dir.)** (2005) *Sciences du texte et analyse de discours. Enjeux d'une interdisciplinarité*. Genève : Slatkine.
- AGNÈS, Y.** (2002) *Manuel de journalisme : écrire pour le journal*. Paris : La Découverte.
- AGUET, J.-P.** (1960) "Le tirage des quotidiens de Paris sous la monarchie de juillet." in *Revue suisse d'histoire*, Vol. 10. pp.216-286. Disponible sur <<http://retro.seals.ch/openurl?rft.issn=0036-834&rft.date=1960&lPage=216>> (Consulté le 10/07/2009).
- AGULHON, M.** (1990) *La République, 1880 à nos jours*. Paris : Hachette.
- AGULHON, M.** (1991) "Le parti républicain." in HAMON, L. (**Dir.**), *Les opportunistes. Les débuts de la République aux républicains*. Paris : Editions de la maison des Sciences de l'Homme, pp. 1-17.
- ALBERT, P.** (1965) "Remarques sur l'histoire de la presse sous la IIIème République." in *Le Mouvement social*, n°53. Paris : Association Le Mouvement Social, pp. 23-37.

- ALBERT, P.** (1973) (**Dir.**) *Documents pour l'histoire de la presse de province dans la seconde moitié du XIXe siècle*. Paris : Editions du CNRS.
- ALBERT, P.** (1975) "Remarques sur les recherches en histoire de la presse." in *Bulletin d'histoire moderne et contemporaine*, Vol. 9. Paris : Société d'histoire moderne et contemporaine. pp. 39-72.
- ALBERT, P.** (1980) *Histoire de la presse politique nationale au début de la Troisième République (1871-1879)*. Paris : Champion.
- ALBERT, P.** (1989a) "Remarques sur l'histoire de la presse de province française." in **CRIQUI, E. & LAPRÉVOTE, L.-P. (Dir.)**, *Le quotidien dévoilé. L'Est-Républicain (1889-1989)*. Jarville-La Malgrange : Editions de l'Est. pp. 15-21.
- ALBERT, P. (Dir.)** (1989b) *Lexique de la presse écrite*. Paris : Dalloz.
- ALBERT, P.** (2003) *Histoire de la presse*. Paris : Presses Universitaires de France, Coll. Que sais-je?
- ALBERT, P., FEYEL, G. & PICARD, J.-F.** (1977) (**Dir.**) *Documents pour l'histoire de la presse nationale aux XIXe et XXe siècles*. Paris : Editions du C.N.R.S.
- ALBERTAN, C. & ALBERTAN, S.** (1989) "Les silences de la presse provinciale en 1788." in **RÉTAT, P. (Dir.)**, *La Révolution du Journal*. Lyon : Editions du CNRS, pp 25-37.
- AMAURY, F.** (1972) *Histoire du plus grand quotidien de la III^e République. Le Petit Parisien, 1976-1944*. Paris : Presses Universitaires de France.
- AMBROISE-RENDU, A.-C.** (1997) *Les Faits divers dans la presse française de la fin du XIX^e siècle (1870-1910)*. Thèse de doctorat en Histoire. Paris : Université de Paris I.
- AMOSSY, R.** (1991) *Les Idées reçues : Sémiologie du stéréotype*. Paris : Nathan.
- AMOSSY, R.** (1999) *Images de soi dans le discours : la construction de l'ethos*. Paris : Delachaux et Niestlé.
- AMOSSY, R.** (2000) *L'Argumentation dans le discours : Discours politique, littérature d'idées, fiction*. Paris : Nathan Université.
- AMOSSY, R. & HERSCHBERG-PIERROT, A.** (1997) *Stéréotypes et clichés : Langue, discours, société*. Paris : Nathan Université.
- AMOSSY, R. & MAINGUENEAU, D.** (2003) (**Dir.**) *L'analyse du discours dans les études littéraires*. Toulouse : Presses universitaires du Mirail.
- ANDRÉ, J. & CHABIN, M.-A.** (1999) "Numériser les documents anciens, et après?" in *Document numérique*, n°3 : "Les documents anciens". Paris : Lavoisier. pp. 7-13.
- ANDRÉANI, R.** (1989) *La presse quotidienne de Montpellier des origines à 1944*. Thèse de doctorat en Histoire. Toulouse : Université Toulouse-Le Mirail.

- ANGENOT, M.** (1982) *La parole pamphlétaire : contribution à la typologie des discours modernes*. Paris : Payot.
- ANGENOT, M.** (1989) *1889. Un état du discours social*. Longueuil : Edition du préambule.
- ANGENOT, M.** (1991) "L'analyse du discours : esquisse d'une problématique générale." in *Bulletin de l'ACLA*, n°13 (1). Montréal : ACLA. pp. 9-20.
- ANSCOMBRE, J.-C.** (1994) "Proverbes et formes adverbiales : valeur évidentielle et argumentative". in *Langue française*, n°102. Paris : Larousse. pp. 95-107.
- ARNAUD, P.-J.-L.** (1991) "Réflexions sur le proverbe." in *Cahiers de lexicologie*, 59 (2). Paris : Honoré Champion. pp. 6-27.
- ARON, P., SAINT-JACQUES, D. & VIALA, A. (Dir.)** (2002) *Le dictionnaire du Littéraire*. Paris : Presses Universitaires de France.
- ASTON, G. & BURNARD, L.** (1998) *The BNC handbook : Exploring the British National Corpus with SARA*. Edinburgh : Edinburgh University Press.
- AUBRY, M. & WESTEEL, I.** (2003) *La numérisation des textes et des images : techniques et réalisations*. Lille : Editions du conseil scientifique Charles de Gaulle.
- AUCLAIR, G.** (1970) *Le Mana quotidien. Structures et fonctions de la chronique des faits divers*. Paris : Anthropos.
- AUROUX, S.** (1994) *La révolution technologique de la grammatisation*. Paris : Mardaga.
- AUROUX, S.** (1998) *La raison, le langage et les normes*. Paris : Presses Universitaires de France.
- AUTHIER, J.** (1978) "Les formes du discours rapporté. Remarques syntaxiques et sémantiques à partir des traitements proposés." in *DRLAV*, n°17, pp.1-78.
- AUTHIER-REVUZ, J.** (1982) "Hétérogénéité montrée et hétérogénéité constitutive : éléments pour une approche de l'autre dans le discours." in *DRLAV*, n°26. pp. 91-151.
- AUTHIER-REVUZ, J.** (1992) "Repères dans le champ du discours rapporté." in *L'information grammaticale*, n°55. Paris : Bibliothèque de l'information grammaticale. pp. 38-42.
- AUTHIER-REVUZ, J.** (1995) *Ces Mots qui ne vont pas de soi. Boucles réflexives et non-coïncidences du dire*. Paris : Larousse.
- AUTHIER-REVUZ, J.** (2000a) "Aux risques de l'allusion." in MURAT, M. (Dir.), *L'Allusion dans la littérature*. Paris : Presses Universitaires de la Sorbonne.

- AUTHIER-REVUZ, J.** (2000b) "Le Fait autonymique : Langage, Langue, Discours — Quelques repères." Communication présentée lors de *Le Fait Autonymique dans les langues et les discours*, Paris, Organisation de SYLED.
- AVENEL, H.** (1900) *Histoire de la Presse française, depuis 1789 à nos jours*. Paris : Flammarion.
- BAKHTINE, M.** (1977) *Le marxisme et la philosophie du langage*. Paris : Editions de Minuit.
- BAKHTINE, M.** (1984) *Esthétique de la création verbale*. Paris : Gallimard.
- BARBOTIN, E. (Dir.)** (1975) *Qu'est-ce qu'un texte? Eléments pour une herméneutique*. Paris : Librairie José Corti.
- BARDIN, L.** (1989) *L'Analyse de contenu*. Paris : Presses Universitaires de France.
- BARTHES, R.** (1967) "Le discours de l'histoire." in *Informations sur les Sciences sociales*, vol. 6, 4. Paris : Maison des Sciences de l'Homme. pp. 65-75.
- BARTHES, R.** (1993-1994), *Oeuvres complètes*. Paris: Seuil.
- BEACCO, J.-C.** (1992) "Les genres textuels dans l'analyse du discours." in *Langages*, n°105. Paris : Larousse. pp. 8-27.
- BEAUVISAGE, T.** (2001) "Exploiter des données morphosyntaxiques pour l'étude statistique des genres textuels : application au roman policier." in *TAL*, Vol. 42, 2. Jussieu : CNRS/Association pour le Traitement Automatique des Langues. pp. 579-608.
- BECKER, J.-J.** (1988) "L'opinion." in RÉMOND, R. (Dir.), *Pour une histoire politique*. Paris : Seuil, pp. 161-183.
- BELLANGER, C., GODECHOT, J. & GUIRAL, P. (Dirs.)** (1969) *Histoire générale de la presse française. Tome 1, Des origines à 1814*. Paris : Presses Universitaires de France.
- BELLANGER, C., GODECHOT, J. & GUIRAL, P. (Dirs.)** (1972) *Histoire générale de la Presse française. Tome 3 : "De 1871 à 1940"*. Paris : Presses Universitaires de France.
- BELLANGER, C., GODECHOT, J., GUIRAL, P., et al. (Dirs.)** (1975) *Histoire générale de la presse française*. Paris : Presses Universitaires de France.
- BELLET, R.** (1976) *Jules Vallès, journaliste, 1857-1885*. Thèse de doctorat en Lettres. Clermont-Ferrand : Université de Clermont-Ferrand.
- BELOIEL, D.** (1994) "Le concile Vatican I dans la presse nantaise (1867-1870)." in *Annales de Bretagne et des pays de l'Ouest*, Vol. 101, 3. Rennes : Presses Universitaires de Rennes. pp. 111-125.

- BENVENISTE, E.** (1966) *Problèmes de Linguistique générale*. Tome 1. Paris : Gallimard.
- BENVENISTE, E.** (1974) *Problèmes de Linguistique Générale*. Tome 2. Paris : Gallimard.
- BENZÉCRI, J.-P.** (1973a) *L'Analyse de données*. Tome 1. Paris : Dunod.
- BENZÉCRI, J.-P.** (1973b) *L'Analyse des données*. Tome 2 : L'analyse des correspondances. Paris : Dunod.
- BENZÉCRI, J.-P.** (1981) *Pratique de l'Analyse des données*. Tome 3 : Linguistique et lexicologie. Paris : Dunod.
- BERSTEIN, S. & RUDELLE, O. (Dir.)** (1992), *Le modèle républicain*. Paris : Presses Universitaires de France.
- BESSON, A.** (1977) *La presse locale en liberté surveillée*. Paris : Editions ouvrières.
- BIBER, D.** (1988) *Variation across Speech and Writing*. Cambridge : Cambridge University Press.
- BIBER, D.** (1989) "A typology of English Texts." in *Linguistics*, n°27. pp. 3-43.
- BIBER, D.** (1993) "Using register-diversified corpora for general language studies." in *Computational Linguistics*, 2, 19. Cambridge (MA) : MIT Press. pp. 243-258.
- BIBER, D.** (1995) *Dimensions of Register Variation*. Cambridge : Cambridge University Press.
- BIBER, D.** (2000) "Investigating language use through corpus-based analyses of association patterns." in BARLOW, M. & KEMMER, S. (Dir.), *Usage-based models of language*. Stanford : CSLI Publications.
- BIBER, D.** (2004) "Conversation text types : a multidimensional analysis." in PURNELLE, G. & al. (éds.), *Actes des 7es Journées internationales d'analyse statistique des données textuelles (JADT 2004)*. Louvain : Presses universitaires de Louvain. pp. 15-34.
- BIBER, D., CONRAD, S. & REPPEN, R.** (2004 [1998]) *Corpus linguistics. Investigating language structure and use*. Cambridge : Cambridge University Press.
- BIBER, D., JOHANSON, S., LEECH, G., et al.** (2000) *The Longman grammar of spoken and written English*. London : Longman.
- BILGER, M. (Dir.)** (2000) *Linguistique sur corpus : études et réflexions*. Perpignan : Presses Universitaires de Perpignan.
- BILGER, M.D.** (2000) *Corpus : méthodologie et applications linguistiques*. Paris : Champion-Slatkine.
- BLOOMFIELD, L.** (1914) *Introduction to the Study of Language*. New-York : Holt.

- BOLLACK, M.** (2003) *Sens contre sens*. Genouilleux-Paris : Ed. La passe du vent.
- BONNAFOUS, S.** (1983) "Bilan critique d'une recherche en analyse du discours." in *Mots*, n°6 : "L'Oeuvre de R-L.Wagner. Vocabulaire et idéologie. Analyses automatiques ". Lyon : ENS Editions. pp. 135-157.
- BONNAFOUS, S.** (1991) *L'immigration prise aux mots*. Paris : Kimé.
- BONNAFOUS, S.** (1999) "Les "immigrés" sont-ils (in)égaux? Étude de presse (1950-1995) " in FIALA, P. (Dir.), *In/égalité/s*. Paris : L'Harmattan, pp. 223-239.
- BONNAFOUS, S. & TOURNIER, M.** (1995) "Analyse du discours, lexicométrie, communication et politique." in *Langages*, n°117. Paris : Larousse. pp. 67-81.
- BONVILLE, J. (de)** (1995) *Les quotidiens montréalais de 1945 à 1985 : morphologie et contenu*. Québec : Institut québécois de recherche sur la culture.
- BONVILLE, J. (de)** (2008) "Le titre des nouvelles locales dans la presse québécoise à la fin du XIXe siècle". in *Semen*, n°25: "Le discours de presse au XIXe siècle : pratiques socio-discursives émergentes". Besançon : Presses Universitaires de Franche-Comté. pp.31-49.
- BOUQUET, S.** (1997) *Introduction à la lecture de Saussure*. Paris : Payot-Rivages.
- BOUQUET, S.** (1998) "Linguistique textuelle, jeux de langage et sémantique du genre." in *Langages*, n°129. Paris : Larousse. pp. 112-123.
- BOURGIN, G.** (1934) "Essai sur l'histoire de la presse française." in *Bulletin du Comité international des sciences historiques*, n°22. pp. 26-70.
- BOZZI, A.** (1999) "Pour un système de philologie numérique." in *Document numérique*, n°3 : "Les documents anciens". Paris : Lavoisier. pp. 93-103.
- BRANCA-ROSOFF, S.** (1988) "Les mots de parti pris. *Citoyen, aristocrate et insurrection* dans quelques dictionnaires (1762-1798)." in *Dictionnaire des usages socio-politiques (1770-1815)*. Fasc.3. Paris : Klincksieck, pp. 47-73.
- BRANCA-ROSOFF, S. (Dir.)** (1998) *Le mot : analyse de discours et sciences sociales*. Aix-en-Provence : Publications de l'Université de Provence.
- BRANCA-ROSOFF, S.** (1998) "Le mot comme notion hétérogène. Linguistique-histoire-discours." in BRANCA-ROSOFF, S. (Dir.), *Le Mot. Analyse de discours et sciences sociales*. Aix-en-Provence : Université de Provence, pp. 1-50.
- BRANCA-ROSOFF, S.** (1999a) "Des innovations et des fonctionnements de langue rapportés à des genres." in *Langage & Société*, n°87. Paris : Maison des Sciences de l'Homme. pp. 115-129.

- BRANCA-ROSOFF, S.** (1999b) "Types, modes et genres : entre langue et discours." in *Langage & Société*, n°87. Paris : Maison des Sciences de l'Homme. pp. 5-24.
- BRANCA-ROSOFF, S.** (2001) "La sémantique lexicale du mot 'quartier' à l'épreuve du corpus Frantext (XIIIe-XXe siècles)." in *Langage & Société*, n°96. Paris : Maison des Sciences de l'Homme. pp. 45-70.
- BRANCA-ROSOFF, S. (Dir.)** (2007) *Langage & Société*. Paris : Editions Maison des Sciences de l'homme.
- BRANCA-ROSOFF, S.** (2007) "Sociolinguistique historique et analyse de discours." in *Langage & Société*, n°121-122 : "La revue a 30 ans. Réflexions et perspectives de recherches". Paris : Maison des Sciences de l'Homme.
- BRANCA-ROSOFF, S. & SCHNEIDER, N.** (1994) *L'écriture des citoyens. Une analyse linguistique de l'écriture des peu-lettrés pendant la Révolution française*. Paris : Klincksieck.
- BRANCA-ROSOFF, S., COLLINOT, A., GUILHAUMOU, J., et al.** (1995) "Question d'histoire et de sens." in *Langages*, n°117. Paris : Larousse. pp. 54-66.
- BRANCA-ROSOFF, S. & GUILHAUMOU, J.** (1998) "De 'société' à 'socialisme' : l'invention néologique et son contexte discursif. Essai de colinguisme appliqué." in *Langage & Société*, n°83/84. Paris : Larousse. pp.39-77.
- BRONCKART, J.-P.** (1996) "Genres de textes, types de discours et opérations discursives." in *Enjeux*, vol. 37/38. Namur : Centre d'études et de documentation pour l'enseignement du français. pp. 31-47.
- BRONCKART, J.-P.** (1997) *Activité langagière, textes et discours. Pour un interactionnisme socio-discursif*. Neuchâtel-Paris : Delachaux et Niestlé.
- BROUCKER, J. (de)** (1995) *Pratique de l'information et écritures journalistiques*. Paris : Presses Universitaires de France.
- BRUNET, E.** (1981) *Le vocabulaire français de 1789 à nos jours, d'après les données du Trésor de la langue française*. Genève-Paris : Slatkine-Champion.
- BRUNET, E.** (1983) *Le vocabulaire de Marcel Proust*. Genève-Paris : Slatkine-Champion.
- BRUNET, E.** (1985) *Le vocabulaire de Zola*. Paris-Genève : Champion-Slatkine.
- BRUNET, E. (Dir.)** (1986) *Actes du Colloque international CNRS "Méthodes quantitatives et informatiques dans l'étude des textes" (en hommage à Charles Muller)*. Genève : Slatkine.

- BRUNET, E.** (1988a) "Une mesure de la distance intertextuelle : la connexion lexicale." in *Revue, Informatique et statistique dans les sciences humaines*, n°24 : "Le nombre et le texte. Hommage à Etienne Evrard". Liège : Centre informatique de Philosophie et Lettres. pp. 81-116.
- BRUNET, E.** (1988b) *Le vocabulaire de Victor Hugo*. Champion : Slatkine.
- BRUNET, E.** (2000) "Qui lemmatise, dilemme attise." in *Lexicometrica*, n°2. Disponible sur <<http://www.cavi.univparis3.fr/lexicometrica/article/numero2/brunet2000.html>> (Consulté le 14/03/2009).
- BRUNET, E.** (2002) "Le lemme comme on l'aime." in MORIN, A., SEBILLOT, P. (éds.) *Actes des 6es Journées internationales d'analyse statistique des données textuelles (JADT 2002)*, Saint-Malo : Irisa. pp. 221-232.
- BRUNET, E.** (2006) "Navigations dans les rafales." in VIPREY, J.-M. (éds.), *Actes des 8es Journées internationales d'Analyse statistique des Données Textuelles (JADT 2006)*. Besançon : Presses Universitaires de Franche-Comté. pp.15-29.
- BURESI, C. & CEDELLE-JOUBERT, L.** (2002) *Conduire un projet de numérisation*. Paris : Editions Tec et Doc.
- BURGER, M.** (2004) "La gestion des activités. Pratiques sociales, rôles interactionnels et actes de discours." in *Cahiers de linguistique française*, n° 26. pp. 177-196.
- BURNARD, L.** (1995) "What is SGML and how does it help?" in *Computers and the Humanities*, n°29. Boston : Kluwer Academic Publishers. pp. 41-50.
- BURNARD, L.** (2007) "TEI + XSL : un mariage fait aux cieus : comment utiliser aujourd'hui la TEI pour numériser les documents en LSH." in MALRIEU, D. (Dir.), *Corpora et questionnements du littéraire. Revue Linx numéro spécial*. Paris : Presses Universitaires de Nanterre.
- BURNARD, L.** (2008) "Une introduction au British National Corpus dans son édition XML." in *Texte et corpus*, n° 3 : "Actes des Journées de la linguistique de Corpus 2007", pp. 17-34. Disponible sur <http://web.univ-ubs.fr/corpus/jlc5/ACTES/ACTES_JLC07_burnard.pdf> (Consulté le 17/03/2009).
- BURNARD, L. & C.M, S.-M.** (1996) "La TEI simplifiée : une introduction au codage des textes électroniques en vue de leur échange." in *Cahiers GUTenberg*, n°24 : "TEI : Text Encoding Initiative". pp. 23-151. Disponible sur <<http://www.gutenberg.eu.org/publications/cahiers/50-cahiers24.html>> (Consulté le 17/03/2009).
- BURNARD, L. & RAHTZ, S.** (2005) "One Document Does it all." Non publié. Disponible sur <<http://www.tei-c.org/Talks/2005/Wuerzburg/odds.pdf>> (Consulté le 12/10/2009).

- BURNARD, L. & ROBINSON, P.** (1999) "Vers un standard européen de description des manuscrits : le projet Master." in *Document numérique*, n°3 : "Les documents anciens". Paris : Lavoisier. pp. 151-169.
- CANVAT, K.** (1996) "Types de textes et genres textuels. Problématiques et enjeux." in *Enjeux*, vol. 37/38. Namur : Centre d'études et de documentation pour l'enseignement du français. pp. 49-75.
- CAPRIOLI, E., CHABIN, M.-A. & RIETSCH, J.-M. (Dir.)** (2006) *Dématérialisation et archivage électronique*. Paris : Dunod.
- CERTEAU, M. (de)** (1990) *L'invention du quotidien*. Paris : Folio essais.
- CHABIN, M.-A.** (1999) *Je pense donc j'archive. L'archive dans la société de l'information, essai*. Paris : L'Harmattan.
- CHABIN, M.-A.** (2007) *Archiver, et après?* Paris : Djakarta.
- CHALABY, J.** (1996) "Journalism as an Anglo-American Invention. A comparison of the Development of French and Anglo-American Journalism." in *European Journal of Communication*, 11, 3. Londres : Sage Publishers. pp. 303-326.
- CHALABY, J.** (1998) *The Invention of Journalism*. Baginsoke : McMillan Press.
- CHAMBURE (DE), A.** (1914) *A travers la presse*. Paris : Fort, Albouy et Cie.
- CHARAUDEAU, P.** (1983) *Langage et discours : éléments de sémiolinguistique : théorie et pratique*. Paris : Hachette.
- CHARAUDEAU, P.** (1992) *Grammaire du sens et de l'expression*. Paris : Hachette.
- CHARAUDEAU, P.** (1997) *Le discours d'information médiatique : la construction du miroir social*. Paris : Nathan.
- CHARAUDEAU, P.** (2005a) *Le discours politique*. Paris : Vuibert.
- CHARAUDEAU, P.** (2005b) Les médias et l'information : l'impossible transparence du discours. Paris : De Boeck.
- CHARAUDEAU, P.** (2006) "Discours journalistique et positionnements énonciatifs. Frontières et dérives." in *Semen*, n°22 : "Énonciation et responsabilité dans les médias". Besançon : Presses Universitaires de Franche-Comté. pp. 29-44.
- CHARAUDEAU, P. & MAINGUENEAU, D. (Dir.)** (2002) *Dictionnaire d'analyse de discours*. Paris : Seuil.
- CHARLE, C.** (2004) *Le siècle de la presse (1830-1939)*. Paris : Seuil.
- CHARLES, M.** (1995) *Introduction à l'étude des textes*. Paris : Seuil.
- CHAROLLES, M.** (1987) "Spécificité et portée des prises en charge en 'Selon A'." in *Pensée naturelle : logique et langage. Hommage à J.-B. Grize*. Neuchâtel : Université de Neuchâtel, pp. 242-269.

- CHAROLLES, M.** (1997) "L'encadrement du discours : univers, champs, domaines et espaces." in *Cahier de recherche linguistique* n°6. Paris : Ophrys. pp.1-73.
- CHARRON, J. & BONVILLE (DE), J.** (1996) "Le paradigme du journalisme de communication." in *Communication* Vol. 17, 2. Paris : Centre d'Etudes Transdisciplinaires Sociologie, Anthropologie, Histoire (EHESS). pp. 51-98.
- CHARTIER, R.** (1990) *Les origines culturelles de la Révolution française*. Paris : Seuil.
- CHARTIER, R.** (1998) *Au bord de la falaise. L'histoire entre certitudes et inquiétude*. Paris : Albin Michel.
- CHARTIER, R., CAVALLO, G. (Dir.)** (2001a) *Histoire de la lecture dans le monde occidental*. Paris : Seuil.
- CHARTIER, R., CAVALLO, G.** (2001b) " Introduction." In CHARTIER, R., CAVALLO, G. (Dir.) *Histoire de la lecture dans le monde occidental*. Paris : Seuil. pp.7-49.
- CHISS, J.-L.** (1987) "Malaise dans la classification." in *Langue Française*, n°74. Paris : Larousse. pp. 10-28.
- CHOLLET, R.** (1983) *Balzac journaliste. Le tournant de 1830*. Paris : Klincksieck.
- CHURCH, K. & HANKS, P.** (1989) "Introduction to the special issue on computational linguistics using large corpora." in *Computational Linguistics*, 19, 1. Cambridge (MA) : MIT Press. pp.1-24.
- CIBOIS, P.** (1983) *L'analyse factorielle*. Paris : PUF. Coll : Que sais-je ?
- CIBOIS, P.** (1994) *L'analyse factorielle : analyse en composantes principales et analyse des correspondantes*. Paris : Hachette.
- CONDAMINES, A.** (2005) *Sémantique et corpus*. Paris : Hermès (Lavoisier).
- COSSUTA, F.** (2004) "Catégories descriptives et catégories interprétatives en analyse du discours." in ADAM, J.-M., GRIZE, J.-B., et al. (Dir.), *Texte et discours : catégories pour l'analyse*. Dijon : Editions universitaires de Dijon, pp. 189-213.
- COTTE, D.** (2004) "Le concept de 'document numérique'." in *Communication & langages*, n°140. Paris : NecPlus. pp. 31-43.
- COTTERET, J.-M. & MOREAU, R.** (1969) *Recherches sur le vocabulaire du Général de Gaulle. Analyse statistique des allocutions radiodiffusées, 1958-1965*. Paris : Armand Colin.
- COURDESSES, L.** (1971) "Blum et Thorez en mai 1936 : Analyse d'énoncé." in MARCELLESI, J.B. (Dir.), *Langue Française*, n°9 : "Linguistique et Société". Paris : Larousse. pp. 22-34.

- COURTINE, J.-J. & MARANDIN, J.-M.** (1981) "Quel objet pour l'analyse de discours?" in CONEIN, B., COURTINE, J.-J., *et al.* (**Dir.**), *Matérialités discursives*. Lille : Presses Universitaires de Lille.
- COUZINET, V.** (2004) "Le document : leçon d'histoire, leçon de méthode." in *Communication & langages*, n°140. Paris : NecPlus. pp. 19-30.
- DAOUST, F., DOBROWOLSKI, G., DUFRESNE, M., et al.**, *Système d'Analyse de Texte par Ordinateur (SATO) version 4.3 (mai 2008)*, 2006. Centre d'analyse de texte par ordinateur, UQAM. Disponible sur <<http://www.ling.uqam.ca/sato/satoman-fr.html>> (Consulté le 12/06/2008).
- DAOUST, F., DOBROWOLSKI, G., DUFRESNE, M., et al.** (2006) "Analyse exploratoire d'entrevues de groupe : quand ALCESTE, DTM, LEXICO et SATO se donnent la main." in VIPREY, J.-M. (éds.), *Actes des 8es Journées internationales d'Analyse statistique des Données Textuelles (JADT 2006)*. Besançon : Presses Universitaires de Franche-Comté. pp. 313-326.
- DAOUST, F. & MARCOUX, Y.** (2006) "Logiciels d'analyse textuelle : vers un format XML-TEI pour l'échange de corpus annotés." in VIPREY, J.-M. (éds), *Actes des 8es Journées internationales d'Analyse statistique des Données Textuelles (JADT 2006)*. Besançon : Presses Universitaires de Franche-Comté. pp. 327-240.
- DAOUST, F. & DUCHASTEL, J.** (2007) "Pluralisme et pluralité des interprétations." in BROSSAUD, C. & REBER, B. (**Dir.**), *Sciences Humaines et technologies de l'information et de la communication*. Paris : Hermès Sciences, pp. 257-268.
- DAOUST, F., DUCHASTEL, J., MARCOUX, Y., et al.** (2008) "Pour un modèle de dépôt de données adapté à la constitution de corpus de recherche." in HEIDEN, S & PINCEMIN, B. (éds.), *Actes des 9es Journées internationales d'Analyse statistique des Données Textuelles (JADT 2008)*. Lyon : Presses Universitaires de Lyon. pp. 355-367.
- DAVALLON, J.** (2004) "Objet concret, objet scientifique, objet de recherche." in JEANNERET, Y. & OLLIVIER, B. (**Dir.**), *Les sciences de l'information et de la communication*. Paris : CNRS éditions, pp. 30-37.
- DELAUNAY, E.** (2007) "Conservation et sauvegarde de la presse ancienne." in DIDELOT, M. (**Dir.**), *Journaux en Lorraine*. Nancy : Presses de l'Université de Nancy 2. pp. 103-110.
- DELEPLACE, M.** (1989) "La notion d'anarchie pendant la Révolution française. Un parcours méthodologique en Analyse de discours." in *Revue d'histoire moderne et contemporaine*, Vol. 43, 2. Paris : Belin. pp. 307-328.

- DELEPLACE, M.** (2000) *L'Anarchie de Mably à Proudhon (1750-1850)*. Lyon : ENS éditions.
- DELORME, G.** (2004) *Ouest-France : histoire du premier quotidien français*. Rennes : Editions Apogée.
- DELPORTE, C.** (1999) *Les journalistes en France. 1880-1950. Naissance et construction d'une profession*. Paris : Seuil.
- DEMONET, M., GEFFROY, A., TOURNIER, M. et al.** (1975 [1978]) *Des tracts en mai 68. Mesures de vocabulaire et de contenu*. Paris : Presses de la Fondation nationale des sciences politiques.
- DENDALE, P.** (1993) *Le conditionnel de l'information incertaine : marqueur modal ou marqueur évidentiel?*, vol. 1. Zürich : A.Francke Verlag. pp. 163-176.
- DENDALE, P. & COLTIER, D.** (2004) "Discours rapporté et évidentialité : comparaison du conditionnel et des constructions en selon N." in LOPEZ MUNOZ, J.-M., MARNETTE, S., et al. (**Dir.**), *Le discours rapporté dans tous ses états*. Bruxelles : L'Harmattan. pp. 587-597.
- DEREYMEZ, J.-W.** (1999) "Un organe socialiste de province : le Droit du peuple (fin XIXe-début XXe siècle)." in GRANGE, D.J. (**Dir.**), *La presse régionale XIXe-XXe siècles*. Grenoble : Presses de l'Université Pierre Mendès France, pp. 135-154.
- DES GRANGES, C.-M.** (1973 [1907]) *Le Romantisme et la critique. La presse littéraire sous la Restauration (1815-1830)*. Genève : Slatkine reprints.
- DÉVELOPTE, C.** (2000) "Titres de presse et représentations d'autrui en temps de crise." in *Mots*, 64. Lyon : ENS Editions. pp. 23-38.
- DIDELOT, M. (Dir.)** (2007) *Journaux en Lorraine*. Nancy : Presses de l'Université de Nancy 2.
- DISEGNI, S.** (1997) *Jules Vallès, du journalisme au roman autobiographique*. Paris : L'Harmattan.
- DORIGNY, M.** (1985) "Honnêtes-gens. L'expression dans la presse girondine, juin-septembre 1792." in *Dictionnaire des usages socio-politiques (1770-1815)*. Fasc. 1 : "Désignants socio-politiques." Paris : Champion. pp. 79-92.
- DOSSE, F.** (2003) *La marche des idées. Histoire des intellectuelles – Histoire intellectuelle*. Paris : La Découverte.
- DUBIED, A.** (1999) *Le faits divers*. Paris : Presses Universitaires de France, coll. Que sais-je?
- DUBIED, A.** (1999) "Sur la notion de genre médiatique." in *Médiatiques*, n°17. Louvain-la-Neuve : Observatoire du récit médiatique (ORM). pp. 42-44.

- DUBIED, A.** (2001) "Invasion péritextuelle et contaminations médiatiques. Le fait divers, une catégorie ancrée dans le champ journalistique." in *Semen*, n° 13. Besançon : Presses Universitaires de Franche-Comté. pp. 49-64.
- DUBIED, A. & LITS, M.** (1997) "L'éditorial : genre journalistique ou position discursive?" in *Pratiques*, n°94. Metz : CRESEF. pp. 49-61.
- DUBOIS, J.** (1962) *Le vocabulaire politique et social en France de 1869 à 1872*. Paris : Larousse.
- DUBOIS, J.** (1969) *Grammaire structurale du français: la phrase et ses transformations*. Paris: Larousse.
- DUCHASTEL, J. & ARMONY, V.** (1993) "Un protocole de description de discours politique." in ANASTEX, S.J. (éds) *Actes des Secondes journées internationales d'analyse statistique de données textuelles (JADT 1993)*. Paris : Telecom, pp. 159-183.
- DUCHASTEL, J. & BOURQUE, G.** (1996) *L'identité fragmentée. Nation et citoyenneté dans les débats constitutionnels canadiens, 1941-1992*. Montréal : Fides.
- DUCHASTEL, J., DAOUST, F. & DELLA FAILLE, D.** (2004) "SATO-XML : une plateforme Internet ouverte pour l'analyse de texte assistée par ordinateur" in PURNELLE, G. & al. (éds.) *Actes des 7es Journées internationales d'analyse statistique des données textuelles (JADT 2004)*. Louvain : Presses universitaires de Louvain. pp. 353-363.
- DURRER, S.** (2001) "De quelques affinités génériques du billet." in *Semen*, n°13. Besançon : Presses Universitaires de Franche-Comté. pp. 163-185.
- ECO, U.** (1989) *Lector in fabula : le rôle du lecteur ou la Coopération interprétative dans les textes narratifs*. Paris : Le Livre de Poche.
- ECO, U.** (1990) *Les limites de l'interprétation*. Paris : Grasset.
- ELUERD, R.** (2000) *La lexicologie*. Paris : Presses Universitaires de France, Coll. "Que sais-je?"
- EVRARD, F.** (1997) *Faits divers et littérature*. Paris : Nathan Université.
- EWALD, F.** (1992) "La politique sociale des opportunistes, 1879-1885." in BERSTEIN, S. & RUDELLE, O. (**Dirs.**), *Le modèle républicain*. Paris : Presses Universitaires de France, pp. 173-188.
- Équipe "XVIII^e et Révolution"** (1985-2006) *Dictionnaire des usages socio-politiques (1770-1815)*. 8 volumes. Paris : Champion.
- FAYE, J.-P.** (1972) *Langages totalitaires*. Paris : Hermann.
- FEBVRE, L.** (1953) *Combats pour l'histoire*. Paris : Armand Colin.
- FERENCZI, T.** (1993) *L'invention du journalisme en France*. Paris : Plon.

- FERNANDEZ SEBASTIAN, J.** (2002) "Historia de los conceptos. Nuevas perspectivas para et estudio de los lenguajes politicos europeos." in *Ayer*, 48. Madrid : Asociación de Historia Contemporanea/Marcial Pons.
- FIALA, P.** (1994) "L'interprétation en lexicométrie." in *Langue Française*, n°103. Paris : Larousse. pp. 113-122.
- FIALA, P. (Dir.)** (1999) *In/égalité/s. Usages lexicaux et variations discursives (XVIIIe-XXe siècles)*. Paris : L'Harmattan.
- FIALA, P.** (2002) "Figement", "Formule", "Langue de bois", "Phraséologie", "Sloganisation". in CHARAUDEAU, MAINGUENEAU (**Dirs.**), *Dictionnaire d'analyse du discours*. Paris : Seuil.
- FIALA, P.** (2006a) "L'analyse de discours, mesures à l'appui." in *Semen* n°21 : "Catégories pour l'analyse du discours politique" Besançon : Presses Universitaires de Franche-Comté. pp. 155-165.
- FIALA, P. (Dir.)** (2006b) *Mots*. n°81 : "Suisse, laboratoire politique européen ?" Paris : ENS Editions. Disponible sur <[http ://mots.revues.org/index70.html](http://mots.revues.org/index70.html)> (Consulté le 12/10/2009).
- FIALA, P. & EBEL, M.** (1983) *Sous le consensus, la xénophobie. Paroles, arguments et contextes (1961-1981)*. Lausanne : Institut de Science Politique.
- FIALA, P., HABERT, B., LAFON, P., et al.** (1987) "Des mots aux syntagmes. Figements et variations dans la résolution générale du congrès de la CGT de 1978." in *Mots*, n°14. Lyon : ENS Editions. pp. 47-87.
- FIALA, P. & HABERT, B. (1989)** "La langue de bois en éclats : les défigements dans les titres de presse." in *Mots*, n°21. Lyon : ENS Editions. pp. 83-99.
- FIALA, P., LAFON, P. & PIGUET, M.-F. (Dirs.)** (1997) *La locution. Entre lexique, syntaxe et pragmatique. Identification en corpus, traitement, apprentissage*. Paris : Klincksieck.
- FOHLEN, C. (Dir.)** (1982) *Histoire de Besançon de la conquête à nos jours*. Vol II. Paris : Nouvelle Librairie de France.
- FONSEGRIVE, G.** (1903) *Comment on lit les journaux*. Paris : V. Lecoffre.
- FONTANIER, P.** (1977) *Les figures du discours*. Paris : Flammarion.
- FOUCAULT, M.** (1966) *Les Mots et les choses : Une archéologie des sciences humaines*. Paris : Gallimard.
- FOUCAULT, M.** (1969) *L'Archéologie du savoir*. Paris : Gallimard.
- FOUCAULT, M.** (1994) *Dits et écrits*. Vol. II (1970-1975). Paris : Gallimard.
- FOURMENT, A.** (1987) *Histoire de la presse des jeunes et des journaux d'enfants (1768-1988)*. Paris : Eole.

- FRANDSEN, F.** (1990) "News Discourse: The Paratextual Structure of News Texts." in *Discourse Structuring. Production and Reception Strategies*, Proceedings from Nord Text Symposium 10-13 may 1990. Åbo : Academic Press.
- FROBERT, L.** (2004) "Le thème de la justice sociale dans l'Echo de la Fabrique." Communication présentée lors de « *Histoire des théories économiques de la justice* », Journées d'étude de l'Association Charles Gide pour l'Etude de la Pensée Economique, Boulogne sur Mer, Université du Littoral Côte d'Opale, 17-18 juin 2004.
- FUCHS, C.** (1996) *Les ambiguïtés du français*. Paris : Ophrys.
- FURET, C.** (1995) *Le titre. Pour donner envie de lire*. Paris : CFPJ.
- FURETIÈRE, A.** (1762 [1727]) *Dictionnaire de l'Académie*. Paris : Académie française.
- FURMAN, N.** (1975) *La Revue des Deux-Mondes et le romantisme (1831-1848)*. Genève : Droz.
- GADAMER, H.-G.** (1991) *L'Art de comprendre*. Paris : Aubier.
- GARDIN, J.-C.** (1971) *Les analyses du discours*. Neuchâtel : Delachaux & Niestlé.
- GEFFROY, A., ARNOLD, N., DOUGNAC, F., et al.** (1995) *Langages de la Révolution (1770-1815)*. Paris : Klincksieck.
- GEFFROY, A., GUILHAUMOU, J. & SALEM, A.** (1975) "L'Histoire sur mesures ou pour une statistique du discours." in *Bulletin du Centre d'analyse du discours de l'université de Lille III*, n°2. Lille : Presses Universitaires de Lille. pp.16-57.
- GENETTE, G.** (1979) *Introduction à l'architexte*. Paris : Seuil.
- GIOVANNONI, A. & GUILHAUMOU, J. (Dir.)** (2008) *Histoire et subjectivation*. Paris : Kimé.
- GIRARDIN, S.-M.** (1859) *Souvenirs et réflexions politiques d'un journaliste*. Paris : Michel Lévy.
- GOGUEL, F.** (1958) "Préface." in KAYSER, J. (Dir.), *La presse de province sous la III^e République*. Paris : Armand Colin.
- GOODY, J.** (1979) *La Raison graphique*. Paris : Ed. de Minuit.
- GOODY, J.** (1986) *La logique de l'écriture. Aux origines des sociétés humaines*. Paris : Armand Colin.
- GRANDMAISON, H. (de)** (1973) *La province trahie*. Les Sables-D'Olonne : Ed. Le Cercle d'Or.
- GRANGE, D.J. (Dir.)** (1999) *La presse régionale XIXe-XXe siècles*. Grenoble : Presses de l'Université Pierre Mendès France.

- GRÉBILLON, A. & MAINGUENEAU, D.** (1984) "Polyphonie, proverbe et détournement." in *Langages*, n°73. Paris : Larousse. pp. 112-125.
- GRÉVY, J.** (1998) *La République des opportunistes, 1870-1885*. Paris : Perrin.
- GROSS, M.** (1975) *Méthodes en syntaxe*. Paris : Hermann.
- GROSSE, E.U. & SEIBOLD, E.** (1996) *Panorama de la presse ordinaire*. Frankfurt : Peter Lang.
- GUÉRY, L.** (1992) *La Presse régionale et locale*. Paris : Editions du CFPJ.
- GUÉRY, L.** (1997) *Visages de la presse. La présentation des journaux des origines à nos jours*. Paris : Editions CFPJ.
- GUESPIN, L.** (1971) "Problématique des travaux sur le discours politique." in *Langages*, n°23. Paris : Larousse. pp. 3-24.
- GUESPIN, L.** (1975) *L'analyse du discours; problèmes et perspectives*. Paris : Ed. de la Nouvelle critique.
- GUILHAUMOU, J.** (1974) "L'idéologie du Père Duschene : les forces adjuvantes (14 juillet-6 septembre 1793)." in **GUILHAUMOU, J., ROBIN, D., et al. (Dir.)**, *Langage et idéologies. Le discours comme objet de l'histoire*. Paris : Editions ouvrières. pp. 81-116.
- GUILHAUMOU, J.** (1986) "L'historien du discours et la lexicométrie. Etude d'une série chronologique : le "Père Duschene" d'Hébert (juillet 1793-mars 1794)." in *Histoire & Mesure*, I, 3/4. Paris : Éditions de l'EHESS. pp. 26-46.
- GUILHAUMOU, J.** (1987) "La Terreur à l'ordre du jour (juillet 1793-1794)." in *Dictionnaire des usages socio-politiques (1770-1815)*. fasc.2. Paris : Klincksieck. pp. 127-161.
- GUILHAUMOU, J.** (1989) *La langue politique et la Révolution française*. Paris : Méridiens-Klincksieck.
- GUILHAUMOU, J.** (1992) *Marseille républicaine (1791-1793)*. Paris : Presses de la Fondation nationale des sciences politiques.
- GUILHAUMOU, J.** (1993) "A propos de l'analyse de discours : les historiens et le 'tournant linguistique'." in *Langage & Société*, n°65. Paris : Maison des Sciences de l'Homme. pp. 5-38.
- GUILHAUMOU, J.** (1996) "Vers une histoire des événements linguistiques. Un nouveau protocole d'accord entre l'historien et le linguiste." in *Histoire Epistémologie Langage*, Vol 18, Fasc. 2. Paris : SHESL. pp.103-126.
- GUILHAUMOU, J.** (1997) "Nation, individu et société chez Sieyès." in *Genèses*, n°26 (Avril 1996). Paris : Belin. pp. 4-24.
- GUILHAUMOU, J.** (1998) *La parole des Sans. Les mouvements actuels à l'épreuve de la Révolution française*. Saint-Cloud : ENS éditions.

- GUILHAUMOU, J.** (2000) "De l'histoire des concepts à l'histoire linguistique des usages conceptuels." in *Genèses*, n°38. Paris : Belin. pp. 105-118.
- GUILHAUMOU, J.** (2001) "L'histoire des concepts : le contexte historique en débat." in *Annales. Histoire, sciences sociales*, mai-juin 2001. Paris : Armand Colin. pp. 685-698.
- GUILHAUMOU, J.** (2002a) "Le corpus en analyse de discours : Perspective historique." in *Corpus*, n°1, pp. 21-50. Disponible sur <<http://corpus.revues.org/index8.html>> (Consulté le 12/10/2009).
- GUILHAUMOU, J.** (2002b) "L'histoire linguistique des usages conceptuels à l'épreuve des événements linguistiques." in BÖDEKER, H.E. (**Dir.**), *Begriffsgeschichte, Diskurgeschichte, Metapherngeschichte*. Göttingen : Wallstein Verlag, pp. 123-158.
- GUILHAUMOU, J.** (2002c) "Histoire/discours", "configuration/archive", "trajet thématique", "événement discursif", "événement linguistique" in CHARAUDEAU, MAINGUENEAU (**dirs.**), *Dictionnaire d'analyse du discours*. Paris : Seuil.
- GUILHAUMOU, J.** (2004) "L'avènement de la 'métaphysique politique'. Sieyès et le nominalisme politique." in GUILHAUMOU, J. & KAUFMANN, L. (**Dirs.**), *L'invention de la société. Nominalisme politique et science sociale au 18ème siècle*. Raisons pratiques. Paris : Editions de l'EHESS, pp. 201-226.
- GUILHAUMOU, J.** (2005) "Où va l'analyse de discours ? Autour de la notion de formation discursive." in MAINGUENEAU, D. (**Dir.**), *Marges linguistiques*, n°9 (Mai 2005) : "Analyse du discours. État de l'art et perspectives", Saint-Chamas, pp. 95-114. Disponible sur <<http://www.revue-texto.net/marges/Marges.html>> (Consulté le 12/10/2009).
- GUILHAUMOU, J.** (2006a) *Discours et événement. L'histoire langagière des concepts*. Besançon : Presses Universitaires de Franche-Comté.
- GUILHAUMOU, J. (Dir.)** (2006b) *Dictionnaire des usages socio-politiques (1770-1815) : Patrie, Patriotisme*. Fascicule 8. Paris : Champion.
- GUILHAUMOU, J.** (2007) "L'analyse du discours du côté de l'histoire. Une démarche interprétative." in *Langage & Société*, n°121-122. Paris : Maison des Sciences de l'Homme. pp. 177-188.
- GUILHAUMOU, J.** (2008a) "Introduction." in GIOVANNONI, A. & GUILHAUMOU, J. (**Dirs.**), *Histoire et subjectivation*. Paris : Kimé, pp. 7-10.
- GUILHAUMOU, J.** (2008b) "La temporalité historique des formes d'individuation. Les figures du moi." in GIOVANNONI, A. & GUILHAUMOU, J. (**Dirs.**), *Histoire et subjectivation*. Paris : Kimé, pp. 219-252.

- GUILHAUMOU, J.** (2008c) "De peuple à prolétaire (s) : Antoine Vidal, porte-parole des ouvriers dans L'Echo de la Fabrique en 1831-1832." in *Semen*, n°25 : "Le discours de presse au XIXe siècle : pratiques socio-discursives émergentes". Besançon : Presses Universitaires de Franche-Comté. pp.101-115.
- GUILHAUMOU, J., MALDIDIER, D., PROST, A., et al.** (1974) *Langage et idéologies. Le discours comme objet de l'histoire*. Paris : Editions ouvrières.
- GUILHAUMOU, J., MALDIDIER, D. & ROBIN, R.** (1989) "Jalons dans l'histoire de l'analyse de discours en France : un trajet des historiens du discours." in *Discours social/Social discourse*, Vol. II, 3. Montréal : Université McGill.
- GUILHAUMOU, J., MALDIDIER, D. & ROBIN, R.** (1994) *Discours et archive. Expérimentations en analyse de discours*. Liège : Mardaga.
- GUILHAUMOU, J. & MONNIER, R. (Dirs.)** (2003) *Des notions concepts en révolution. Autour de la liberté politique à la fin du 18ème siècle*. Etudes révolutionnaires, Vol. 4. Paris : Société des études robespierristes.
- GUILHAUMOU, J. & KAUFMANN, L. (Dirs.)** (2004) *L'invention de la société. Nominalisme politique et science sociale au 18ème siècle*. Vol. 4, Raisons pratiques. Paris : Editions de l'EHESS.
- GUILHAUMOU, J. & BERTRAND, O. (Dirs.)** (2005) *Le politique en usages (XIVème siècle-XIXème siècle)*. Langages et Société, n°113. Paris : Maison des Sciences de l'Homme.
- GUILHAUMOU, J. & GIOVANNONI, A. (Dirs.)** (2008) *Histoire et subjectivation*. Paris : Kimé.
- GUIRAL, P.** (1971) "Problèmes d'histoire de la presse." in *Revue d'histoire moderne et contemporaine*, Vol. 18, 4. Paris : Belin. pp. 481-488.
- GUIRAUD, P.** (1954) *Les caractères statistiques du vocabulaire*. Paris : Presses Universitaires de France.
- GUIRAUD, P.** (1960) *Problèmes et méthodes de la statistique linguistique*. Paris : Larousse.
- HABERMAS, J.** (1962) *L'espace public. Archéologie de la publicité comme dimension constitutive de la société bourgeoise*. Paris : Payot.
- HABERMAS, J.** (1978) *L'Espace public*. Paris : Payot.
- HABERT, B.** (1985) "L'analyse des formes spécifiques : bilan critique et propositions d'utilisation." in *Mots*, n°85. Lyon : ENS Editions. pp. 127-154.
- HABERT, B.** (2000) "Des corpus représentatifs : de quoi, pour quoi, comment?" in BILGER, M. (Dir.), *Linguistique sur Corpus. Etudes et réflexions (Actes de la journée Linguistique et corpus)*. Perpignan : Presses de l'université de Perpignan, pp. 11-58.

- HABERT, B.** (2005) *Instruments et ressources électroniques pour le français*. Gap/Paris : Ophrys.
- HABERT, B., FABRE, C. & ISSAC, F.** (1997) *De l'écrit au numérique : constitution, annotation et analyse de corpus*. Paris : Masson.
- HABERT, B., NAZARENKO, A. & SALEM, A.** (1997) *Les linguistiques de corpus*. Paris : Armand Colin-Masson.
- HABERT, B., ILLOUZ, G., LAFON, P. et al.** (2000) *Profilage de textes : cadre de travail et expérience*. Lausanne : Lexicométrica. Disponible sur <<http://www.cavi.univ-paris3.fr/lexicométrica/jadt/jadt2000/pdf/56/56.pdf>> (Consulté le 07/10/2009).
- HAILLET, P.** (1995) *Le conditionnel dans le discours journalistique : essai de linguistique descriptive*. Canada : Bref.
- HAILLET, P.** (1998) "Le conditionnel d'altérité énonciative et les formes du discours rapporté dans la presse écrite." in *Pratiques*, n°100. Metz : CRESEF. pp. 63-79.
- HALÉVY, D.** (1972) *La fin des notables*. Paris : Le livre de poche.
- HALLIDAY, M.A.K. & HASAN, R.** (1976) *Cohesion in English*. Londres : Longman.
- HARISMENDY, P.** (2000) "L'évolution de la maquette de l'Ouest-Eclair." in LAGRÉE, M., HARISMENDY, P., et al. (Dir.), *L'Ouest-Eclair, naissance et essor d'un grand quotidien régional*. Rennes : Presses Universitaires de Rennes, pp. 136-137.
- HAROUCHE, C., HENRY, P. & PÊCHEUX, M.** (1971) "La sémantique et la coupure saussurienne : langue, parole, discours." in *Langages*, n°24. Paris : Larousse. pp. 93-106.
- HARRIS, Z.** (1969) "Analyse du discours." in *Langages*, n°61. Paris : Larousse. pp. 8-45.
- HATIN, E.** (1853) *Histoire du journal en France, 1631-1853*. Paris : P. Jannet.
- HEIDEN, S. & LAFON, P.** (1998) "Cooccurrences, La CFDT de 1973 à 1992." in *Des mots en liberté, Mélanges Maurice Tournier*. Vol. 1. Fontenay-aux-Roses : ENS Éditions. pp. 65-83.
- HEIDEN, S.** (2003) "Illustration d'une méthode lexicométrique des cooccurrences sur un corpus historique." in GUILHAUMOU, J. & MONNIER, R. (éds), *Des notions-concepts en révolution*. vol. 4 des Études révolutionnaires. Paris : Société des études robespierristes. pp. 105-122.

- HEIDEN, S.** (2004) "Interface hypertextuelle à un espace de cooccurrences : implémentation dans Weblex.", in PURNELLE, G. & al. (éds.), *Actes des 7es Journées internationales d'analyse statistique des données textuelles (JADT 2004)*. Louvain : Presses universitaires de Louvain. pp. 577-588.
- HEIDEN, S.** (2006) "Modèles de données et formats d'échange pour l'interopérabilité des outils de textométrie." in VIPREY, J.-M. (éds.), *Actes des 8es Journées internationales d'Analyse statistique des Données Textuelles (JADT 2006)*. Besançon : Presses Universitaires de Franche-Comté. pp. 487-498.
- HERMAN, T. & LUGRIN, G.** (1999a) *Formes et fonctions des rubriques dans les quotidiens romands*. Fribourg : Institut de journalisme et des communications sociales.
- HERMAN, T. & LUGRIN, G.** (1999b) "La rubrique : un outil de description du marché de la presse écrite." in *Communication & langages* n°122. Paris : NecPlus.
- HIRTZ, C.** (1973) *L'Est républicain 1889-1914*. Grenoble : Presses Universitaires de Grenoble.
- HOEK, L.-H** (1981) *La marque du titre. Dispositifs sémiotiques d'une pratique textuelle*. La Haye-Paris-New-York : Mouton.
- HUMMEL, P.** (2000) *Histoire de l'histoire de la philologie : étude d'un genre épistémologique et bibliographique*. Genève : Droz.
- ILLOUZ, G. & HABERT, B.** (1999) "Maîtriser les déluges de données hétérogènes." in CONDAMINES, A., et al. (éds). *Corpus et TAL : pour une réflexion méthodologique, 6^e Conférence annuelle sur le Traitement Automatique des Langues Naturelles (TALN 99)*. Cargèse. pp. 37-46.
- JACOUTY, J.-F. (Dir.)** (2006) *Presse, politique, culture et société du XVIIIe au XXe siècle (France-Languedoc). Mélanges en l'honneur du professeur Roland Andréani*. Montpellier : Presses universitaires de la Méditerranée.
- JACQUESSON, A. & RIVIER, A.** (1999) *Bibliothèques et documents numériques, concepts, composantes techniques et enjeux*. Paris : Electre - Cercle de la librairie.
- JAMATI, V.** (1906) *Pour devenir journaliste. Comment se rédige et s'administre un journal*. Paris : Librairie J. Victorien.
- JAMET, C. & JANNET, A.-M.** (1999) *La mise en scène de l'information*. Paris : L'Harmattan.

- JANSEN, P.** (1981) "Le dépouillement des principaux périodiques parus en 1768 et son exploitation par ordinateur." in BONNET, J.-C. (**Dir.**), *L'année 1768 à travers la presse traitée par ordinateur (Travaux du Centre d'étude du XVIIe et du XVIIIe siècle)*. Paris : Editions du CNRS.
- JANSEN, P., LECLERCLE, J.-L., BARIDON, M., et al. (Dir.)** (1982) *L'année 1778 à travers la presse traitée par ordinateur*. Paris : Presses Universitaires de France.
- JEANDILLOU, J.-F.** (1997) *L'analyse textuelle*. Paris : Armand Colin.
- JEANNENEY, J.-N.** (1996) *Une histoire des médias des origines à nos jours*. Paris : Ed. du Seuil.
- JEANNERET, Y.** (2000) *Y'a-t'il vraiment des technologies de l'information ?* Villeneuve d'Asq : Presses universitaires du septentrion.
- JEANNERET, Y.** (2003) *Le numérique, une pratique du texte*. Lille, mars 2003. Disponible sur <<http://halma-ipel.recherche.univ-lille3.fr/fichierspdf/pubenligne/2003mediation/jeanneret.pdf>> (Consulté le 12/10/2009).
- JONES, G.S.** (1983) *Languages of Class. Studies in English Working Class History (1832-1982)*. Cambridge : Cambridge University Press.
- JOST, F. & LEBLANC, G.** (1998) *La télévision française au jour le jour*. Paris : INA-Anthropos.
- JOUBE, V.** (1993) *La lecture*. 112 pages. Paris : Hachette.
- JUPER, N., HERMAN, T.** (2000) "L'éditorial 'vitrine idéologique du journal'." in *Semen*, n°13, Besançon : Presses Universitaires de Franche-Comté. pp.139-169.
- KALIFA, D.** (1995) *L'encre et le sang. Récits de crimes et société à la Belle Epoque*. Paris : Fayard.
- KALIFA, D.** (2000) "L'entrée de la France en régime médiatique : l'étape des années 1860." in MIGOZZI, J. (**Dir.**), *De l'écrit à l'écran. Littérature populaire : mutations génériques, mutations médiatiques*. Limoges : Pulim, pp. 39-51.
- KALIFA, D.** (2005) *Crime et culture au XIX^e siècle*. Paris : Editions Perrin.
- KALIFA, D. & VAILLANT, A.** (2004) "Pour une histoire culturelle et littéraire de la presse française au XIX^e siècle." in *Le Temps des Médias*, 2004/1, 2. Paris : Nouveau Monde éditions. pp. 197-214.
- KARLGREN, J. & CUTTING, D.** (1994) "Recognizing Text Genres with Simple Metrics Using Discriminant Analysis." in *Proceedings of the 15th International Conference on Computational Linguistics (COLING'94)*. vol. II. Kyoto, Japon. pp. 1071-1075. Disponible sur <<http://eprints.sics.se/56/01/cmplglixcol.pdf>> (Consulté le 07/10/2009).

- KASTBERG SJÖBLOM, M.** (2003) "Comment l'ordinateur peut-il servir dans l'étude stylistique d'un texte littéraire ?" in *Texto!* Disponible sur <<http://www.revue-texto.net/Corpus/Publications/Kastberg/Kastberg.html>> (Consulté le 12/10/2009).
- KASTBERG-SJÖBLOM, M.** (2006) *L'écriture de J.M.G Le Clézio. Des mots aux thèmes*. Paris : Honoré Champion.
- KASTBERG-SJÖBLOM, M. & ABLALI, D. (Dir.)** (2009 (à paraître)) *Actes du colloque Linguistique et littérature. Cluny, 40 ans après*. Besançon : Presses Universitaires de Franche-Comté.
- KASTBERG-SJÖBLOM, M.** (2007) "La sémantique lexicale et les genres : Analyse systématique d'un corpus québécois." in WILLIAMS, G. (Dir.), *Les 4èmes journées de la linguistique de corpus*. Disponible sur <http://web.univ-ubs.fr/corpus/jlc4/acteJLC2005_7_katsberg.pdf> (Consulté le 12/05/2007).
- KAYSER, J.** (1957) "L'historien et la presse." in *Revue historique*, 218, 2. Paris : Presses Universitaires de France. pp. 284-309.
- KAYSER, J. (Dir.)** (1958) *La presse de province sous la III^e République*. Paris : A. Colin.
- KAYSER, J. (Dir.)** (1963) *Le quotidien français*. Paris : Armand Colin.
- KENNEDY, G.** (1998) *An introduction to corpus linguistics. Studies in language and linguistics*. London : Longman.
- KERBRAT-ORECCHIONI, C.** (1977) *La Connotation*. Lyon : Presses Universitaires de Lyon.
- KERBRAT-ORECCHIONI, C.** (1980) *L'énonciation de la subjectivité dans le langage*. Paris : A. Colin.
- KESSLER, B., NUNBERG, G. & SCHÜTZE, H.** (1997) "Automatic detection of text genre." in *Proceedings of the European Chapter of the Association for Computational Linguistics (EACL'97)*. San Francisco: Morgan Kaufmann. pp. 32-38.
- KOREN, R.** (1996) *Les enjeux éthiques de l'écriture de presse et la mise en mots du terrorisme*. Paris : L'Harmattan.
- KOSELLECK, R.** (1990) *Le futur passé. Contribution à la sémantique des temps historiques*. Paris : Editions de l'Ecole des Hautes Etudes en Sciences Sociales.
- KOSELLECK, R.** (1997) *L'Expérience de l'histoire*. Paris : Gallimard/Seuil.
- KRIEG-PLANQUE, A.** (1999) "Vacance argumentative : l'usage de (sic) dans la presse d'extrême droite contemporaine." in *Mots*, n°58. Lyon : ENS Editions. pp. 11-34.

- KRIEG-PLANQUE, A.** (2000) "Analyser le discours de presse. Mises au point sur le "discours de presse" comme objet de recherche." in *Communication*, n° 20, 1. Paris : Centre d'Etudes Transdisciplinaires Sociologie, Anthropologie, Histoire (EHESS). pp. 75-97.
- KRIEG-PLANQUE, A.** (2003) *"Purification ethnique". Une formule et son histoire.* Paris : CNRS Editions.
- KRIEG-PLANQUE, A.** (2006) ""Formules" et "lieux discursifs" : propositions pour l'analyse du discours politique." in *Semen*, n°21 : "Catégories pour l'analyse du discours politique". Besançon : Presses Universitaires de Franche-Comté. pp.19-48.
- KRIEG-PLANQUE, A.** (2008) "Travailler les discours dans la pluridisciplinarité. Exemples d'une "manière de faire" en analyse du discours." in BONNAFOUS, S. & TEMMAR, M. (Dir.), *Analyse du discours et sciences humaines et sociales.* Paris : Ophrys.
- KRONNING, H.** (2005) "Polyphonie, médiation et modalisation : le cas du conditionnel épistémique." in BRES, J., MELLET, S., et al. (Dir.), *Dialogismes et polyphonie. Approches linguistiques.* Bruxelles : De Boeck Duculot, pp. 287-312.
- LA CAPRA, D. & KAPLAN, S. (Dir.)** (1982) *Modern European Intellectual History : Reappraisals and New Perspectives.* Ithaca : Cornell University Press.
- LABROSSE, C. & RÉTAT, P. (Dir.)** (1990) *Textologie du journal.* Paris : Minard.
- LAFON, P.** (1981) "Analyse lexicométrique et recherche des cooccurrences." in *Mots*, n°3. Lyon : ENS Editions. pp. 95-148.
- LAFON, P.** (1984) *Dépouillements et Statistiques en lexicométrie.* Genève, Paris : Slatkine-Champion.
- LAFON, P. & SALEM, A.** (1983) "L'inventaire des segments répétés d'un texte." in *Mots*, n°6. Lyon : ENS Editions. pp. 161-173.
- LAGOUILLE, P.** (1997) *La vie politique en France au XIX^e siècle.* Paris : Ophrys.
- LAGRÉE, M., HARISMENDY, P. & DENIS, M. (Dir.)** (2000) *L'Ouest-Éclair : naissance et essor d'un grand quotidien régional, 1899-1933.* Rennes : Presses Universitaires de Rennes.
- LAMALLE, C. & SALEM, A.** (2002) "Types généralisés et topographie textuelle dans l'analyse quantitative des corpus textuels. " in MORIN, A., SEBILLOT, P. (éds.) *Actes des 6es Journées internationales d'analyse statistique des données textuelles (JADT 2002)*, Saint-Malo : Irisa. pp. 403-412.
- LAPLACE, R. & TOURNIER, M.** (1970) "Deux siècles de vocabulaire politique : élaboration d'une nomenclature." in *Cahiers de Lexicologie*, n°17. Paris : Honoré Champion. pp. 74-86.

- LEBART, L., MORINEAU, A. & PIRON, M.** (2000) *Statistique exploratoire multidimensionnelle*. Paris : Dunod.
- LEBART, L. & SALEM, A.** (1988) *Analyse statistique des données textuelles*. Paris : Dunod.
- LEBART, L. & SALEM, A.** (1994) *Statistique textuelle*. Paris : Dunod.
- LEBLANC, J.-M.** (2005) *Les Vœux présidentiels sous la Cinquième République (1959-2001). Recherches et expérimentations lexicométriques à propos de l'ethos dans un genre discursif rituel*. Thèse de Doctorat en Sciences du langage, sous la direction de P. Fiala. Paris : Université Paris X Nanterre.
- LEECH, G.** (1991) "The state of the art in corpus linguistics." in AIJMER, K. & ALTENBERG, B. (Dir.), *English Corpus Linguistics*. London : Longman, pp. 8-29.
- LEGALLOIS, D.** (2006a) "Hypertextualité et virtualité comme modes de la construction des discours et des connaissances." in *Pratiques*, n°129-130. Metz : CRESEF. pp. 139-156.
- LEGALLOIS, D.** (2006b) (Dir.) *Langages*, n°163 : "Unité(s) du texte." Paris : Larousse.
- LEIMDORFER, F. & SALEM, A.** (1994) "L'approche lexicométrique en analyse de discours." in *Les Cahiers des Sciences humaines de l'ORSTOM*, n°3. Paris : IRD/ORSTOM.
- LENOBLE, M.** (1986) "Statistique lexicale et critique littéraire : le mariage impossible?" in *Actes du Colloque International CNRS*, Université de Nice, 5-8 juin 1985. Paris : Slatkine-Champion. pp. 565-575.
- LEPREUX, G.** (1896) *Nos journaux. Histoire et bibliographie de la presse périodique dans le département du Nord (Flandres, Hainaut, Cambrésis, 1746-1889 -1895)*. Douai : L. et G. Crépin frères.
- LERICHE, F.** (2006) "Quel balisage pour les corpus épistolaires numériques? De l'annotation traditionnelle du "document" à une analyse générique et pragmatique." in RASTIER, F. & BALLABRIGA, M. (Dir.), *Corpus en Lettres et Sciences sociales. Des documents numériques à l'interprétation*, Paris : Texto! Disponible sur <http://www.revue-texto.net/Parutions/Livres-E/Albi-2006/Actes_ALBI-06.pdf> (Consulté le 12/10/2009).
- LERNER, H.** (1978) *La Dépêche, journal de la démocratie*. Toulouse : Publications de l'Université de Toulouse-le-Mirail.
- LETHIER, V.** (2008) "Formes et fonctions du rubriquage d'un quotidien régional du XIX^e siècle : *Le Petit Comtois* (1883-1903)." in *Semen*, n°25 : "Le discours de presse au XIX^e siècle : pratiques socio-discursives émergentes". Besançon : Presses Universitaires de Franche-Comté. pp. 115-133.

- LETHIER, V.** (2009) "*Multidimensional Analysis based on morpho-lexical features : The example of a 19th Century regional Press Corpus along with its columns*". Actes du Colloque thématique 2008 du CBL/BKL. Anvers : John Benjamins. pp. 175-190.
- LETHIER, V. & VIPREY, J.-M. (Dirs.)** (2008) *Le discours de presse au XIXe siècle : pratiques socio-discursives émergentes*. Semen, n°25. Besançon : Presses Universitaires de Franche-Comté.
- LOCHARD, G. & BOYER, H.** (1998a) *La communication médiatique*. Paris : Seuil.
- LOCHARD, G. & BOYER, H.** (1998b) *Scènes de télévision en banlieues*. Paris : INA-L'Harmattan.
- LOISEAU, S.** (2002) "Les standards : autour d'XML et de la TEI". Disponible sur <<http://www.revue-texto.net/Corpus/Corpus.html>> (Consulté le 10/05/2007).
- LONGRÉE, D., LUONG, X., MELLET, S. (2004)** "Temps verbaux, axe syntagmatique, topologie textuelle : analyse d'un corpus lemmatisé." in PURNELLE, G. & al. (éds.), *Actes des 7es Journées internationales d'analyse statistique des données textuelles (JADT 2004)*. Louvain : Presses universitaires de Louvain. pp. 743-752.
- LONGRÉE, D., LUONG, X., MELLET, S.** (2008) "Les motifs : un outil pour la caractérisation topologique des textes. " in HEIDEN, S & PINCEMIN, B. (éds.), *Actes des 9es Journées internationales d'Analyse statistique des Données Textuelles (JADT 2008)*. Lyon : Presses Universitaires de Lyon. pp. 733-744.
- LORDA, C.-U.** (2001) "Les articles dits d'information : la relation de déclarations politiques." in *Semen*, n°13. Besançon : Presses Universitaires de Franche-Comté. pp. 123-138.
- LUCAY, R.** (1885) *Les ministères sous Jules Grévy jusqu'en 1885*. Lyon : Imprimerie nouvelle.
- LUGRIN, G.** (2000) "Critères de typologisation des genres de la presse écrite." in *Études de lettres*, vol. 4, 4. Lausanne : IASA.
- LUGRIN, G.** (2000) "Les ensembles rédactionnels : multitexte et hyperstructure." in *Médiatiques*, n°11. Louvain-la-Neuve : Observatoire du récit médiatique (ORM). pp. 34-36.
- M' SILI, M.** (2000) *Le fait divers en République. Histoire sociale de 1870 à nos jours*. Paris : CNRS Editions.
- MAINGUENEAU, D.** (1976) *Initiation aux méthodes de l'analyse du discours*. Paris : Hachette.
- MAINGUENEAU, D.** (1991) *L'Analyse du discours. Introduction aux lectures de l'archive*. Paris : Hachette.

- MAINGUENEAU, D.** (1996) *Les Termes-clés de l'analyse du discours*. Paris : Seuil.
- MAINGUENEAU, D.** (2005a) "L'analyse du discours et ses frontières." in MAINGUENEAU, D. (**Dir.**), *Marges linguistiques*, n°9 (Mai 2005) : "Analyse du discours. État de l'art et perspectives", Saint-Chamas, pp. 64-75. Disponible sur <<http://www.revue-texto.net/marges/Marges.html>> (Consulté le 12/10/2009).
- MAINGUENEAU, D.** (2005b) "Philologie et analyse du discours." in ADAM, J.M. & HEIDMANN, U. (**Dirs.**), *Sciences du texte et analyse de discours : enjeux d'une interdisciplinarité*. Genève : Slatkine Erudition, pp. 37-51.
- MALDIDIER, D.** (1971) "Lecture des discours de De Gaulle par six quotidiens parisiens : 13 mai 1958." in MARCELLESI, J.B. (**Dir.**), *Langue Française*, n°9 : "Linguistique et Société". Paris : Larousse. pp. 34-46.
- MALDIDIER, D.** (1990) *L'inquiétude du discours*. Paris : Editions des Cendres.
- MALDIDIER, D.** (1994) "Éléments pour une histoire de l'analyse de discours en France." in GUILHAUMOU, J., MALDIDIER, D., ROBIN, R., *Discours et archive*. Liège : Mardaga. pp. 173-184.
- MALDIDIER, D. & ROBIN, D.** (1977) "Du spectacle au meurtre de l'événement : reportages, commentaires et éditoriaux de presse à propos de Charléty (mai 1968)." in *Pratiques*, n°14. Metz : CRESEF. pp. 21-65.
- MALRIEU, D.** (2004) "Linguistique de corpus, genres textuels, temps et personnes". in *Langages*, n°153. Paris : Larousse. pp. 73-85.
- MALRIEU, D. (Dir.)** (2007) *"Corpora et questionnements du littéraire". Numéro Spécial de Linx*. Paris : Presses Universitaires de Paris X.
- MANEVY, R.** (1955) *La presse de la III^e République*. Paris : Editions Foret.
- MANEVY, R.** (1956) *L'évolution des formules de présentation de la presse*. Paris : Editions Estienne.
- MANEVY, R.** (1958) *La presse française de Renaudot à Rochefort*. Paris : Forest.
- MARCELLESI, J.-B.** (1970) "Problèmes de socio-linguistique : Le congrès de Tours." in *La Pensée*, n°153. Toulouse : Académie de Toulouse. pp. 68-80.
- MARCELLESI, J.-B.** (1971) *Le Congrès de Tours (déc. 1920) : études socio-linguistiques*. Paris : Le Pavillon.
- MARCELLESI, J.-B.** (1976) "Analyse de discours à entrée lexicale. Application à un corpus de 1924-1925." in *Langages*, n°41. Paris : Larousse. pp. 79-124.

- MARCOUX, Y.** (2006) *A natural-language approach to modeling : why is some XML so difficult to write?* Disponible sur <<http://conferences.idealliance.org/extreme/html/2006/Marcoux01/EML2006Marcoux01.html>> (Consulté le 07/10/2009).
- MARCOUX, Y.** (2008) "*Graph characterization of overlap-only TexMECS and other overlapping markup formalisms.* Balisage Series on Markup Technologies, vol. 1. Montréal, Canada. Disponible sur <<http://www.balisage.net/Proceedings/vol1/html/Marcoux01/BalisageVol1-Marcoux01.html>> (Consulté le 07/10/2009).
- MARLIN, R.** (1965a) *Bibliographie de la presse française politique et d'information générale, 1865-1944, 25-Doubs et territoire de Belfort.* Paris : Bibliothèque Nationale.
- MARLIN, R.** (1965b) "Courte histoire de la presse politique du Doubs aux XIX^e et XX^e siècles." in *Mémoires de la Société d'Emulation du Doubs*, Nouvelle série, n°7. Besançon : Société d'Emulation du Doubs. pp. 179-214.
- MARNETTE, S.** (2004) "L'effacement énonciatif dans la presse contemporaine." in *Langages*, n°156. Paris : A. Colin. pp. 51-64.
- MARTIN, M.** (1981) "Journalistes parisiens et notoriété (vers 1830-1870) : pour une histoire sociale du journalisme." in *Revue historique*, n°589. Paris : Presses Universitaires de France. pp. 31-74.
- MARTIN, M.** (1997) *Médias et journalistes de la République.* Paris : Editions Odile Jacob.
- MARTIN, M.** (2002) *La presse régionale. Des Affiches aux grands quotidiens.* Paris : Fayard.
- MARTINEZ, W.** (2003) *Contribution à une méthodologie de l'analyse des cooccurrences lexicales multiples dans les corpus textuels.* Thèse de doctorat en Sciences du Langage. Paris : Université Paris 3 - Sorbonne Nouvelle.
- MARTIN-LAGARDETTE, J.-L.** (2000) *Le guide de l'écriture journalistique. Concevoir, rédiger, présenter l'information.* Paris : Syros.
- MASSONIE, J.-P.** (1990a) *Analyse informatisée des textes.* Besançon : Annales littéraires de l'Université de Besançon.
- MASSONIE, J.-P.** (1990b) *Pratique de l'analyse des correspondances.* Paris : Les Belles lettres.
- MATHIEN, M.** (1986) *La presse quotidienne régionale.* Paris : Presses Universitaires de France.
- MAYAFFRE, D.** (2002a) "Les corpus réflexifs : entre architextualité et intertextualité." in *Corpus* n°1, pp. 51-70. Disponible sur <<http://corpus.revues.org/index11.html>> (Consulté le 12/10/2009).

- MAYAFFRE, D.** (2002b) "L'Herméneutique numérique." in *L'Astrolabe. Recherche littéraire et informatique*. Disponible sur <<http://www.uottawa.ca/academic/arts/astrolabe>> (Consulté le 12/10/2009).
- MAYAFFRE, D.** (2000c) *Le poids des mots. Le discours de gauche et de droite dans l'entre-deux-guerres*. Paris : Champion.
- MAYAFFRE, D.** (2000d) "Temps lexical ou temps politique ? " in *Lexicometrica*, n°2. Disponible sur <<http://www.cavi.univ-paris3.fr/lexicometrica/article/numero2/mayaffre2000.PDF>> (Consulté le 10/10/2009).
- MAYAFFRE, D.** (2002c) "Discours politique, genres et individuation socio-linguistique." in MORIN, A., SEBILLOT, P.(éds.) *Actes des 6es Journées internationales d'analyse statistique des données textuelles (JADT 2002)*, Saint-Malo : Irsa. pp. 517-529.
- MAYAFFRE, D.** (2004) *Paroles de président. Jacques Chirac (1995-2003) et le discours présidentiel sous la Vème république*. Paris : Champion.
- MAYAFFRE, D.** (2005a) "De la lexicométrie à la logométrie." in *L'Astrolabe. Recherche littéraire et Informatique*. Disponible sur <<http://www.uottawa.ca/academic/arts/astrolabe/articles/art0048/Logometrie.htm>> (Consulté le 07/10/2009).
- MAYAFFRE, D.** (2005b) "Les corpus politiques : objet, méthode et contenu. Introduction." in MAYAFFRE, D. (Dir.), *Corpus*, n°4 : "Les corpus politiques : objet, méthode et contenu", pp. 5-19. Disponible sur <<http://corpus.revues.org/index292.html>> (Consulté le 07/10/2009).
- MAYAFFRE, D.** (2006) "Faut-il prendre en compte la composition grammaticale des textes ? " in VIPREY, J.-M. (éds.), *Actes des 8es Journées internationales d'Analyse statistique des Données Textuelles (JADT 2006)*. Besançon : Presses Universitaires de Franche-Comté. pp. 673-681.
- MAYAFFRE, D.** (2007a) *Philologie et/ou herméneutique numérique : nouveaux concepts pour de nouvelles pratiques*. Toulouse : Presses Universitaires de Toulouse. pp. 15-26.
- MAYAFFRE, D.** (2007b) "Analyses logométriques et rhétoriques des discours." in OLIVÉSI, S. (Dir.), *Introduction à la recherche en SIC*. Grenoble : Presses Universitaires de Grenoble, pp. 153-180.
- MAYAFFRE, D.** (2007c) "L'analyse de données textuelles aujourd'hui : du corpus comme une urne au corpus comme un plan. Retour sur les travaux actuels de topographie/topologie textuelle." in *Lexicometrica*, n°7 : "Topographie et topologie textuelle". Disponible sur <<http://www.cavi.univ-paris3.fr/lexicometrica/numspeciaux/special9/mayaffre.pdf>> (Consulté le 07/10/2009).

- MAYAFFRE, D.** (2008a) "L'entrelacement lexical des textes, co-occurrences et lexicométrie." *Texte et corpus*, n°3 / août 2008. pp.91-102. Disponible sur <http://web.univ-ubs.fr/corpus/jlc5/ACTES/ACTES_JLC07_mayaffre.pdf> (Consulté le 12/10/2009).
- MAYAFFRE, D.** (2008b) "*Quand "Travail" "famille" "patrie" co-occurrent dans le discours de Nicolas Sarkozy. Etude de cas et réflexion théorique sur la co-occurrence.*" in HEIDEN, S & PINCEMIN, B. (éds.), *Actes des 9es Journées internationales d'Analyse statistique des Données Textuelles (JADT 2008)*. Lyon : Presses Universitaires de Lyon. pp. 811-822.
- MAYEUR, J.-M.** (1973) *Les Débuts de la IIIe République, 1871-1898*. Paris : Seuil.
- MAYEUR, J.-M.** (1984) *La vie politique sous la Troisième République (1870-1940)*. Paris : Editions du Seuil.
- MAZÉDIER, R.** (1945) *Histoire de la Presse Parisienne. De Théophraste Renaudot à la IVe République*. Paris : Editions du Pavois & Paris.
- MAZIÈRE, F.** (2005) *L'Analyse du discours*. Paris : Presses Universitaires de France, coll. Que sais-je?
- MCENERY, T. & WILSON, A.** (1996) *Corpus Linguistics*. Edinburgh : Edinburgh Press.
- MEALAND, D.** (1997) "Measuring Genre Differences in Mark with correspondance analysis." in *Literary and Linguistic Computing*, 12. Oxford : Oxford Journal. pp. 227-243.
- MELLET, S. (Dir.)** (2002) *Corpus*, n°1 : *Corpus et recherches linguistiques*. Disponible sur <<http://corpus.revues.org/index49.html>> (Consulté le 07/09/2009).
- MELLET, S. & BARTHÉLEMY, J.-P.** (2007) "La topologie textuelle : légitimation d'une notion émergente." in *Lexicométrica*, numéro thématique "Topographie et topologie textuelle". Disponible sur <<http://www.cavi.univ-paris3.fr/lexicométrica/numspeciaux/special9/mellet.pdf>> (Consulté le 12/10/2009).
- MELLET, S. (Dir.)** (1998) *JADT 1998, 4èmes Journées Internationales d'Analyses des Données Textuelles*. Nice : Université de Nice Sophia-Antipolis, CNRS, INALF.
- MELMOUX-MONTAUBIN, M.-F.** (2001) "Leurres et incertitudes : Lire la presse du XIXe siècle." in THÉRENTY, M.-E. (Dir.), *1836, l'An I de l'ère médiatique*. Paris : Nouveau monde éditions, pp. 327-335.

- MEZZASALMA, P. & JOANNIC-SETA, F.** (2007) "La numérisation de la presse à la BNF." Communication présentée lors de "*Hier à la une : les collections de presse dans les bibliothèques françaises.*" in *3èmes Journées Patrimoine écrit 2007, à Rennes, les 13 et 14 septembre 2007.*
- MOIRAND, S.** (2000a) "Du traitement différent de l'intertexte selon les genres convoqués dans les événements scientifiques à caractère politique." in *Semen*, n°13. Besançon : Presses Universitaires de Franche-Comté. pp. 97-118.
- MOIRAND, S.** (2000b) "Les indices dialogiques de contextualisation dans la presse ordinaire." in *Cahiers de praxématique*, n°33 : "Sémantique de l'intertexte". Montpellier : Presses universitaires de la Méditerranée. pp. 154-184.
- MOIRAND, S.** (2003) "Les lieux d'inscription d'une mémoire interdiscursive." in JUHANI, H. (Dir.), *Le langage des médias : discours éphémères?* Paris : L'Harmattan, pp. 83-111.
- MOIRAND, S.** (2004a) "L'impossible clôture des corpus médiatiques ou la construction des observables entre catégorisation et contextualisation." in *TRANEL*, n°40. Neuchâtel : Centre de Linguistique Appliquée. pp. 71-92.
- MOIRAND, S.** (2004b) "Le texte et ses contextes." in ADAM, J.-M., ALI BOUACHA, M., GRIZE, J.-B., et al. (Dir.) *Texte et discours : catégories pour l'analyse.* Dijon : Editions Universitaires de Dijon. pp. 129-143.
- MOLLIER, J.-Y. & GEORGES, J.** (1994) *La Plus Longue des Républiques (1870-1940).* Paris : Fayard.
- MONVILLE-BURSTON, M.** (1992) "Les verba dicendi dans la presse d'information." in *Langue Française*, n°93. Paris : Larousse. pp. 48-66.
- MOUILLAUD, M.** (1982) "Grammaire et idéologie du titre de journal." in *Mots*, n°4. Lyon : ENS Editions. pp.69-91.
- MOUILLAUD, M. & TÉTU, J.-F.** (1989) *Le journal quotidien.* Lyon : Presses Universitaires de Lyon.
- MOUILLAUD, M.** (1990) "Le journal un texte sous tension." in RÉTAT, P. (Dir.) *Textologie du journal*, in *Cahiers de textologie*, pp. 141-155.
- MOURIQUAND, J.** (1997) *L'écriture journalistique.* Paris : Presses Universitaires de France.
- MUHLMANN, G.** (2004) *Une histoire politique du journalisme XIXe-XXe siècle.* Paris : Presses Universitaires de France.
- MÜLLER, B.** (2005) "Linguistic turn, histoire." in *Dictionnaire des idées.* Encyclopedia Universalis.
- MULLER, C.** (1964) *Essai de statistique lexicale.* Paris : Klincksieck.

- MULLER, C.** (1993) *Initiations aux méthodes de la statistique linguistique*. Paris : Champion.
- MULLER, C.** (1993) *Principes et méthodes de statistique lexicale*. Paris : Champion.
- MULLER, P.-E.** (1999) "L'égalité dans le discours socialiste autour de 1900." in FIALA, P. (**Dir.**), *In/égalité/s*. Paris : L'Harmattan, pp. 147-169.
- NARDIN, M.** (2000) "Le Petit Comtois et le faits divers criminel de 1880 à 1914." in STORA LAMARRE, A. (**Dir.**), *Archives de la peur. Les "populations à risque" dans la Franche-Comté au XIXe siècle*. Besançon : Presses Universitaires franc-comtoises, pp. 103-121.
- NÉRÉ, J.** (1964) *Le boulangisme et la presse*. Paris : Armand Colin.
- NETTEMENT, A.** (1846) *Etudes critiques sur le roman-feuilleton*. Paris : De Perrodie.
- NETTEMENT, A.** (1846) *La Presse Parisienne*. Paris : Dentu.
- NEVEU, E.** (2001) *Sociologie du journalisme*. Paris : La Découverte.
- NICOLET, C.** (1982) *L'idée républicaine en France*. Paris : Gallimard.
- NOIRIEL, G.** (1996) *Sur la 'crise' de l'histoire*. Paris : Belin.
- NOIRIEL, G.** (1998) *Qu'est-ce que l'histoire contemporaine?* Paris : Hachette.
- NOIRIEL, G.** (2003) *Penser avec, penser contre. Itinéraire d'un historien*. Paris : Belin.
- OGIER, J.-M., MULLOT, R., LABICHE, J., et al.** (2000) "The Semantic Coherency : The Basis of an Image Interpretation Device. Application to the French Cadastral Map Interpretation." in *IEEE Transactions on Systems, Man, and Cybernetics*, Part B "Cybernetics", 30 (2). pp. 323-338.
- PAITIER, E.** (2003) *L'image de l'armée dans la presse nancéenne*. Mémoire de Maîtrise en Sciences du Langage sous la direction de J.-P. Bois. Nantes : Université de Nantes.
- PALMER, M.** (1983) *Des petits journaux aux grandes agences : naissance du journalisme moderne, 1863-1914*. Paris : Aubier.
- PALMER, M.** (2006) "Plumes à louer, combats à mener, carrière à faire : Les journalistes en République, à travers leurs associations et les campagnes électorales (1880-1914)." in *Le Temps des Médias*, n°7, 2. Paris : Nouveau Monde éditions. pp. 35-47.
- PAVEAU, M.-A.** (2007) "Analyse du discours et histoire. Rencontres et oublis." in BONNAFOUS, S. & TEMMAR, M. (**Dirs.**), *Analyse du discours et sciences humaines et sociales*. Paris : Ophrys.

- PAVEAU, M.-A.** (2007) "Interdiscours et intertexte : Généalogie scientifique d'une paire de faux jumeaux." Communication présentée lors de *Linguistique et littérature : Cluny, 40 ans après, Besançon, 29-31 octobre 2007*, Organisation de KATSBURG-SJOBLOM, M. & ABLALI, D.
- PECHEUX, M.** (1975) "Analyse du discours, langue et idéologies." in *Langages*, n°37. Paris : Larousse. pp. 3-6.
- PÊCHEUX, M.** (1969) *Analyse automatique du discours*. Paris : Dunod.
- PÉRY-WOODLEY, M.-P.** (1995) "Quels corpus pour quels traitements automatiques?" in *TAL*, Vol. 36, 1-2. Jussieu : CNRS/Association pour le Traitement Automatique des Langues. pp. 213-232.
- PESCHANSKI, D.** (1989) *Et pourtant ils tournent : vocabulaire et stratégie du PCF (1934-1936)*. Paris : Klincksieck.
- PETITJEAN, A.** (1987) "Les faits divers : polyphonie énonciative et hétérogénéité textuelle." in *Langue Française*, n°74. Paris : Larousse. pp. 73-96.
- PEYTARD, J.** (1975) "Lecture(s) d'une "aire scripturale"; "la page de journal"." in *Langue Française*, n° 28. Paris : Larousse. pp. 39-59.
- PEYTARD, J.** (1992) "Analyse informatisée du texte littéraire : Pertinences et limites." in PEYTARD, J. (Dir.), *Syntagmes*, 4. Besançon : Annales Littéraires de l'Université de Besançon, Les Belles Lettres. pp. 173-185.
- PEYTARD, J.** (1993) "D'une sémiotique de l'altération." in *Semen*, 8 : "Configurations discursives". Besançon : Presses Universitaires de Franche-Comté. pp. 143-177.
- PEYTARD, J.** (1999) "Écriture et pointillés de sens : lecture-analyse de deux pages de Proust (*La fin de la Jalousie*). " in *Semen*, n°11. Besançon : Presses Universitaires de Franche-Comté. pp. 357-376.
- PIÉGAY-GROS, N.** (1996) *Introduction à la textualité*. Paris : Dunod.
- PINCEMIN, B.** (1999) "Construire et utiliser un corpus : le point de vue d'une sémantique textuelle interprétative" in CONDAMINES, A., et al. (éds.), *Corpus et TAL : pour une réflexion méthodologique, 6^e Conférence annuelle sur le Traitement Automatique des Langues Naturelles (TALN 99)*. Cargèse. pp. 26-36.
- PLOUX, F.** (2003) *De bouche à oreille. Naissance et propagation des rumeurs dans la France du XIXe siècle*. Paris : Aubier.
- POCOCK, J.G.A.** (1971) *Politics, Language and Time*. New-York : Methuen.
- POCOCK, J.G.A.** (1997) *Le Moment machiavélien. La pensée politique florentine et la tradition républicaine antique*. Paris : PUF.
- POCOCK, J.G.A.** (1998) *Vertu, commerce et histoire*. Paris : PUF.

- PONTY, J.** (1974) "La presse quotidienne et l'affaire Dreyfus en 1898-1899." in *Revue d'histoire moderne et contemporaine*, Vol. 21. Paris : Belin. pp. 194-220.
- POPKIN, J.D.** (2002) *Press, Revolution and Social Identity in France, 1830-1835*. Pennsylvania : The Penn State University Press.
- POUDAT, C.** (2006) *Étude contrastive de l'article scientifique de revue linguistique dans une perspective d'analyse des genres*. Thèse de Doctorat en Sciences du langage, sous la direction de G. Bergounioux. Orléans : Université d'Orléans. Disponible sur < <http://www.revue-texto.net/Corpus/Publications/Poudat/Etude.html> > (Consulté le 01/05/2008).
- PROST, A.** (1974a) *Vocabulaire des proclamations électorales de 1881, 1885 et 1889*. Paris : Presses Universitaires de France.
- PROST, A.** (1974b) "Combattants et politiciens. Le discours mythologique sur le politique entre les deux guerres." in GUILHAUMOU, J., MALDIDIÉ, D., PROST, A., et al. (1974) *Langage et idéologies. Le discours comme objet de l'histoire*. Paris : Editions ouvrières. pp. 117-149.
- PROST, A.** (1988) "Les mots." in RÉMOND, R. (Dir.), *Pour une histoire politique*. Paris : Seuil, pp.255-286.
- PROST, A.** (1996) *Douze leçons sur l'histoire*. Paris : Editions du Seuil.
- PROVOST-CHAUVEAU, G.** (1971) "Problèmes théoriques et méthodologiques en analyse du discours." in MARCELLES, J.B. (Dir.), *Langue Française*, n°9 : "Linguistique et Société". Paris : Larousse. pp. 6-21.
- QUEFFELEC, L.** (1989) *Le Roman-feuilleton français au XIXe siècle*. Paris : Presses Universitaires de France, Coll. Que sais-je?
- RABATEL, A.** (1997) *Une histoire de point de vue*. Paris-Metz : Klincksieck-CELTED.
- RABATEL, A.** (1998) *La construction textuelle du point de vue*. Lausanne-Paris : Delachaux & Niestlé.
- RABATEL, A.** (2003) "La valeur de 'on' pronom indéfini/pronom personnel dans les perceptions représentées." in *L'information grammaticale*, n°88. Paris : Bibliothèque de l'information grammaticale. pp. 28-32.
- RABATEL, A.** (2004a) "La déliaison des énonciateurs et des locuteurs dans la presse satirique." in *Langage & Société*, n° 110. Paris : Maison des Sciences de l'Homme. pp. 7-23.
- RABATEL, A.** (2004b) "L'effacement énonciatif dans les discours rapportés et ses effets pragmatiques." in *Langages*, n°156. Paris : Larousse. pp. 3-17.

- RABATEL, A.** (2005) "Effacement énonciatif et argumentation indirecte. 'on-perceptions', 'on-représentations' et 'on-vérités' dans les points de vue stéréotypés." in *Signes, langues et cognition*, RACCAH, P.-Y. (Dir.). Paris : L'Harmattan. pp. 85-116.
- RABATEL, A. & CHAUVIN-VILENO, A. (Dir.)** (2006) *Énonciation et responsabilité dans les médias*. Semen n°22. Besançon : Presses Universitaires de Franche-Comté.
- RABATEL, A. & CHAUVIN-VILENO, A.** (2006) "La question de la responsabilité dans l'écriture de presse." in *Semen*, n°22 : "Énonciation et responsabilité dans les médias". Besançon : Presses Universitaires de Franche-Comté. pp. 7-28.
- RAMEL, J.-Y., BOISSIER, G. & EMPTOZ, H.** (1999) *Automatic Reading of Handwritten Chemical Formulas from a Structural Representation of the Image*. Bangalore : IEEE Computer Society. pp. 83-86.
- RASTIER, F.** (1987) *Sémantique interprétative*. Paris : Presses Universitaires de France.
- RASTIER, F.** (1989) *Sens et textualité*. Paris : Hachette.
- RASTIER, F. (Dir.)** (1995) *L'Analyse thématique des données textuelles*. Paris : Didier.
- RASTIER, F.** (1996) "Pour une sémantique des textes. Questions d'épistémologie." in RASTIER, F. (Dir.), *Texte et sens*. Paris : Didier Erudition, pp. 9-35.
- RASTIER, F.** (1998) "Le problème épistémologique du contexte et le statut de l'interprétation dans les sciences du langage." in *Langages*, n°129. Paris : Larousse. pp. 97-111.
- RASTIER, F.** (2000) "Philologie numérique." in TYVAERT, J.-E. (Dir.), *Recherches en Linguistique et Psychologie cognitive*, n°16 : "Philologie électronique et assistance à l'interprétation des textes". Reims : Presses Universitaires de Reims. pp. 117-149.
- RASTIER, F.** (2001) *Arts et sciences du texte*. Paris : Seuil.
- RASTIER, F.** (2007) "Passages." in PINCEMIN, B. (Dir.), *Corpus*, n°6 : "Interprétation, contextes, codage", pp. 25-54. Disponible sur <<http://corpus.revues.org/index832.html>> (Consulté le 12/10/2009).
- RASTIER, F. & MALRIEU, D.** (2001) "Genres et variations morpho-syntaxiques." in *TAL*, n°42, 2. Jussieu : CNRS/Association pour le Traitement Automatique des Langues. pp. 547-577.
- RASTIER, F. & PINCEMIN, B.** (1999) "Des genres à l'intertexte " in *Cahiers de Praxématique*, n°23 : " La négation et ses marges : absence, écart, rupture". Montpellier : Presses Universitaires de la Méditerranée. pp. 90-111.

- RÉBÉRIOUX, M.** (1975) *La République radicale ? 1898-1914*. Paris : Seuil.
- REICHARDT, R.** (1982) "Pour une histoire des mots-thèmes socio-politiques." in *Mots*, n°5. Lyon : ENS Editions. pp. 189-202.
- REICHARDT, R., LÜSEBRINK, H.-J., SCHMITT, E.** (1985-2000) *Handbuch politisch-sozialer Grundbegriffe in Frankreich (1680-1820)*. München : Olgenbourg. Heft 1-20.
- REINERT, M.** (1993) "Les 'mondes lexicaux' et leur logique." in *Langage & Société*, n°66. Paris : Maison des Sciences de l'Homme. pp. 5-39.
- REINERT, M.** (1998) "Quelques interrogations à propos de 'l'objet' d'une analyse de discours de type statistique et de la réponse 'Alceste'." in *Langage & Société*, n°90. Paris : Maison des Sciences de l'Homme. pp. 57-70.
- RÉTAT, P. (Dir.)** (1982) *Le journalisme d'Ancien Régime*. Lyon : Presses Universitaires de Lyon.
- RÉTAT, P. (Dir.)** (1989) *La Révolution du Journal (1788-1794)*. Lyon : Editions du CNRS.
- RÉTAT, P. (Dir.)** (1990) *Cahiers de Textologie*, n°3 : "Textologie du journal." Paris : Minard.
- REVAZ, F.** (1997) "Le récit dans la presse écrite." in *Pratiques*, n°94. Metz : CRESEF. pp. 19-33.
- RIEGEL, M., et al.,** (1994) *Grammaire méthodique du français*. Paris : Presses Universitaires de France.
- RIGOOT, R. & ROCHARD, Y.** (2005) "Proximité éditoriale : normes et usages des genres journalistiques." in *Mots*, n°77. Lyon : ENS Editions. pp. 73-90.
- RIGOT, H. (2006)** "(en)-jeux de corpus pour la recherche en SHS. Énoncés, textes et documents." in RASTIER, F. & BALLABRIGA, M. (Dir.), *Corpus en Lettres et Sciences sociales. Des documents numériques à l'interprétation*, Paris : Texto! Disponible sur <<http://www.revue-texto.net/Parutions/Livres-E/Albi-2006/Rigot.pdf>>. (Consulté le 08/07/2009).
- RINGLET, G.** (1981) *Le mythe au milieu du village. Comprendre et analyser la presse locale*. Bruxelles : Editions Vie Ouvrière.
- ROBERT, P.** (2004) "Critique de la dématérialisation." in *Communication & langages*, n°140. Paris : NecPlus. pp. 55-68.
- ROBIN, R.** (1973) *Histoire et Linguistique*. Paris : Colin.
- ROBIN, D.** (1975) "Le champ sémantique de féodalité dans les Cahiers de doléances généraux de 1789." in *Bulletin du Centre d'analyse du discours de Lille III*, n°2. Lille : Presses Universitaires de Lille. pp.61-86.

- ROLE, F.** (1999) "Représentation et exploitation des métadonnées complexes : le cas des documents anciens." in *Document numérique*, n°3 : "Les documents anciens". Paris : Lavoisier. pp. 135-151.
- ROSANVALLON, P.** (2000) *La démocratie inachevée. Histoire de la souveraineté du peuple en France*. Paris : Gallimard.
- RORTY, R.** (1967) *The linguistic turn: essays in Philosophical method*. Chicago : Chicago University Press.
- ROSIER, L.** (1999) *Le discours rapporté, Histoire, théories, pratiques*. Paris : Duculot.
- ROTH, F. (Dir.)** (1983) *Le temps des journaux : presse et cultures nationales en Lorraine mosellane : 1860-1940*. Metz : Ed. Serpenoise.
- ROTH, F. (Dir.)** (1989) *Le quotidien dévoilé. L'Est Républicain (1889-1989)*. Jarville-La Malgrange : Ed. de l'Est.
- SALANSKIS, J.-M.** (2003) *Herméneutique et cognition*. Villeneuve-D'Ascq : Presses universitaires du Septentrion.
- SALAÜN, J.-M.** (2002) "Documents numériques et Universités françaises." in GUICHARD, E. (Dir.), *Comprendre les usages de l'Internet*. Lyon : ENS-Ulm, pp. 112-117.
- SALAÜN, J.-M.** (2004) "Chronique inachevée d'une réflexion collective sur le document." in *Communication & langages*, n°140. Paris : NecPlus. pp. 9-18.
- SALEM, A.** (1982) "Analyse factorielle et lexicométrie: synthèse de quelques expériences." *Mots*, n°4. Lyon : ENS Editions. pp. 147-168.
- SALEM, A.** (1987) *Pratique des segments répétés : essai de statistique textuelle*. Paris : Klincksieck.
- SALEM, A.** (1988) "L'étude quantitative des corpus chronologiques ." in *Actes des Rencontres de Nancy les 21, 22 et 23 novembre 1985*. Paris : Klincksieck. pp. 49-62.
- SALEM, A.** (1991) "Les Séries textuelles chronologiques." in *Histoire & Mesure*, VI, 1/2 : "Séries temporelles". Paris : Éditions de l'EHESS, pp. 149-175. Disponible sur <<http://histoiremesure.revues.org/index600.html>> (Consulté le 07/09/2009).
- SALEM, A.** (1994) "Lexicométrie chronologique." in *Actes du colloque de lexicologie politique : "Langages de la Révolution"*. Paris : Klincksieck. Collection "St. Cloud."
- SALVAN, J.** (1982) "Le journal, champ d'investigation des évolutions lexicales et intellectuelles." in RÉTAT, P. (Dir.), *Le journalisme d'Ancien Régime*. Lyon : Presses Universitaires de Lyon, pp. 43-53.
- SARFATI, G.E.** (1997) *Éléments d'analyse du discours*. Paris : Nathan.

- SARFATI, G.E.** (2003) "L'analyse du discours et la question du texte : l'horizon du sens commun et de la doxa." in MAINGUENEAU, D. & AMOSSY, R. (**Dirs.**), *L'analyse du discours dans les études littéraires*. Toulouse : Cerisy-Presses universitaires du Mirail, pp. 429-438.
- SCHEPENS, P.** (2004) "Le rapport de l'Inserm sur les psychothérapies : quelques points d'analyse." in *Cliniques Méditerranéennes*, n°71, Toulouse : Erès. Disponible sur <<http://www.cairn.info/revue-cliniques-mediterraneennes-2005-1.htm>> (Consulté le 07/09/2009).
- SCHEPENS, P. (Dir.)** (2006) *Catégories pour l'analyse du discours politique*. Semen n°21. Besançon : Presses Universitaires de Franche-Comté.
- SCHEPENS, P.** (2006) "Médias et responsabilité : pour un point de vue bakhtinien." in CHAUVIN-VILENO, A. & RABATEL, A. (**Dirs.**), *Semen*, n°22 : "Enonciation et responsabilité dans les médias". Besançon : Presses Universitaires de Franche-Comté. pp.61-66.
- SCHEPENS, P.** (2008) *La construction des représentations franc-comtoises de l'Europe à travers les bulletins professionnels, confessionnels, syndicaux, politiques, corporatifs, etc. Analyse de contenu et analyse du discours*. Besançon : Rapport de recherche pour le SGAR de Franche-Comté et la Maison des Sciences de l'Homme et de l'Environnement.
- SCHEPENS, P. & VIPREY, J.M.** (A paraître) "A la jonction des disciplines, l'analyse de discours aujourd'hui." Cambridge : Cambridge Scholars Publishing.
- SCHOR, R.** (2004) *Histoire de la société française au XXe siècle*. Paris : Belin.
- SEBASTIAN, J.-F. & FUENTES, J.-F. (Dirs.)** (2002) *Diccionario político y social del siglo XIX español*. Madrid : Alianza Editorial.
- SÉGU, F. (Dir.)** (1932) *Le premier Figaro, 1826-1833*. Paris : Les Belles Lettres.
- SGARD, J.** (1981) "Journale und Journalisten im Zeitalter der Aufklärung " in GUMBRECHT, H.-U., REICHARDT, R., *et al. (Dirs.)*, *Sozialgeschichte der Aufklärung in Frankreich*. München-Wien : Oldenbourg, pp.3-33.
- SGARD, J.** (1999) "Les lecteurs des *Affiches du Dauphiné* à la fin du XVIIIe siècle." in GRANGE, D.J. (**Dir.**), *La presse régionale XIXe-XXe siècles*. Grenoble : Presses de l'Université Pierre Mendès France, pp.15-23.
- SHAPIRA, C.** (1999) *Les Stéréotypes en français. Proverbes et autres formules*. Paris : Ophrys.
- SHAPIRA, C.** (2000) "Proverbe, proverbialisation et déproverbialisation." in *Langages*, n°139. Paris : Larousse. pp. 81-97.
- SIMON-VANDENBERGEN, A-M.** (1981) *The grammar of headlines in The Times 1870-1970*. Brussel : AWLSK.

- SINCLAIR, J.** (1996) *Preliminary recommendations on Corpus Typology*. Rapport technique pour EAGLES (Expert Advisory Group on Language Engineering Standards).
- SKINNER, Q.** (1996) *Reason and Rhetoric in the Philosophy of Hobbes*. Cambridge : Cambridge University Press.
- SKINNER, Q.** (2001) *Les fondements de la pensée politique moderne*. Paris : Albin Michel.
- SKINNER, Q.** (2002) *Vision of Politics*. Vol I "Regarding Method", Vol. II "Renaissance Virtues, Vol. III "Hobbes and Civil Science". Cambridge : Cambridge University Press.
- SORLIN, P.** (1967) *La Croix et les juifs (1880-1899)*. Paris : Grasset.
- SOUCHIER, E.** (2004) "Mémoires - outils - langages. Vers une société du texte?" in *Communication & langages*, n°139. Paris : NecPlus. pp. 41-52.
- SOUCHIER, E., JEANNERET, Y. & LE MAREC, J. (Dir.)** (2003) *Lire, écrire, récrire. Objets, signes et pratiques des médias informatisés*. Paris : Bibliothèque publique d'information.
- SPERBERG-MCQUEEN, C.M. & BURNARD, L. (Dir.)** (2001) *The TEI Consortium : Text Encoding initiative, The XML Version of TEI Guidelines*. Disponible sur <<http://www.tei-c.org/index.xml>> (Consulté le 10/10/2008).
- STAROBINSKY, J.** (1970) "L'interprète et son cercle." in STAROBINSKY, J., *La relation critique*. Paris : Gallimard.
- STONE, L.** (1979) "Revival of narrative: reflections on a new old History." in *Past and Present*, n°85. pp. 3-24.
- STORA LAMARRE, A. (Dir.)** (2000) *Archives de la peur : les "populations à risque" dans la Franche-Comté au XIXe siècle*. Besançon-Paris : Presses universitaires franc-comtoises-Les Belles Lettres
- SULLET-NYLANDER, F.** (1998) *Le titre de presse. Analyses syntaxique, pragmatique et rhétorique*. Thèse de doctorat en Sciences du Langage. Stockholm : Stockholm Universitet.
- SULLET-NYLANDER, F.** (2005) "Jeux de mots et défigements à la Une de *Libération* (1973-2004)." in *Langage & Société*, n°112. Paris : Maison des Sciences de l'Homme. pp. 111-139.
- SULLET-NYLANDER, F.** (2006) "Paratexte, contexte et intertexte dans *Le Monde* (2005) : Conditionnels journalistiques et discours rapportés." in *Pratiques*, n°129-130. Metz : CRESEF. pp. 123-138.
- SUEUR, J.-P.** (1968) *Étude de la structure syntaxique de titres de journaux*. Mémoire d'études et de recherches en vue de l'obtention d'une maîtrise d'enseignement de Lettres Modernes. Paris : Université de Paris Nanterre.

- SZONDI, P.** (1989) *Introduction à l'herméneutique littéraire*. Paris : Ed. du Cerf.
- TAÏEB, E.** (2006) "La 'rumeur' des journalistes." in *Diogène*, n° 213: "Rumeurs et légendes urbaines". Paris : PUF. pp. 133-152.
- TÉTU, J.-F.** (2002) "L'Analyse de discours en France." in VIALON, P. & WEILAND, U. (Dir.), *Kommunikation Medien Gesellschaft. Eine Bestandsaufnahme deutscher und französischer Wissenschaftler*. Berlin : Avinus Verlag, pp. 205-217. Disponible sur <http://hal.archives-ouvertes.fr/index.php?halsid=i939rkkd9uieh22grosqo82n0&view_this_doc=halshs-00396398&version=1> (Consulté le 12/10/2009).
- TÉTU, J.-F.** (2008) "L'illustration de la presse au XIXème siècle." in LETHIER, V. & VIPREY, J.M. (Dir.), *Semen*, n°25 : "Le discours de presse au XIXe siècle : pratiques socio-discursives émergentes". Besançon : Presses Universitaires de Franche-Comté. pp. 49-73.
- THÉRENTY, M.-E.** (2001) "Physiologie du chercheur sur la presse." in THÉRENTY, M.-E. (Dir.), *1836, L'An I de l'ère médiatique*. Paris : Nouveau monde éditions, pp. 320-326.
- THÉRENTY, M.-E.** (2003) "Pour une histoire littéraire de la presse." in *Revue d'histoire littéraire de la France*, n°3. Paris : Presses Universitaires de France. pp. 625-636.
- THÉRENTY, M.-E.** (2007) *La littérature au quotidien : poétiques journalistiques au XIX^e siècle*. Paris : Editions du Seuil.
- THÉRENTY, M.-E. & VAILLANT, A. (Dir.)** (2001) *1836, L'an I de l'ère médiatique. Analyse littéraire et historique de la Presse de Girardin*. Paris : Nouveau Monde Editions.
- THÉRENTY, M.-E. & VAILLANT, A. (Dir.)** (2004) *Presse et Plumes. Journalisme et littérature au XIXe siècle*. Paris : Nouveau Monde Editions.
- THOMPSON, E.P.** (1963) *The Making of the English Working Class*. New-York : Pantheon.
- TODOROV, T.** (1978) *Les genres du discours*. Paris : Seuil.
- TODOROV, T.** (1981) *Mikhaïl Bakhtine, le principe dialogique*. Paris : Seuil.
- TOGEBY, K.** (1982) *Grammaire française*. Volume II. Copenhague : Akademisk Forlag.
- TOPALOV, C.** (2008) "En finir avec la société? Un débat historiographique." in *Revue d'histoire du XIX^e siècle*, n°37, 2. Paris : Société d'histoire de la révolution de 1848. pp. 167-182.
- TORRÈS, F.** (1995) *Un journal dans sa région* Paris : Albin Michel.
- TORRÈS, F.** (2002) *La Dépêche du midi : histoire d'un journal en république : 1870-2000*. Paris : Hachette.

- TOURNIER, M.** (1975) *Un vocabulaire ouvrier en 1848. Essai de lexicométrie.* Thèse de Doctorat en Sciences du langage. 4 vol. Saint-Cloud : Publications de l'ENS.
- TOURNIER, M.** (1980a) "D'où viennent les fréquences de vocabulaire ?" in *Mots*, n°1. Lyon : ENS Editions. pp. 189-212.
- TOURNIER, M.** (1980b) "En souvenir de Lagalo." in *Mots*, n°1. Lyon : ENS Edition. pp. 5-9.
- TOURNIER, M.** (1985) "Textes propagandistes et cooccurrences. Hypothèses et méthodes pour l'étude de la sloganisation " in *Mots*, n°11. Lyon : ENS Editions. pp. 155-187.
- TOURNIER, M.** (1998) "Des mots en histoire." in BEAUVOIS, Y. & BLONDEL, C. (Dir.), *Qu'est-ce qu'on ne sait pas en histoire ?* Lille : Presses universitaires du Septentrion, pp. 131-143.
- TOURNIER, M.** (2003) "Le label, entre usage et concept." in GUILHAUMOU, J. & MONNIER, R. (Dir.), *Des notions-concepts en révolution.* Paris : Société des études robespierristes, pp. 7-8.
- TRÉNARD, L.** (1989) "Les périodiques lyonnais dans les luttes révolutionnaires (1792-1794)." in RÉTAT, P. (Dir.), *La Révolution du Journal.* Lyon : Editions du CNRS, pp. 297-309.
- UTARD, J.-M.** (2004) "Journalisme et publicité, entre rationalités techniques et talent personnel. Quelle place pour les 'savoirs savants'?" in *Études de communication*, n°27. Disponible sur <<http://edc.revues.org/index202.html>> (Consulté le 18/06/2008).
- VAN DEN DUNGEN, P.** (2008) "Ecrivains du quotidien : Journalistes et journalisme en France au XIXe siècle." in LETHIER, V. & VIPREY, J.M. (Dir.), *Semen*, n°25 : "Le discours de presse au XIXe siècle : pratiques socio-discursives émergentes". Besançon : Presses Universitaires de Franche-Comté. pp. 13-29.
- VAN DIJK, T.A.** (1981) "Le texte : structures et fonctions. Introduction élémentaire à la science du texte." in VARGA, A.K. (Dir.), *Théorie de la littérature.* Paris : Picard.
- VAN DIJK, T.A.** (1982) "Relevance in Text and Context." in ALLEN, S. (Dir.), *Text Processing.* Stockholm : Almqvist & Wiksell, pp. 415-432.
- VAN DIJK, T.A.** (1984) "Texte." in BEAUMARCHAIS, J.P., COUTY, D., et al. (Dir.), *Dictionnaire des littératures de langue française.* Paris : Bordas.
- VAN DIJK, T.A.** (1985) "Structures of news in the press." in VAN DIJK, T.A. (Dir.) *Discourse and communication.* Berlin-New York: de Gruyter. pp.69-93.
- VAN DIJK, T.A.** (1985) (Dir.) *Discourse and communication.* Berlin-New-York : de Gruyter.

- VAN DIJK, T.A.** (1995) "Ideological Discourse Analysis." in VENTOLA, E. & SOLIN, A. (Dir.), *New Courant*, Vol. 4 : "Special issue Interdisciplinary approaches to Discourse Analysis". Helsinki : English Dept, University of Helsinki. pp. 135-161.
- VANDENDORPE, C.** (1999) *Du papyrus à l'hypertexte : essai sur les mutations du texte et de la lecture*. Paris : La Découverte.
- VAUELLES (DE), L.** (1971) *Le "Nouvelliste de Lyon" et la défense religieuse (1879-1889)*. Paris : Les Belles-Lettres.
- VÉRON, E.** (1981) *Construire l'événement : les médias et l'accident de three mile island*. Paris : Editions de Minuit.
- VÉRONIS, J.** (2004a) "Hyperlex : lexical cartography for information retrieval." in *Computer, Speech and language*, 18, 3. London : Elsevier. pp. 223-252.
- VÉRONIS, J.** (2004b) "Quels dictionnaires pour l'étiquetage sémantique? " in *Le français Moderne*, 72,1. pp. 27-38.
- VEYNE, P.** (1979) *Comment on écrit l'histoire*. Paris : Seuil.
- VILLEMESSANT (DE), H.** (1873) *Mémoires d'un journaliste*. Paris : Dentu.
- VIPREY, J.-M.** (1997) *Dynamique du vocabulaire des "Fleurs du mal"*. Paris : Seuil.
- VIPREY, J.-M.** (2004) "Analyse séquencée de la micro-distribution lexicale. " in PURNELLE, G. & al. (éds.), *Actes des 7es Journées internationales d'analyse statistique des données textuelles (JADT 2004)*. Louvain : Presses universitaires de Louvain. pp.1165-1176.
- VIPREY, J.-M.** (2005a) "Corpus et sémantique discursive : éléments de méthode pour la lecture des corpus " in CONDAMINES, A. (Dir.), *Sémantique et corpus*. Paris : Hermès, pp. 245-276.
- VIPREY, J.-M.** (2005b) "Philologie numérique et herméneutique intégrative." in ADAM, J.M. & HEIDMANN, U. (Dir.), *Sciences du texte et analyse de discours. Enjeux d'une interdisciplinarité*. Paris : Etudes de Lettres. pp. 51-58.
- VIPREY, J.-M.** (2006a) "Un de ces [syntagmes] qui (à propos de la locution un(e) de ces [...] qui)." in *Corpus*, n°5 : "Corpus et stylistique". Disponible sur <<http://corpus.revues.org/document713.html>> (Consulté le 12/10/2009).
- VIPREY, J.-M.** (2006b) "Quelle place pour les sciences des textes dans l'analyse de discours?" in *Semen*, n°21: "Catégories pour l'analyse du discours politique" Besançon : Presses Universitaires de Franche-Comté. pp.167-183.

- VIPREY, J.-M.** (2006c) "Ergonomiser la visualisation AFC dans un environnement d'exploration textuelle : une projection 'géodésique'." in VIPREY, J.-M. (éds.), *Actes des 8es Journées internationales d'Analyse statistique des Données Textuelles (JADT 2006)*. Besançon : Presses Universitaires de Franche-Comté. pp. 981-992.
- VIPREY, J.-M.** (2006d) "Structure non séquentielle du texte." in *Langages* n°163 : "Unités du texte". Paris : Larousse. pp. 71-85.
- VIPREY, J.-M. & ZELLAGUI, K.** (2000) "Problèmes soulevés par l'étiquetage des textes littéraires." in *Revue Informatique et Statistique dans les Sciences humaines*. Disponible sur <<http://promethee.philo.ulg.ac.be/RISSHpdf/annee2000/Articles/KZellaguietc.pdf>> (Consulté le 12/10/2009).
- VIPREY, J.-M. & LETHIER, V.** (2008a) "Annotation linguistique de corpus : vers l'exhaustivité par la convivialité" in HEIDEN, S & PINCEMIN, B. (éds.), *Actes des 9es Journées internationales d'Analyse statistique des Données Textuelles (JADT 2008)*. Lyon : Presses Universitaires de Lyon. pp.1151-1161.
- VIPREY, J.-M. & LETHIER, V.** (2008b) "*Lire l'archive : presse écrite du XIXe siècle.*" Disponible sur <http://web.univ-ubs.fr/corpus/jlc5/ACTES/ACTES_JLC07_lethier_viprey.pdf> (Consulté le 12/10/2009).
- VISSE, J.-P.** (2004) *La presse du Nord et du Pas-de-Calais au temps de l'Écho du Nord*. Villeneuve d'Ascq : Presses universitaires du Septentrion.
- VOGNE, M.** (1977 [1972]) *La Presse périodique en Franche-Comté des origines à 1970*. (7 volumes). Paris : Les Belles Lettres.
- WALTER, E.** (1983) *La Presse picarde, mémoire de la République : luttes de mémoire et guerres scolaires à travers la presse de la Somme : 1876-1914*. Amiens : Centre universitaire de recherche sociologique.
- WARTELLE, F.** (1984) "Honnêtes gens en l'an IV : Les mots, enjeu et reflet des luttes politiques." in *Mots*, n°9. Lyon : ENS Editions. pp. 167-188.
- WARTELLE, F.** (1985) "Honnêtes-gens. La dénomination comme enjeu des luttes politiques (1795-1797)." in *Dictionnaire des usages socio-politiques (1770-1815)*. Fasc. 1 : "Désignants socio-politiques." Paris : Champion. pp. 93-138.
- WEILL, G.** (1934) *Le Journal. Origines, évolution et rôle de la presse périodique*. Paris : La Renaissance du livre.
- WEINRICH, H.** (1989) *Grammaire textuelle du français*. Paris : Didier.
- WHITE, H.** (1973): *Metahistory: The Historical Imagination in Nineteenth-Century Europe*. Baltimore, MD: Johns Hopkins University Press.

- WHITE, H.** (1978): *Tropics of Discourse: Essays in Cultural Criticism*". Baltimore, MD : Johns Hopkins University Press.
- WHITE, H.** (1987) *The Content of the Form: Narrative Discourse and Historical Representation*". Baltimore, MD: Johns Hopkins University Press.
- WIERZBICKA, A.** (1987) *English Speech Act Verbs. A Semantic Dictionary*. Sydney : Academic Press.
- WILL, N.** (1976) *Essai sur la presse et le capital*. Paris : Union générale d'Éditions.
- WILLIAMS, G. (Dir.)** (2005) *La linguistique de corpus*. Rennes : Presses Universitaires de Rennes.
- WILLIAMS, G. & ROPERS, C.** (2007) "Textes électroniques et corpus : pourquoi utiliser les normes de la TEI ?" in HÉDIARD, M. **(Dir.)**, *Linguistica dei corpora : strumenti e applicazioni*. Cassino : Edizioni Università di Cassino, pp. 23-44.
- WINOCK, M.** (1992) "Le mythe fondateur : l'affaire Dreyfus." in BERSTEIN, S. & RUDELLE, O. **(Dir.)**, *Le modèle républicain*. Paris : Presses Universitaires de France, pp. 131-146.
- WOGAN (DE), T.** (1898) *Manuel des gens de lettres : le journal, le livre, le théâtre*. Paris : Firmin-Didot.
- TEI**, <http://www.tei-c.org>
- TEI P5**, <http://www.tei-c.org/P5>

Table des matières

[TOME 1]

INTRODUCTION GÉNÉRALE	1
CHAPITRE I.	
EXPLORER LA PRESSE RÉGIONALE DE LA TROISIÈME	
RÉPUBLIQUE : ÉTAT DE L'ART	9
1.1. Approches de la presse régionale de la Troisième République	
par les sciences humaines.....	10
1.1.1. Un champ de recherche peu investi par les sciences de	
l'information et de la communication	
ainsi que par l'analyse du discours	10
1.1.1.1. En analyse du discours.....	10
1.1.1.2. En sciences de l'information et de la communication	12
1.1.2. Une presse essentiellement étudiée par les sciences	
historiques	13
1.1.2.1. Bref panorama des recherches en sciences historiques	
sur et/ou par la presse régionale de la Troisième République	14
1.1.2.2. Les facteurs qui accentuent le retard de développement	
de l'étude de la presse régionale de la Troisième République	19
1.2. Le discours et sa dimension matérielle.....	24
1.2.1. Regards sur le <i>linguistic turn</i>	24
1.2.1.1. Une remise en cause des moyens	
de la connaissance historique.....	25
1.2.1.2. La controverse sur l'objectivité et la vérité.....	28
1.2.1.3. <i>L'histoire des concepts</i>	30
1.2.2. <i>L'analyse du discours du côté de l'histoire</i>	34

1.2.2.1. Retour sur la démarche inaugurale de l'analyse du discours du côté de l'histoire	35
1.2.2.2. La démarche configurationnelle	39
1.2.2.3. Vers l'histoire linguistique des concepts.....	43
1.2.3. L'analyse textuelle du discours	47
1.2.4. La dynamisation de notre champ de recherche est conditionnée par la constitution de vastes ensembles de données textuelles d'intérêt historique	51
1.2.4.1. Peu d'expériences inaugurales prennent la PQR3R pour objet	51
1.2.4.2. Retour sur des entreprises pionnières de constitution de base textuelle de presse ancienne	53
1.3. Construire une base textuelle de presse ancienne : aspects méthodologico-techniques	56
1.3.1. Constitution de données textuelles	57
1.3.1.1. Renforcer le statut scientifique de la presse ancienne comme ressource	57
1.3.1.2. Assurer portabilité et pérennité aux données textuelles	60
1.3.1.3. Le XML-TEI	62
1.3.2. La numérisation de la presse ancienne.....	66
1.3.2.1. La numérisation en mode image.....	67
1.3.2.2. La numérisation en mode texte	71
Conclusion.....	82
CHAPITRE II.	
CONSTITUTION DU DISPOSITIF D'EXPLORATION DU <i>PETIT COMTOIS</i>.....	85
2.1. Du support papier à la base textuelle.....	86
2.1.1. Brève présentation du projet de base textuelle PQR3R.....	86
2.1.2. Acquisition des données image.....	88
2.1.3. Techniques d'aide à l'OCR	96
2.1.4. Océrisation et correction automatique, semi-automatique et manuelle	100
2.1.4.1. Ouverture et traitement des images.....	101
2.1.4.2. Enregistrement du texte reconnu	102
2.1.4.3. Options de reconnaissance optique des caractères	102
2.1.4.4. Qualité de l'océrisation effectuée	102
2.1.4.5. Vérification et correction orthographique.....	105

2.1.4.6. La correction manuelle.....	108
2.1.5. Documenter les données	108
2.1.6. Établissement et normalisation des données textuelles : vers une philologie numérique	113
2.1.6.1. XML-TEI et philologie numérique.....	114
2.1.6.2. Encodage des données du <i>Petit Comtois</i> en XML-TEI.....	117
2.1.7. Bilan et perspectives	124
2.2. De la base textuelle au corpus	130
2.2.1. Retour sur notre projet de recherche	131
2.2.1.1. Bref historique des débuts de la Troisième République (1883-1903)	132
2.2.1.2. Un contexte de mutation de la pratique journalistique	140
2.2.2. Présentation du <i>Petit Comtois</i>	145
2.2.2.1. Contexte général de parution	145
2.2.2.2. <i>Le Petit Comtois</i> : un exemple typique de journal régional sous la Troisième République ?.....	148
2.2.3. Description du corpus	159
Conclusions partielles.....	173
 CHAPITRE III.	
LA MISE EN FORME DE L'INFORMATION	
DANS LE PETIT COMTOIS	
	175
3.1. Description matérielle de l'objet imprimé	176
3.2. Le rubriquage du <i>Petit Comtois</i> (1883-1903).....	185
3.2.1. Approche générale du rubriquage	185
3.2.1.1. Retour sur l'étymologie du mot « rubrique ».....	185
3.2.1.2. Le point de vue des théoriciens sur la rubrique	187
3.2.2. Approche quantitative des rubriques du <i>Petit Comtois</i>	189
3.2.2.1. Méthodologie d'analyse quantitative des rubriques	189
3.2.2.2. La « cuisine » du <i>Petit Comtois</i>	196
3.2.3. Des critères sous-jacents au système de rubriquage qui se révèlent composites	211
3.2.4. Remarques conclusives sur le rubriquage du <i>Petit Comtois</i>	214
3.3. Formes et fonctions des titres.....	218
3.3.1. Approche générale du titre	218
3.3.1.1. Retour sur les fonctions du titre	219
3.3.1.2. Lectures du titre	220
3.3.2. Les titres dans <i>Le Petit Comtois</i>	223

3.3.3. Analyse des titres de la rubrique des « <i>Dépêches de nuit</i> »	229
3.3.3.1. Méthode de classification des titres.....	230
3.3.3.2. Aperçu sur les caractéristiques formelles générales des titres en diachronie.....	245
3.3.3.3. Répartition des fonctions des titres dans le sous-corpus.....	247
3.3.3.4. Remarques conclusives sur l'appareil titulaire des <i>Dépêches de nuit</i>	297
Conclusions sur la mise en forme de l'information dans <i>Le Petit Comtois</i>	299

[TOME 2]

CHAPITRE IV. EXPLORATIONS DU VOCABULAIRE DU <i>PETIT COMTOIS</i>	304
4.1. Retour sur la pratique d'analyse statistique textuelle.....	305
4.1.1. Lexicométrie, statistique linguistique, analyse des données textuelles.....	305
4.1.1.1. Une pratique interprétative objectivée du discours.....	306
4.1.1.2. Esquisse d'une praxis lexicométrique	308
4.1.2. Vers une herméneutique numérique	313
4.1.2.1. Une approche statistique des unités linguistiques plurielles du texte	314
4.1.2.2. Une approche statistique syntagmatique	317
4.1.2.3. Une statistique textuelle aux prises avec le sémantique	318
4.1.3. Les outils d'analyse textuelle assistée par informatique	320
4.1.3.1. Les outils disponibles	320
4.1.3.2. Astartex.....	324
4.2. Manipulations pré-exploratoires sur le vocabulaire du <i>Petit Comtois</i>	333
4.2.1. Consultation de l'index hiérarchique	334
4.2.2. Examen de la distribution du vocabulaire du <i>Petit Comtois</i> en diachronie.....	337
4.2.2.1. Analyse de la distribution du vocabulaire du <i>Petit Comtois</i> en diachronie.....	337
4.2.2.2. Examen du vocabulaire propre des séquences chronologiques.....	345

4.2.3. Remarques conclusives sur l'examen de la ventilation du vocabulaire en diachronie.....	378
4.3. Analyse de la micro-distribution du vocabulaire du <i>Petit Comtois</i>	379
4.3.1. Description de la configuration fine du vocabulaire du corpus	380
4.3.2. Synthèse sur la configuration fine du vocabulaire	403
4.4. Analyse de la structure fine du vocabulaire en diachronie.....	404
4.4.1. Premières expérimentations.....	404
4.4.2. Comparaison simultanée	410
4.4.3. Remarques conclusives provisoires	428
4.5. Le rubriquage, facteur structurant de la configuration fine du vocabulaire.....	429
4.5.1. Méthodologie d'analyse des dominantes lexico-thématiques des rubriques.....	429
4.5.2. Validation de la corrélation entre le système de rubriquage et la configuration fine du vocabulaire	432
Conclusion.....	444

CHAPITRE V.

APPROCHES DU NIVEAU MORPHOSYNTAXIQUE SUR LE CORPUS *PETIT COMTOIS*.

ESSAI DE TYPOLOGISATION TEXTUELLE.....	448
---	------------

5.1. Bref cadrage de la pratique de typologisation textuelle à partir de variables morphosyntaxiques.....	450
--	------------

5.1.1. Méthodes inductives et analyses statistiques des données	450
5.1.2. Genres et typologies textuelles.....	453

5.2. Méthodologie générale	454
---	------------

5.2.1. Recueil et sélection des variables	455
5.2.2. Remarques sur l'étiquetage effectué par <i>Cordial</i>	457
5.2.3. Méthodologie	472

5.3. Analyse de la distribution des catégories grammaticales dans les rubriques du <i>Petit Comtois</i>	475
--	------------

5.4. Analyse de la distribution d'un panel large de variables morphosyntaxiques dans les rubriques du <i>Petit Comtois</i>	482
---	------------

5.4.1. Analyse de la distribution de 52 variables morphosyntaxiques dans les rubriques du <i>Petit Comtois</i>	483
5.4.2. Description des facteurs fortement discriminants	486
5.4.3. Synthèse sur les facteurs discriminants	511
5.4.4. Caractérisation morphosyntaxique des rubriques	513
5.4.4.1. Les dépêches	513
5.4.4.2. Les chroniques régionales.....	520
5.4.4.3. Les rubriques <i>Petite chronique, Chronique, Chronique de la semaine, Lettres hebdomadaires</i>	528
5.4.4.4. Les rubriques <i>Faits Divers, Tribunaux, Variétés, Causerie</i>	532
5.4.4.5. <i>Courrier de la mode, Chronique parisienne, Revue agricole, Agriculture, Bibliographie, Bulletin</i>	537
Conclusion	543
CONCLUSION GENERALE	545
BIBLIOGRAPHIE	I
INDEX	TOME 3
ANNEXES	TOME 3
TABLE DES FIGURES	TOME 3

Exploration textuelle du discours d'un quotidien régional de presse au carrefour des XIX^e et XX^e siècles : Le *Petit Comtois* (1883-1903)

Mots-clés : analyse du discours, presse, base, corpus, textométrie, numérisation.

À la différence des tendances actuelles de l'analyse du discours médiatique, majoritairement tournée vers une presse du XXI^e siècle nationale, cette thèse vise un titre régional du XIX^e siècle : *Le Petit Comtois* (1883-1944). Plongeant ses racines dans différents terrains de recherche, à savoir la linguistique de corpus, l'analyse du discours du côté de l'histoire (Guilhaumou, Maldidier, Robin), la statistique textuelle (Lebart, Salem) et l'analyse textuelle du discours (Adam & Heidmann), cette thèse interroge les apports d'une *nouvelle philologie numérique* (Rastier ; Viprey) se fixant pour programme de renforcer l'accès aux matérialités discursives.

Cette thèse articule ainsi deux axes de recherche complémentaires : d'une part, les conditions de la constitution d'une base de données textuelles finement établies, contrôlées et normalisées selon les recommandations XML-TEI, visant à favoriser le renouveau des lectures interdisciplinaires de la presse régionale de la Troisième République. Il s'agit, ce faisant, de proposer des pistes méthodologiques en vue de la levée des obstacles techniques pesant sur l'acquisition et le partage de grands volumes de données textuelles, dont dépendent la confrontation heuristique des points de vue, des outils et des méthodes des sciences des textes.

D'autre part, notre objectif est d'apporter un éclairage sur le discours du *Petit Comtois* (1883-1903) à partir d'un corpus de 5,5 millions de mots en faisant dialoguer diverses méthodologies et niveaux d'analyse, de sorte à appréhender dans toute l'hétérogénéité et la complexité du *texte*. Si notre recherche s'appuie principalement sur les outils de la statistique textuelle qui permettent de faire monter du corpus lui-même ses reliefs linguistiques, textuels, discursifs, envisagés comme des vecteurs d'exploration, elle vise à dépasser l'entrée du vocabulaire sur laquelle se fonde la lexicométrie, et ses modes d'approche traditionnels sur cette entrée même.

Textual exploration of the discourse of a French newspaper at the meeting point of the 19th and the 20th century : *Le Petit Comtois* (1883-1903).

Keywords: discourse analysis, newspapers, archive, corpus, lexical statistics, digitalization.

In contrast to topical trends of mediatic discourse analysis (predominantly focused on newspapers from the 21st century), this thesis aims at analyzing the discourse of a French newspaper called *Le Petit Comtois* (1883-1944). Rooted in various fields of research, namely corpus linguistics, "discourse analysis from a historical viewpoint" (Robin, Maldidier, Guilhaumou), "textual analysis of discourse" (Adam, Heidmann) and lexical statistics (Lebart, Salem), this thesis questions the contributions of a *new numerical philology* (Rastier; Viprey). This work is based on two complementary axes of research: on the one hand, we worked on the constitution of a textual database with the aim of accurately editing, controlling and normalizing it in order to allow for a revival of interdisciplinary studies led on regional 19th century newspapers. Therefore, our point is to propose some methodological approaches in order to overcome technical obstacles which occur during the acquisition and the sharing of large volume of textual data. The construction and the normalization of the database condition the heuristic confrontation of points of views and the choice of textual sciences methods for the analysis.

On the other hand, our objectives are to highlight the specificities of *Le Petit Comtois* discourse during the period 1883-1903. Exploring a 5.5 million-word corpus, we mainly use the textual statistical methods to produce the emergence of the linguistic, textual and discursive reliefs of the text itself in order to use them as clues for the continuation of the analysis, with the aim to overtake the traditional entrance of the vocabulary used in lexicometry and its ways to approach textuality.