



HAL
open science

Toutes choses égales par ailleurs. Comparer deux congrès de l'Association française de sociologie

Didier Torny, Patrick Trabal

► To cite this version:

Didier Torny, Patrick Trabal. Toutes choses égales par ailleurs. Comparer deux congrès de l'Association française de sociologie. Bulletin de Méthodologie Sociologique / Bulletin of Sociological Methodology, 2007, 94, pp.57-75. halshs-01678177

HAL Id: halshs-01678177

<https://shs.hal.science/halshs-01678177>

Submitted on 9 Jan 2018

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

Toutes choses égales par ailleurs.
Comparer deux congrès de l'Association française de sociologie

Didier Torny, Patrick Trabal

Bulletin de méthodologie sociologique, 94 2007

<http://journals.openedition.org/bms/476>

L'objet de cet article est de questionner la pertinence d'une comparaison de corpus, d'en délimiter les conditions et d'en saisir l'interprétation. Que fait-on lorsqu'on décide d'isoler une variable ou une série de variables et de comparer, toutes choses égales par ailleurs ? Pour répondre à cette question, nous avons choisi comme matériel les résumés des communications présentées aux congrès de l'Association Française de Sociologie, qui ont déjà fait l'objet de plusieurs analyses. Un premier numéro du BMS (n°85, janvier 2005) présentait des analyses réalisées avec quatre logiciels. Nous explicitons dans ce numéro les modalités du travail avec Prospéro (Trabal, 2005). Un ouvrage collectif (Demazière et al., 2006) donnait aux mêmes auteurs la possibilité de pousser plus loin l'analyse. A cette occasion, nous proposons de construire une série d'actes de congrès qui permettrait de mettre en histoire et de saisir les évolutions du travail sociologique (Trabal, 2005, Torny et Trabal, 2006). Le deuxième congrès de 2006 donne une première opportunité de mettre en perspective les deux corpus et les auteurs de l'ouvrage précédemment cité ont déjà livré quelques éléments comparatifs (Boudesseul, 2006, Van Meter et De Saint Léger, 2007).

Mais plus que de simplement comparer les congrès 2004 et 2006, il s'agit pour nous d'aborder la question même de la comparabilité : dans quel espace commun projette-t-on les deux corpus, et complémentaiement, quelles règles, indices et mesures servent à représenter leurs différences ? Nous avons choisi de suivre trois stratégies comparatives distinctes, dont les objectifs sont mis en œuvre à l'aide d'algorithmes bien différents : on établira tout d'abord une comparaison structurelle entre les deux corpus. Dans quelle mesure compare-t-on des productions issues des mêmes professionnels ou, du moins, issus des mêmes collectifs ? La population des auteurs, a-t-elle évolué et dans quel sens ? Après avoir répondu à ces questions, on travaillera ensuite la question de la comparaison lexicale entre les deux congrès : quelles sont les nouveautés apparues en 2006 ou quels sont, à l'inverse, les disparitions et quel sens donner à ces variations ? Enfin, on cherchera à tester la permanence de quelques propriétés fondées sur une analyse comparative entre trois congrès de sociologie antérieurs, se basant sur des ensembles d'indices complexes

UNE COMPARAISON STRUCTURELLE : QUI PARLE A L'AFS ?

Dans cette partie, la comparaison vise à saisir, à partir de variables portant sur les auteurs et la structure du congrès, des évolutions entre les deux congrès¹. On ne s'intéresse donc pas ici directement au contenu des résumés, mais aux propriétés de ceux qui les portent et des lieux où ils sont rassemblés, ce qui suppose de s'appuyer sur des variables les décrivant. Certaines

¹ Les deux congrès apparaissent comparables au premier abord avec un nombre de communications du même ordre de grandeur : 1.067 en 2004 et 1.190 en 2006. Pour ce dernier, on a retiré du corpus de départ un résumé « vide » et deux résumés écrits en langue anglaise.

sont directement données dans l'objet « résumé de communication » ; d'autres sont partiellement renseignées, voire demandent des enquêtes spécifiques. Les choix associés à chaque type de variable conditionnent fortement l'interprétation : par exemple, enquêter sur le lieu de formation des personnes (thèse actuelle ou thèse passée) vise à reconstituer des trajectoires et/ou mesurer la hiérarchie des établissements formateurs dans la profession. En revanche, privilégier les variables explicitement partagées par l'institution et les auteurs, comme les réseaux thématiques, invite à suivre les politiques scientifiques de l'AFS, en travaillant par exemple sur la question du taux de sélection ou la capacité d'attraction d'une thématique donnée. Pour notre part, nous avons choisi de nous appuyer sur quatre variables : l'identité du premier auteur, son genre, son statut professionnel et le réseau thématique dans lequel la communication est présentée. Chacune de ces variables pose des problèmes de codage spécifique que nous allons exposer rapidement ; nous indiquerons alors les résultats comparatifs obtenus et leur interprétation.

Nous avons d'abord fait le choix de ne retenir que le premier auteur de la communication, et non l'ensemble des coauteurs. Cette réduction a un effet relativement marginal : en effet, comme pour la majorité des articles écrits (Pontille, 2004), les communications en sociologie n'émanent que rarement d'un collectif d'auteurs - 88% des communications du congrès 2006 ont un unique auteur signataire. Pour distinguer ces auteurs entre eux, on retient leurs noms et prénoms déclarés, limitant ainsi les cas d'homonymie² Par ailleurs, lorsqu'ils présentent deux communications, il peut arriver que les informations entrées sur le site de l'AFS soient lexicalement divergentes³. En prenant en compte ces deux facteurs, on parvient au résultat suivant en distinguant les auteurs présents aux deux congrès des autres :

	AFS 2004	AFS 2006
Premiers auteurs communs	264	264
Premiers auteurs spécifiques	675	766
Total	939	1030

Tableau 1 Répartition des premiers auteurs suivant leur présence aux congrès

A notre surprise, nous constatons donc qu'il n'y a que 25% de premiers auteurs communs entre les deux congrès, ce qui modifie bien évidemment la nature des interprétations comparatives que l'on tire dans la suite de cet article : les stabilités entre les deux congrès sont per se significatives d'une certaine permanence disciplinaire au-delà des auteurs, alors qu'il faudra systématiquement vérifier que les différences ne résultent pas simplement de l'arrivée de nouveaux entrants.

Notre deuxième variable, le genre des auteurs, pose d'autres problèmes de codage : contrairement aux congrès de l'AFS, il n'est en effet pas explicitement déclaré par les auteurs. On peut estimer que cette situation évoluera, étant donné que le nouveau bureau de l'AFS élu au terme d'une discussion houleuse lors de l'Assemblée Générale du congrès de Bordeaux, a décidé de « veiller à la mise en oeuvre de la parité entre hommes et femmes à l'AFS »⁴, ce qui suppose la déclaration du genre des auteurs. En son absence, nous avons reposé sur trois indices distincts pour le coder : tout d'abord, sur le statut déclaré par les premiers auteurs (voir plus bas), puisque 319 d'entre elles l'ont sexué (« Doctorante »,

² Il y a par exemple sept auteurs différents (cinq femmes et deux hommes) nommés Martin dans le corpus AFS 2006.

³ Par exemple, des noms composés avec et sans tiret, ou des noms propres avec et sans accent.

⁴ Site de l'AFS, <http://www.afs-socio.fr>, page vue pour la dernière fois le 31 mars 2007.

« Maîtresse de conférences », « Chercheuse »...), pratique qui se répand sur de nombreux marchés du travail (Marchal & Torny, 2003). Ensuite, nous nous sommes appuyés sur le codage manuel effectué sur le corpus 2004 pour les auteurs communs, enfin sur un codage manuel pour les nouveaux auteurs, s'appuyant sur le genre des prénoms et sur une recherche web en cas de prénoms mixtes⁵. Sur la base de ces trois indices, on obtient les résultats suivants pour les premiers auteurs, qui sont similaires à ceux présentés par Boudesseul (2006), en montrant une inversion du rapport Homme/Femme des auteurs, au profit de ces dernières :

Genre	AFS 2004	AFS 2006
Femme	452	540
Homme	487	490
Tous	939	1030

Tableau 2 Répartition des premiers auteurs par genre

Néanmoins, la variation de ce rapport n'est pas si importante, surtout quand on la rapporte au nombre de nouveaux entrants. Il faut croiser cette information avec d'autres variables pour pouvoir l'interpréter plus en avant. Or, pour le congrès 2006, on dispose d'une variable supplémentaire, renseignée de manière presque exhaustive : le statut des auteurs des communications. Cependant, ce statut a été rempli sous la forme d'un champ libre lors de l'inscription au congrès ; aussi on constate une hétérogénéité importante des données introduites avec 404 statuts différents déclarés.

Aussi, on a procédé à un recodage en trois modalités, en veillant à limiter au maximum les cas ambigus⁶ : Doctorant (en incluant les rares « DEA » et « M2 »), Docteur/Post-doc en incluant les personnes se déclarant sans statut fixe (« Chercheur contractuel », « en recherche d'emploi ») et enfin Permanent pour l'ensemble des personnes déclarant un poste fixe. Ce choix découle d'une double hypothèse : d'une part, nous nous intéressons à la question de la carrière des individus-sociologues, et les trois modalités constituent une trajectoire contemporaine « réussie », d'autre part, nous pensons qu'en termes de conditions de travail/de statut, la distinction entre CDI/poste de fonctionnaire et CDD/précaire est un élément essentiel, plus important par exemple qu'une distinction de type CR/DR ou CR/MDC. Sur les 1030 auteurs, 26 n'ont pas déclaré de statut, on les a écarté dans le tableau suivant qui, signalons-le, est légèrement divergent par rapport aux données de Boudesseul (2006), sans qu'on ait pu déterminer les causes de ces écarts :

Statut 2006	Premiers auteurs
Permanent	429
Doc/postdoc	209
Doctorant	366
Total	1004

Tableau 3 Répartition des premiers auteurs par statut

⁵ Il existe toujours des cas-limites, une collègue se prénommant par exemple Yannick. Dans ce genre de cas, seule la connaissance directe ou indirecte de la personne permet de déterminer son genre. Ce qui nous permet de préciser qu'on entend ici par genre une variable dichotomique Homme/Femme, qui est précisément à l'œuvre lorsque l'AFS évoque la question de la parité.

⁶ Pour cette variable, il est difficile d'effectuer une enquête complémentaire en croisant les sources, le statut retenu étant le statut déclaré au moment de la soumission de la communication, soit six mois avant le congrès et un an avant l'écriture de cet article. On peut supposer qu'en dépit d'un contexte difficile, nombre d'auteurs, en particulier parmi les plus jeunes dans la carrière professionnelle, ont entre-temps changé de statut.

Plus de 55% des premiers auteurs communicants n'ont donc pas de poste fixe, ce qui contraste fortement avec le taux avancé par l'AFS pour ses membres : « Elle compte actuellement plus de 1300 membres regroupés en une quarantaine de Réseaux Thématiques, certains déjà bien stabilisés tandis que d'autres sont encore en formation. Un quart environ de ses membres sont des doctorants ou jeunes docteurs. »⁷ Les deux constats ne sont pas nécessairement incompatibles, mais tendent plutôt à souligner l'importance professionnelle variable de ces communications suivant le statut des auteurs, qui serait plus important pour des personnes en voie d'insertion professionnelle⁸. On peut désormais croiser les deux dernières variables et constater une structuration de statut bien différente suivant le genre des premiers auteurs :

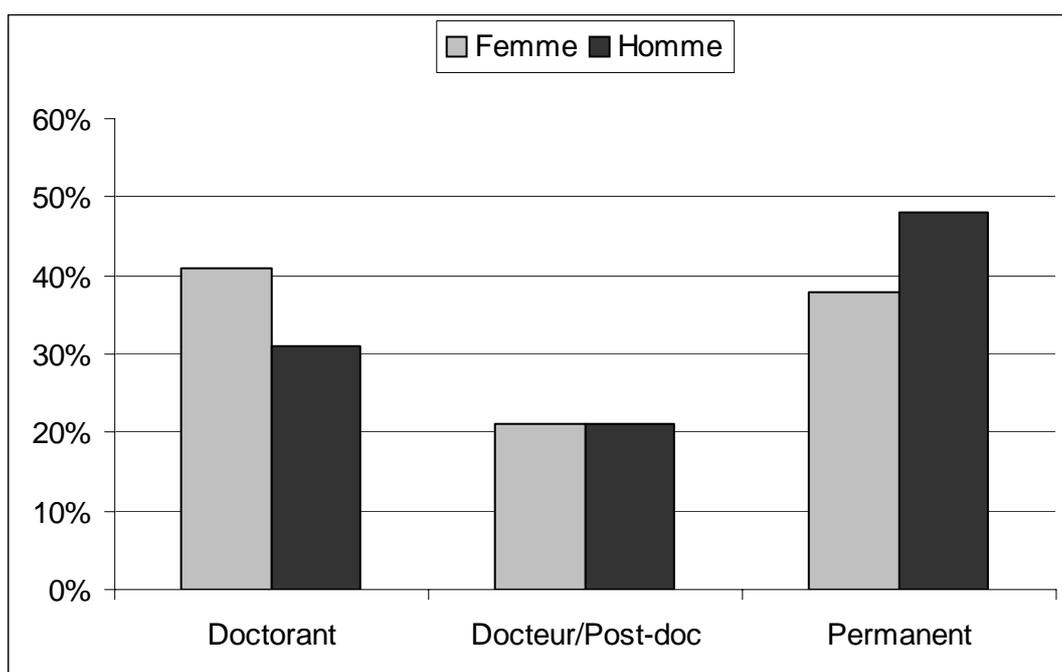


Figure 1 Répartition des premiers auteurs par statut et par genre en 2006

En l'absence de données pour le congrès précédent, on ne peut déterminer si on recueille un phénomène de féminisation de la profession sur une durée moyenne ou de sélection des personnes durant la carrière sur la base du genre. A moins qu'il ne s'agisse d'un phénomène plus récent comme le montrerait l'inversion du rapport de genre décrit plus haut, ou un simple artefact de distribution lié à ce congrès particulier (localisation, date). Nous ne pouvons que souligner une fois de plus la nécessité de constituer des données sur la longue durée pour pouvoir répondre à ce type de questions.

En revanche, si on se focalise sur les premiers auteurs communs aux deux congrès (264), leurs spécificités s'imposent : ils communiquent plus (1,20 communications/ auteur contre 1,14 aux nouveaux auteurs), forment une population beaucoup plus masculine (55% d'hommes), et plus avancée dans la carrière (64% de permanents). On ne peut néanmoins pas en conclure

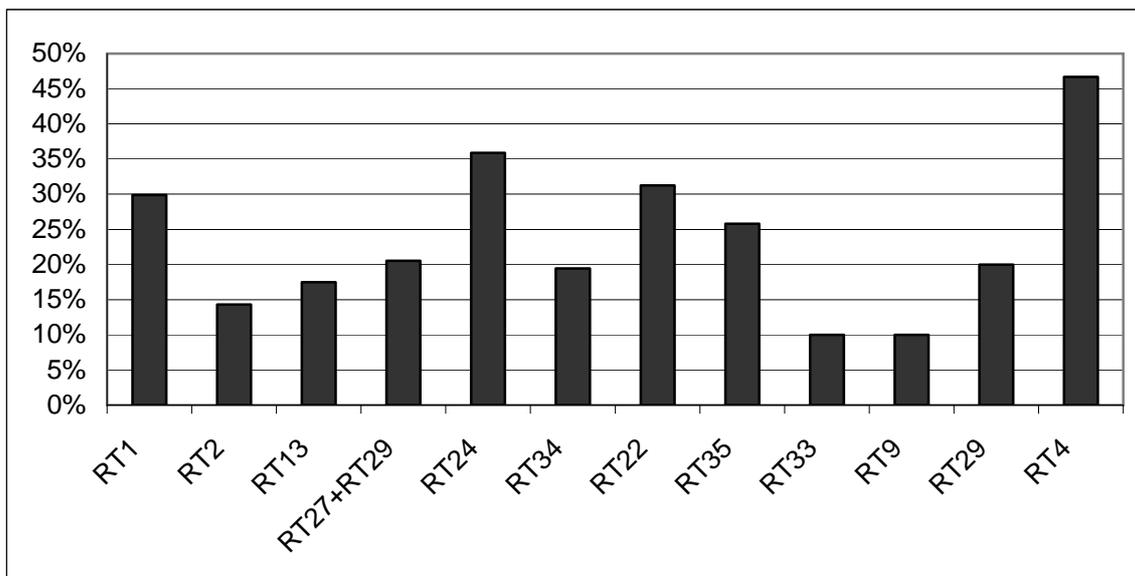
⁷ Site de l'AFS, <http://www.afs-socio.fr>, page vue la dernière fois le 31 mars 2007.

⁸ Un élément plus conjoncturel peut également expliquer ce décalage : le congrès de Bordeaux a lieu la semaine de la rentrée scolaire et à la veille de la rentrée universitaire, étant dès lors en concurrence avec nombre d'obligations professionnelles et familiales, *a priori* plus fortes pour les auteurs au statut *Permanent*.

que seuls les doctorants formeraient, autour d'un noyau stable, les nouveaux entrants, puisqu'on trouve tout de même 34% de permanents dans ce second groupe. Là encore, une analyse sur un plus grand nombre de congrès permettra de déterminer de manière beaucoup plus fine les trajectoires professionnelles des auteurs et savoir si les congrès de l'AFS (et l'appartenance à l'association) deviennent un rendez-vous régulier pour une majorité de sociologues ou rassemble une population professionnelle de manière passagère.

Intéressons-nous maintenant à notre dernière variable structurante : le réseau thématique (RT) dans lequel s'inscrit la communication⁹. Etant donnée la jeunesse de l'institution, on pourrait s'attendre à des modifications importantes de structures en comparant simplement la liste de ces réseaux. Or, on constate plutôt la forte stabilité de la structure des RT mise en place préalablement au congrès de Villeteuse : seuls la disparition du RT18 (Approche plurielle du sujet), l'absorption du RT41 dans le RT 27 dont on rend compte par le nouveau nom de ce dernier (Sociologie des intellectuels et de l'expertise : savoirs et pouvoirs) et la création de deux Groupes ad hoc la modifient. A quoi s'ajoute les changements de nom du RT3 « Sociologie du droit » en « Sociologie du droit et de la justice », celui du RT29 de « Sciences, innovations technologiques et société » en « Sciences et techniques en société : savoirs, pratiques, instruments et institutions » et celui du RT 24 « Travail (productif et reproductif), rapports sociaux, rapport de genre » en « Genre, Classe, Race. Rapports sociaux et construction de l'altérité », dont nous mesurerons le poids plus loin.

Cette image de stabilité structurelle se renforce si on considère maintenant l'importance quantitative de chacun des réseaux. Pour cela, on se détache d'une approche en termes d'auteurs, pour adopter le point de vue des communications (rappelons-le, au nombre de 1190 en 2006). A l'exception du RT1 – Savoirs, Travail, Professions – dont le nombre de communications passe de 53 à 87¹⁰, l'ensemble des RT se situe dans une fourchette allant de 15 à 50 communications. Il faut utiliser la variable « nouvel auteur » pour mesurer des dynamiques contrastées entre RT :



⁹ On ne fait pas ici de distinction entre RT et RTf (RT en formation), les RTf devenant peu à peu RT, à mesure qu'ils valident les conditions d'institutionnalisation édictées par le bureau de l'AFS.

¹⁰ S'agit-il d'une attirance pour les thématiques du RT1 ou d'un choix de sélection différents des autres réseaux ? Il faudrait reposer sur d'autres documents et d'autres indices, comme par exemple le taux d'acceptation des communications.

Figure 2 Pourcentage de premiers auteurs présent dans les deux congrès sur les 12 RT abritant le plus de communications

Alors qu'en moyenne, 30% des communications sont données par des premiers auteurs déjà présents en 2004, on constate sur les RT les plus importants en nombre de communications (dans l'ordre décroissant sur la figure 2, allant de 87 à 30 communications) une distribution très inégale allant de 10% (RT33, RT9) - et donc 90% de nouveaux auteurs - à 47% (RT4). Cette figure nous permet simplement d'indiquer que des dynamiques diverses sont à l'œuvre dans le renouvellement des communications et des auteurs, à l'intérieur d'une stabilité structurelle.

Cette première partie, uniquement fondée sur les propriétés paratextuelles des résumés de communications, doit nous servir de base dans une étude de plus longue durée pour établir différentes propriétés à la fois populationnelles et individuelles de la profession. Néanmoins, on a pu déjà établir que si les corpus sont structurellement et quantitativement comparables, les auteurs qui ont signé les textes sont aux trois quarts différents. Il s'agit maintenant de voir en quoi ce mouvement d'entrée et de sortie a influé sur le contenu même des communications du congrès de Bordeaux.

UNE COMPARAISON LEXICALE ET SEMANTIQUE : DE QUOI PARLE-T-ON A L'AFS ?

La comparaison terme à terme du contenu des résumés des deux congrès inviterait à reprendre les dictionnaires, les jeux de catégories et les collections construits précédemment¹¹ et à les projeter sur ce nouveau corpus. Mais en maintenant cette seule posture, on se heurte à la possibilité même de saisir des changements. Car si on fait le pari que des transformations ont eu lieu en deux ans - hypothèse qui justifie le travail comparatif - il reste que l'on peut manquer les formes émergentes : nouvelles entités, autres façons de décrire le social ou de rendre compte du travail sociologique. Est-ce à dire qu'il faut reprendre intégralement la construction des catégories, déniaient ainsi tout caractère cumulatif à l'analyse sociologique s'appuyant sur des textes (Chateauraynaud, 2003) ? Cette tension entre la nécessaire réévaluation de l'indexation et l'impératif de rapporter les différences à une base commune pour la comparaison est une constante de toute analyse comparative, en particulier dans sa dimension quantitative et historique.

On développe ici des stratégies intermédiaires entre ces deux extrêmes. A l'intérieur de la langue, au-delà de la diversité des formes et des inventions lexicales, on peut miser sur certaines stabilities des façons d'agencer les mots, les expressions et les arguments. Les "formules"¹², à l'intérieur du logiciel Prospéro, permettent une économie cognitive puisque les formes pertinentes lors du premier travail de codification ont été sauvegardées. Ainsi, par

¹¹ Le travail présenté ici s'effectue comme lors de nos précédents articles, (Trabal, 2005, Torny et Trabal, 2006), avec le logiciel Prospéro. Pour mieux connaître cet outil, on peut consulter le site prosperologie.org ; on peut aussi se reporter à l'ouvrage de son auteur pour saisir les fondements épistémologiques qui sous-tendent le programme (Chateauraynaud, 2003).

¹² Il s'agit d'un vocabulaire composé de quelques mots-clés servant à désigner les objets du logiciel (les entités, les catégories, les collections...) qui permet de repérer des agencements caractéristiques.

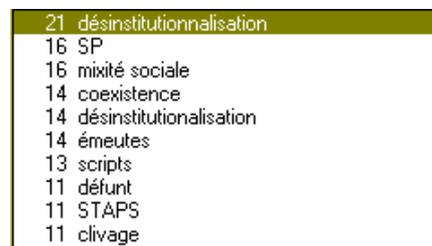
exemple, alors que nous cherchions à coder les marques de parité dans le corpus 2004, nous avons créé une série de formules comme celle-ci :

/EF=HOMME@ /BLA /EF=FEMME@.

Elle vise à identifier tous les segments composés d'un représentant de l'être-fictif HOMME@¹³, suivi d'un empan variable jusqu'à ce qu'il repère un représentant de l'être-fictif FEMMES@¹⁴. On voit alors apparaître des formes qui viennent compléter notre catégorie de parité, comme par exemple "filles-garçons", "filles et garçons" ou "les jeunes femmes et les jeunes hommes" - agencements absents du corpus 2004.

Cette démarche visant à repérer des nouveaux agencements peut être lue comme un premier travail de comparaison puisqu'il s'agit bien de chercher des formes, jugées pertinentes lors de l'analyse du précédent corpus, qui révèlent des nouvelles manières d'exprimer ces idées dans les textes récemment introduits. Elle permet également de compléter à moindre coût le travail de codification déjà effectué; c'est donc sur un corpus correctement indexé, que l'on pourra mobiliser d'autres outils de comparaison.

Mais il existe également un vocabulaire réellement nouveau, qui n'était aucunement présent dans le corpus de 2004. En sélectionnant ce type de propriété, on adopte une stratégie interprétative ouverte, qui s'oppose au choix *a priori* de thématiques que nous avons effectué dans un précédent article. Il s'agit là d'accorder une attention particulière, et rarement importante sur le plan statistique, à des éléments qui surgissent et dont il s'agit de déterminer la pertinence : simple variation lexicale, voire graphique, ou apparition de nouveaux objets ou concepts pour l'analyse sociologique ?



21	désinstitutionnalisation
16	SP
16	mixité sociale
14	coexistence
14	désinstitutionnalisation
14	émeutes
13	scripts
11	défunt
11	STAPS
11	clivage

Figure 3 Entités nouvelles dans le corpus 2006 par rapport au corpus 2004, classées par occurrences

Sous ce rapport, la notion de désinstitutionnalisation, avec ces deux orthographes utilisées (avec ou sans le doublement du "n") émerge de façon significative. En rassemblant les différentes formes graphiques sous un être-fictif, nous recensons 36 occurrences employées par 16 auteurs. D'un point de vue purement lexicométrique ou statistique, ce n'est certes pas un objet important du corpus (il ne s'agit que du 231ième score); mais son surgissement dans le corpus 2006 intrigue. En regardant les réseaux thématiques mobilisant ce concept, on repère qu'il s'agit essentiellement du RT 40 "Sociologie des institutions" (à 32 reprises), ce qui est certes attendu. Mais pourquoi aucun auteur de ce réseau n'employait-il pas ce terme en 2004 dans leurs 27 communications, alors qu'on l'évoque maintenant à propos de la prison, de l'hôpital ou de l'armée, et alors même que cinq de ces auteurs étaient déjà présents au

¹³ Un être fictif est regroupement de termes fixé par le chercheur lorsqu'il considère que ces mots désignent une même entité. Ainsi, l'être fictif HOMME@ rassemble des termes comme "homme", "hommes", "garçon", "jeunes gens"...

¹⁴ Parmi les représentants de l'être fictif FEMME@, on recense "femme", "femmes", "filles"...

congrès de Villeteuse ? Cela est d'autant plus surprenant que cette notion est pourtant bien installée dans la tradition sociologique¹⁵. Si on s'arrête au matériel des résumés, cette variation demeure ininterprétable, ce qui nous rappelle que le logiciel a moins vocation à administrer une preuve de façon définitive, que d'ouvrir de nouvelles pistes d'enquête. Nous pouvons le faire en nous appuyant sur un autre matériel : l'appel à communication du RT 40 permet d'interpréter de manière satisfaisante l'émergence du terme puisque son titre est « La désinstitutionnalisation des institutions ? ». On recueille donc là un phénomène d'alignement par déclaration d'une politique scientifique d'un RT, et on pourra mesurer à l'avenir si cette impulsion a des effets de plus long terme.

D'autres entités émergent en 2006 ne méritent pas le même intérêt. C'est le cas, notamment du SP, acronyme de "Service Public" qui n'est porté que par deux auteurs. L'effet du nombre (16 occurrences sur un nouveau thème) ne renvoie pas à la montée d'une préoccupation nouvelle, mais à l'usage répété d'un acronyme par deux sociologues (dans le corpus, plusieurs auteurs parlent du service public, en 2006 comme en 2004). La notion de script est plus surprenante, car, comme pour la désinstitutionnalisation, il ne s'agit que d'une demie-nouveauté. Comme le montre le réseau spécifique de cette entité, la notion de script est ici connectée aux deux auteurs du concept définissant la notion de "script sexuel" en 1973 (Gagnon et Simon). Il s'agit donc d'une entité dont le surgissement pourrait faire penser à l'émergence d'un nouveau concept mais qui ne concerne que quatre auteurs réutilisant une notion datant de plus de 30 ans.

Enfin, parmi les entités surgissant en 2006, nous notons également les "émeutes". Une simple lecture de la figure 3 laisserait à penser que cette apparition est indexée sur les émeutes de novembre 2005. Ce qui permettrait d'envisager une analyse sur la distance entre intérêts académiques et questions sociales au cœur de l'actualité : y aurait-il des communications en 2008 portant sur les nombreuses élections ayant lieu d'ici là et les mobilisations qui y sont liées alors que les élections françaises récentes – et singulièrement le fameux choc du 21 avril 2002 – ne font l'objet que de trois communications dans le corpus 2006 en dépit de l'existence d'un RT « Sociologie du politique » ? La lecture des énoncés contenant le terme « émeute » invite à plus de prudence car, sur les 14 occurrences de "émeutes", la moitié provient d'un sociologue travaillant sur les émeutes de 1981 aux Minguettes.

En dehors des nouveautés pures, une analyse du même type peut être développée en s'appuyant uniquement sur des augmentations et des diminutions de poids affectant les quelques 13 691 entités des deux corpus, qui présentent un poids comparable. On se concentre ainsi sur des variations statistiques d'objets suffisamment présents dans le corpus, puisque pour des mesures de type relatif, les variations des "petits" objets peuvent être fortes sans que leur variation absolue soit importante¹⁶.

¹⁵ Voir par exemple, sur des objets très différents, (Castel, 1976) ou (Dubet, 2002)

¹⁶ Le mot "méfiance" apparaît, par exemple, une fois dans le corpus 2004 et à deux reprises dans celui de 2006. Compte-tenu de la variation du nombre de pages (respectivement 857 et 917 pour le premier et le deuxième congrès), on obtient une augmentation de 86% pour cette entité.

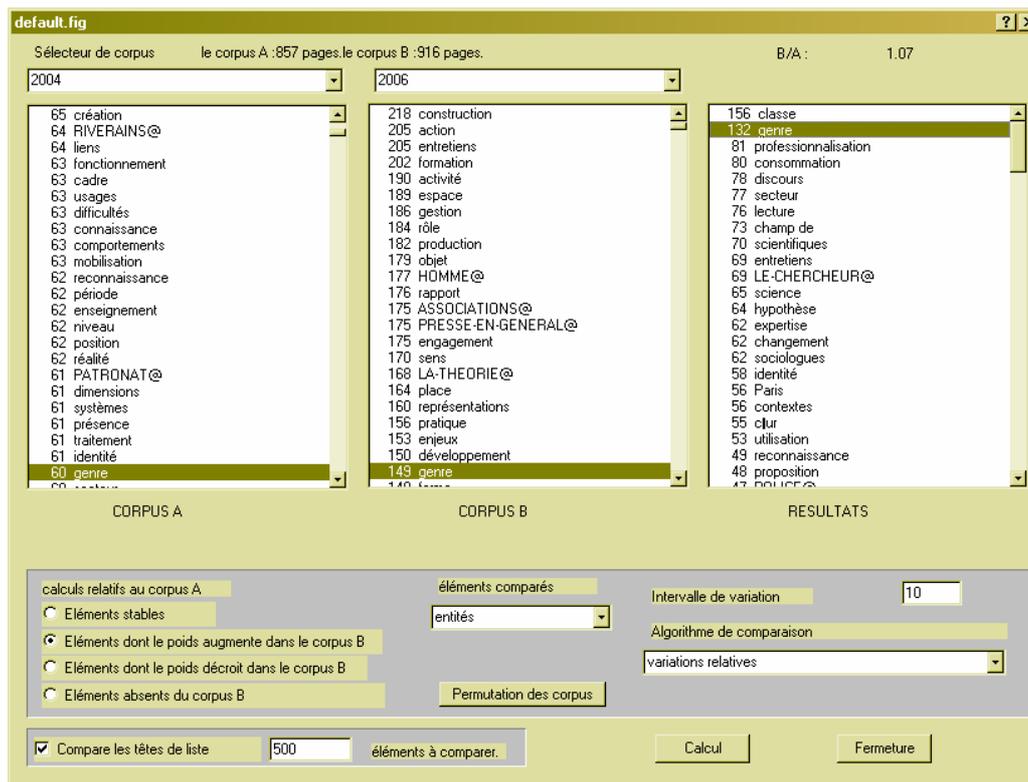


Figure 4 Eléments dont le poids relatif augmente dans le corpus 2006

Pour les 500 premières entités et être-fictifs du corpus 2004, on constate que des entités comme "classe"¹⁷, "genre", "professionnalisation" ou "consommation" figurent parmi celles dont le poids augmente le plus au corpus AFS 2006. On sait que le mot "classe" est assurément polysémique puisqu'il renvoie tout autant à un vocabulaire marxien, à un groupe scolaire faisant l'objet de travaux en sociologie de l'éducation ou encore à une discussion sur les opérations de codification statistique. Si on a déjà créé pour les traitements précédents des expressions telles que « classes populaires », « classes scolaires » ou « classes sociales », les formes élidées ou contextuelles demeurent présentes. Quant au "genre", malgré des rares allusions à des expressions comme "genre littéraire", renvoie généralement à l'étude du genre. Quand on regarde quelles sont les entités liées à "genre", on repère qu'il s'agit surtout de "classe"¹⁸. Qui plus est, cette relation entre "classe" et "genre" est nouvelle puisque tous les énoncés contenant simultanément ces deux entités proviennent du corpus 2006. Cette remarque invite à pousser l'analyse sur ces objets ce qui nous permettra d'examiner si la thématique du genre constitue une voie pour renouveler les approches marxistes que nous avons longuement étudié précédemment. Ainsi, d'un regard purement statistique qui analyse la variation quantitative, on peut bifurquer vers une logique différente, celle du suivi systématique de thèmes ou de concepts sociologiques. Il n'y a là rien d'étonnant, la variation statistique n'étant pas, par elle-même, porteuse de signification, il faut qu'elle s'appuie sur un travail interprétatif autour d'objets pertinents pour pouvoir faire sens¹⁹.

¹⁷ L'illustration mentionne le taux d'augmentation : cela signifie donc que l'entité "classe" connaît une augmentation de 156% lorsque l'on passe du corpus 2004 à celui de 2006.

¹⁸ L'autre entité fortement connectée à "genre" est "sexe", ce qui est moins surprenant.

¹⁹ Un très bon exemple en est donné par l'étude d'un très gros corpus de recettes « exotiques » saisi par des questionnements de nature anthropologique et où les indices lexicaux – ingrédients, marqueurs linguistiques - ne font sens que parce qu'ils sont reliés aux questions du prototype, de la distance sociale ou de l'authenticité supposée du plat (Regnier, 2004).

UNE COMPARAISON SOCIOLOGIQUE : LES GENDER STUDIES SONT-ELLES UNE SOURCE D'INNOVATION CONCEPTUELLE ?

Au-delà de l'analyse lexicale, comparer deux corpus peut aussi revenir à se focaliser sur la reprise des propriétés déjà dégagées pour en éprouver la robustesse, suite à la constitution de nouveaux corpus. Dans cette perspective, c'est moins les propriétés du corpus lui-même qu'il s'agit de tester que la validation, sur de nouvelles enquêtes, de la portée et de la généralité des jugements sociologiques déjà établis sur un matériel différent ou plus restreint.

Ainsi, sur le corpus de 2004, nous avons montré qu'une part importante des communications (31%) utilisait des modalisations verbales du type « je vais tenter de montrer », « vise à définir », ou encore « on propose d'analyser » (Torny et Trabal, 2006). Nous avons analysé ces modalisations, qui s'opposent à un discours scientifique objectivant, comme des marques spécifiques du résumé de proposition de communication en tant que genre littéraire, parce qu'il constitue un projet, un programme ou une promesse. Comme l'ensemble de ces modalisations verbales est indépendant des objets sur lesquels il porte et qu'il est construit à partir de « formules », on peut projeter sans modification la même grille d'analyse sur le corpus 2006 où on obtient un résultat similaire (28%), ce qui aurait tendance à indiquer une constante largement indépendante des groupes de sociologues concernés. De plus, la connaissance du statut de l'auteur permet de répondre à une interrogation que nous posons : s'agit-il de formes particulièrement employées par des auteurs jeunes dans la carrière ? Il n'en est rien, puisque ces formes sont autant employées par les permanents, les docteurs ou les doctorants²⁰.

D'autres résultats peuvent être aussi confortés ou infléchis : dans quelle mesure la montée en puissance des études liées au genre et leur effet se manifeste-t-elle au congrès de Bordeaux ? La rhétorique marxienne, dont on pressentait le renouveau dans le corpus de l'AFS 2004, est-elle à nouveau en développement ? Comme nous l'avons montré dans notre analyse de l'impact des gender studies (Torny et Trabal, 2006), il convient de dépasser l'étude du seul mot "genre", puisqu'il y existe bien des manières d'évoquer cet objet (marqueurs de parité, adjectifs du type féminin, masculin). L'objet le plus spécifique était une catégorie que nous avons construit pour ce corpus, "Rapport de sexe", contenant pas moins de 133 représentants, et que nous avons ajusté avec la création de 29 nouvelles expressions (« stéréotypes de genre », « mixité sexuée »,...) construites à partir de jeux de formules et issues du corpus 2006.²¹ Cet ensemble a été mobilisé à 207 reprises dans 90 résumés en 2004, et ce sont surtout les femmes (69%) et les auteurs du RT 24 (43% des usages) qui ont utilisé les entités de cette catégorie. En 2006, la catégorie est présente à 209 reprises dans 106 résumés, le taux d'auteurs utilisant cette catégorie augmente fortement (près de 84%) et celui des sociologues du RT 24 légèrement (46%). L'évolution n'est donc pas quantitativement marquante, même si elle est marquée par une féminisation accrue.

²⁰ Nous sommes en effet très proches du cas d'indépendance ($Khi^2 = .78$; $ddl=2$)

29 Mais analyser l'évolution de la thématique du genre revient aussi à étudier les changements dans les liens entre cet objet et d'autres problématiques. Le réseau spécifique de "genre" dans le corpus 2004 pointait des discussions sur le concept, les politiques sociales et la professionnalisation. Les résumés du congrès 2006 font apparaître que le genre est spécifiquement lié à la notion de classe, de la domination et de la race.

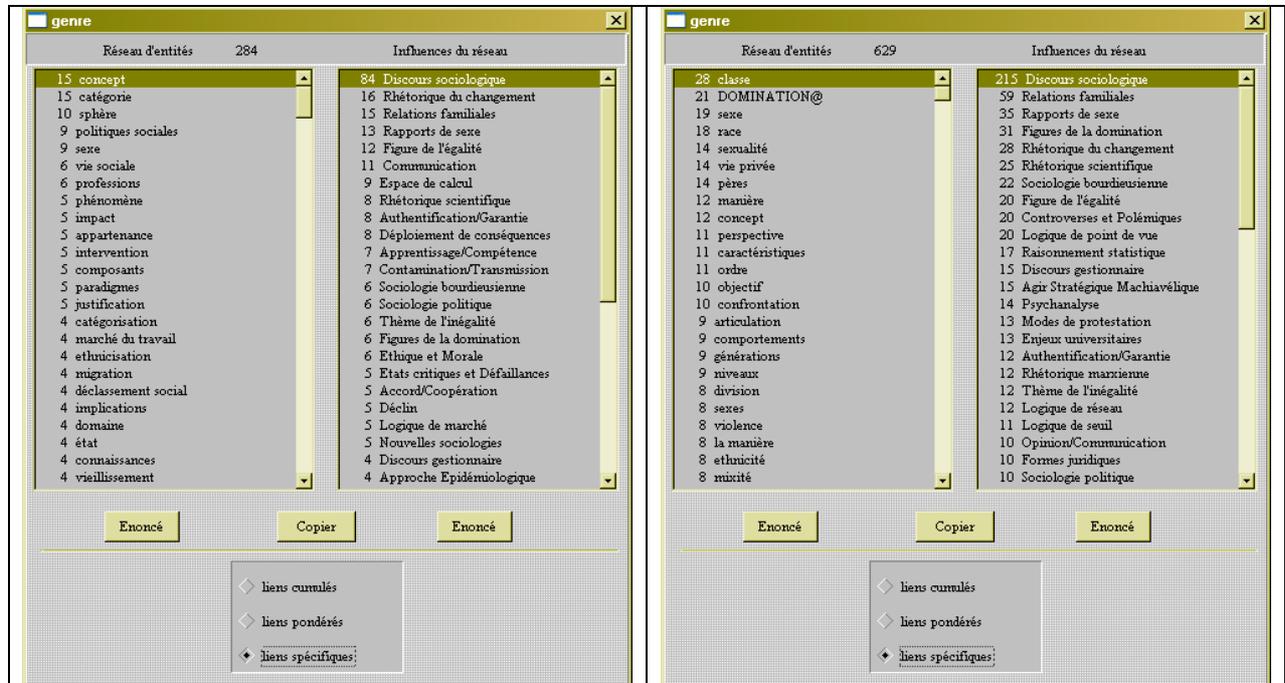


Figure 5 Réseaux spécifiques de "genre" dans les corpus 2004 (gauche) et 2006 (droite)

La notion de domination est désormais assez centrale et, comme dans le cas de la désinstitutionnalisation, on recueille là une modification de la politique scientifique du RT lui-même comme le confirme la lecture de l'appel à proposition publié par ce réseau thématique, qui explicitement invite à "penser l'articulation des rapports de genre à d'autres rapports de domination":

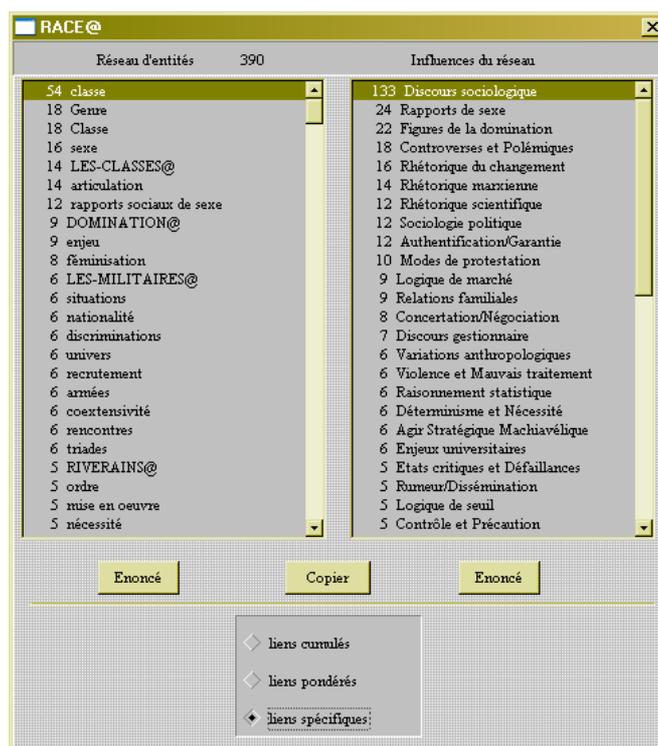
Sexe et classe comme rapport social doivent être croisés avec celui de race. Car si l'on admet que le genre construit le sexe, le capitalisme la classe et le racisme la race, on se doit, dans un tel contexte de reconfiguration mondiale, de saisir ces trois systèmes sociaux de domination dans leur interaction afin de comprendre comment, à partir d'une naturalisation de l'altérité, se (re)construisent les rapports de hiérarchie et de domination, mais aussi comment naissent et se développent résistances et révoltes. (Extrait de l'appel du RT 24).

Ce changement d'approche – on vient de le voir, parfaitement, assumé par les animateurs (trices) de ce réseau – invite à travailler sur les questions de la classe et de la race. On saisit, du coup, la nouvelle connexion entre genre et classe décrite plus haut. Les sociologues utilisant les termes caractéristiques de la rhétorique marxienne, évoluent principalement dans les RT 5 ("Classes, inégalités, fragmentations") et dans le RT 24 qui contient également le mot "classe" dans son titre. Mais au-delà des transformations dans le découpage et les revendications des notions sociologiques, la présence de "race" étonne.

En effet, travailler sur la race ou sur l'ethnie sur une base statistique fait l'objet de polémiques récurrentes en France, contrairement par exemple à la Grande-Bretagne, particulièrement pour mesurer des phénomènes tels que les discriminations (Stavo-Debaugue, 2003). Du coup, on peut s'interroger sur les manques de notre analyse précédente : l'objet "race", figurait-il dans

le corpus 2004 ? Nous ne trouvons en fait que deux énoncés dans les résumés du premier congrès, dont l'un est d'ailleurs discutabile puisqu'il s'agit du titre d'une revue du début du XXème siècle (La race et les mœurs), que l'on aurait pu coder en tant que tel. Dans les actes du deuxième congrès, nous relevons 28 occurrences du mot "race", se répartissant dans 20 textes.²¹ Ces 28 apparitions émanent à 27 reprises de sociologues communiquant dans le RT 24 ou dans une session mixte dans lequel ce même RT participait, ce qui marque clairement le point d'origine de la nouveauté et son absence de diffusion pour le moment.

33L'émergence de ce thème, qui constitue un « signal faible » et en conséquence peu visible par les lexicomètres, apparaît tout autant comme une surprise pour le chercheur qu'une nouvelle piste de travail. En effet, la nouveauté de type lexical, on l'a vu dans la deuxième partie, peut n'avoir qu'un faible porté sémantique et a fortiori sociologique. Aussi, pour en tester l'importance, il convient avant tout de reconstruire cette notion en créant un être-fictif rassemblant les différentes formes graphiques de "race" et des termes proches : nous avons identifié 31 représentants parmi lesquels "origine ethnique", "minorité ethnique", "domination ethnicisée"... Le score cumulé atteint 78 dans le corpus 2006, alors que la projection du même regroupement ne donne que neuf occurrences sur les résumés de 2004, confirmant ainsi le caractère « nouveau » de cette thématique dans le cadre de l'AFS. Si l'on s'emploie à discriminer les usages de cet être-fictif, on note que ses représentants sont souvent utilisés par les femmes (77%) et par les sociologues du RT 24 (62%). L'analyse du réseau spécifique de RACE@ montre l'existence d'énoncés liant les notions de genre, de classe et de race.



²¹ On pourrait s'étonner du fait que cette augmentation n'apparaisse pas dans l'illustration 4. En fait, comme nous le signalions, à la note 17, nous pouvons décider de n'afficher que les entités figurant parmi les 500 premières du corpus (cette valeur est d'ailleurs paramétrable). L'intérêt est de limiter le bruit que constituerait une augmentation importante d'une petite entité passant de une ou deux occurrences entre les deux sous-corpus. Les limites de cette contrainte sont de manquer l'émergence d'entités réalisant un très faible score dans le premier sous-corpus. Le cas de l'objet "race" constitue une illustration de l'intérêt de cette option permettant de réaliser les deux expériences.

En utilisant l'algorithme de "configuration discursive" qui permet de régler le degré de contraintes sur le niveau de présence des catégories "Rhétorique marxienne" et "Rapport de sexe", d'une part, et des êtres-fictifs RACE@ et LES-DOMINES@, d'autre part, on peut donc identifier des énoncés ou des textes mobilisant simultanément l'ensemble de ces thématiques. Ainsi, peut-on pointer quelques textes prototypiques, dont un parvient en quelques énoncés à convoquer tous ces objets :

Ces femmes apparaissent comme le signe d'une certaine modernité sociale face aux ouvriers du bâtiment, jugés, par certains responsables d'entreprise et de représentants syndicaux patronaux, comme des êtres dangereux, bruts, voire primitifs, jugement par ailleurs non exempt de références ethnicisées discriminantes. La féminisation du secteur du bâtiment introduit donc nombre d'enjeux sociaux. Ces enjeux se construisent autour de trois rapports sociaux : de genre par l'introduction de femmes dans un univers d'hommes ; de classes sociales par la féminisation vue comme un moyen de management ; de race, enfin, par la confrontation, sur le sens de la féminisation, des références des ouvriers et des managers issus d'origines ethniques différentes.

On poursuit alors l'enquête en utilisant les différents éléments bâtis au cours de l'analyse, les variables externes telles que les nouveaux auteurs ou le statut des communicants. Nous avons construit un sous-corpus rassemblant l'ensemble des résumés convoquant l'un des quatre thèmes décrits ci-dessus. On repère que 70% des auteurs des 184 textes ainsi rassemblés n'ont pas communiqué au congrès de 2004, taux comparable à celui des nouveaux auteurs sur l'ensemble du corpus. Mais en discriminant les auteurs évoquant les questions de race, on obtient des chiffres différents : les auteurs déjà présents en 2004 sont deux fois plus représentés dans l'utilisation de ce thème. Symétriquement, on note que les doctorants sont deux fois moins nombreux dans l'emploi de cet être-fictif. On peut donc considérer que la race comme un nouvel objet d'étude pour des sociologues français, qui malgré l'absence d'éléments statistiques sur l'hexagone, ont importé cette préoccupation présente ailleurs dans le monde, et particulièrement dans la sociologie anglo-saxonne. Il s'agira ensuite pour nous de suivre l'avenir de cette importation, de voir si elle s'étend à d'autres RT, à d'autres auteurs, à de nouveaux objets ou si elle reste confinée, voire si elle disparaît, constituant dans ce dernier cas une nouveauté sans lendemain.

*

**

*

A travers les trois formes de comparaison que nous avons développées dans cet article, on saisit la tension constante entre l'application d'une grille préfixée – même si elle est très sophistiquée en raison d'un long travail préliminaire de codage au plus près des textes – qui favorise une économie de codage et la volonté de suivre au plus près les variations les plus fines en laissant ouverte une partie des codages, au risque que les résultats apparaissent moins robustes ou plus artefactuels. Au cœur de cette tension réside la double conception d'une statistique tout autant réaliste que construite (Desrosières, 1993). En nous appuyant sur plusieurs algorithmes – en particulier, ceux reposant sur des formules – nous avons montré qu'on pouvait réduire cette tension en enrichissant les codages existants, réduisant ainsi une asymétrie créée de facto par un appui fondé sur une exploration systématique des premiers corpus, au détriment des derniers entrés.

Cette asymétrie historique, avec sa préférence pour le lien avec le passé et la conservation des grilles comparatives, existe aussi bien du côté des appareils statistiques d'Etat que des personnes ordinaires qui développent des activités de codage (Boltanski et Thévenot, 1983). A contrario, nous avons défendu une position symétrique qui vise à être aussi attentif aux variations

quantitatives et qualitatives à l'intérieur de codages préétablis, et donc à des interrogations sociologiques déjà là, qu'à l'émergence de nouveautés lexicales ou sémantiques qui ne peuvent être réduites à l'existant, et donc susceptibles de déclencher des questions inédites. Cette position implique aussi un détachement toujours possible entre le constat de variations lexicales, relationnelles ou statistiques et la série d'inférences qui peut en être tirée : réponse à une commande explicite, modification des populations d'auteurs, adoption de nouveaux thèmes de recherche ou simple effet d'accumulation aléatoire.

On a insisté à plusieurs reprises sur la nécessité de construire des séries historiques longues – qui le seront au fur et à mesure des congrès de l'AISLF et de l'AFS. Le sens des évolutions décrites dans cet article pourra évidemment en être modifié, la séparation entre épiphénomènes et tendances lourdes étant toujours faite à l'échelle du temps considéré. De même, en l'absence de statistiques précises et publiques sur la profession, en se restreignant à ces congrès, on ne pourra mesurer la « masse noire » des collègues qui ne s'y rendent jamais – par manque d'intérêt, de moyens ou parce qu'ils privilégient d'autres arènes de communication scientifique. L'important est donc bien, à chaque fois, de connaître les limites des codages engagés, du matériel recueilli et de les ajuster aux questions posées. Avec la mise en série progressive du matériel et la mise à disposition de nouvelles variables (taux d'acceptation, âge des auteurs...), on sera à même de maintenir ce cahier des charges en répondant à de nouvelles questions sur les trajectoires professionnelles des auteurs, leurs choix de thèmes et de réseaux thématiques, ou l'évolution des politiques à l'intérieur de l'AFS. Le foisonnement de ces axes comparatifs n'est pas une perte d'objectivité, mais la possibilité d'une relecture pertinente d'un matériel dont on pensait avoir tout dit.

Bibliographie

- Boltanski L., Thévenot L. (1983), "Finding One's Way in Social Space: A Study based on Games", *Social Science Information*, 22 (4-5), pp. 631-680.
- Boudesseul G. (2006). "Brève sociographie du Second Congrès de l'AFS – Bordeaux Septembre 2006". *Bulletin de Méthodologie Sociologique*. n°92. octobre, pp. 45-54.
- Castel R. (1976). *L'ordre psychiatrique : l'âge d'or de l'aliénisme*. Paris, Minuit.
- Chateauraynaud F. (2003). *Prospéro. Une technologie littéraire pour les sciences humaines*, Paris, CNRS éditions.
- Demazière D., Brousseau C., Trabal P., Van Meter K. (eds.). (2006). *Analyses textuelles en sociologie*. Rennes, Presses Universitaires de Rennes.
- Desrosières A. (1993), *La politique des grands nombres : histoire de la raison statistique*, Paris, La Découverte.
- Dubet, F. (2002). *Le déclin de l'institution*. Paris, Seuil.
- Gagnon J., Simon W. (1973), "Sexual Scripts : Permanence and Change". *Society*, 22, pp. 55-60.
- Marchal E., Torny D. (2003). "Des petites aux grandes annonces. Evolution du marché des offres d'emploi (1960-2000)". *Travail et emploi*. n°95. juillet. pp.78-93.
- Pontille D. (2004), *La signature scientifique. Une sociologie pragmatique de l'attribution*, Paris, CNRS.
- Regnier F. (2004), *L'exotisme culinaire. Essai sur les saveurs de l'autre*, Paris, PUF.
- Stavo-Debaugé, J. (2003), "Prendre position contre les catégories ethniques. Le sens commun constructiviste, une manière de se figurer un danger", in P. Laborier et D. Trom (eds.), *Historicité de l'action publique*, Paris, PUF.
- Torny D., Trabal P. (2006). "Le résumé de communication comme objet sociologique - Une analyse thématique, ontologique et littéraire à l'aide du logiciel Prospéro". In Demazière D. et al. (eds.) *Analyses textuelles en sociologie*. Rennes, Presses Universitaires de Rennes, pp. 23-80.
- Trabal, P. (2005). "Le logiciel Prospéro à l'épreuve d'un corpus de résumés sociologiques". *Bulletin de Méthodologie Sociologique*. n°85. pp. 10-43.
- Van Meter K., Saint Léger (de) M. (2007). "Preliminary co-world analysis of the 2006 congress of the 'Association Française de Sociologie'", *Bulletin de Méthodologie Sociologique*. n°93. janvier, pp. 55-69.