

Mots. Les langages du politique

69 (2002)
Révolutions

Jean-Paul Honoré

« N'oublie pas le cœur du débutant ». Remarques empiriques sur la formation des corpus en AD

Avertissement

Le contenu de ce site relève de la législation française sur la propriété intellectuelle et est la propriété exclusive de l'éditeur.

Les œuvres figurant sur ce site peuvent être consultées et reproduites sur un support papier ou numérique sous réserve qu'elles soient strictement réservées à un usage soit personnel, soit scientifique ou pédagogique excluant toute exploitation commerciale. La reproduction devra obligatoirement mentionner l'éditeur, le nom de la revue, l'auteur et la référence du document.

Toute autre reproduction est interdite sauf accord préalable de l'éditeur, en dehors des cas prévus par la législation en vigueur en France.

revues.org

Revues.org est un portail de revues en sciences humaines et sociales développé par le Cléo, Centre pour l'édition électronique ouverte (CNRS, EHESS, UP, UAPV).

Référence électronique

Jean-Paul Honoré, « N'oublie pas le cœur du débutant ». Remarques empiriques sur la formation des corpus en AD », *Mots. Les langages du politique* [En ligne], 69 | 2002, mis en ligne le 14 mai 2008. URL : <http://mots.revues.org/10633>

DOI : en cours d'attribution

Éditeur : ENS Éditions
<http://mots.revues.org>
<http://www.revues.org>

Document accessible en ligne sur : <http://mots.revues.org/10633>

Ce document est le fac-similé de l'édition papier.

© ENS Éditions

Jean-Paul HONORÉ^o

« N’oublie pas le cœur du débutant » Remarques empiriques sur la formation des corpus en AD

*Shoshin wasuru bekarazu*¹

On ne s’étonnera pas de trouver surtout des questions, et certaines délibérément naïves, dans le texte ci-dessous. Les propos qui suivent, en effet, n’étaient pas destinés à publication, mais devaient contribuer à animer une table ronde sur le thème de la constitution et du traitement des corpus. Ils reflètent une pratique qui n’a rien d’exemplaire (peut-être rien de rare non plus), et témoignent de quelques-unes des interrogations qui surgissent lorsque le travail de recherche s’effectue sur la base de corpus non numérisés (ou dont la numérisation n’est que partielle), impropres à une enquête statistique approfondie, et constitués ou interrogés de façon empirique.

Je n’ai sans doute jamais assez théorisé pour moi-même les conditions nécessaires à l’établissement d’un « bon » corpus, et cela contribue à expliquer la grande variété que présentent ces objets en ce qui me concerne.

Différences de durée : certains, privilégiant une variable chronologique, embrassent un intervalle relativement long (cinquante ans pour celui qui concerne les changements de noms). D’autres saisissent une durée courte (les corpus relatifs aux tremblements de terre de Kobé et de San Francisco ne s’étendent qu’aux trois semaines consécutives à

^o CEDITEC / Université Paris Val de Marne, 61 av. du Général de Gaulle, 94010 Créteil Cedex.

1. « N’oublie pas le cœur du débutant » (proverbe japonais).

chacun des évènements)² : les variables sont ici constituées par la dualité des évènements, et la diversité des locuteurs.

Différences d'homogénéité : le corpus consacré à l'étude des changements de noms est extrêmement homogène (il n'est fait que des demandes de changements de nom publiées dans la rubrique « Annonces » du *Journal Officiel*). En revanche, le corpus iconographique consacré au Japon est plus hétérogène : il rassemble des images publicitaires, des caricatures, des photos, des graphiques, le tout articulé ou non à du texte. Tous ces objets ont des fonctionnements différents. Fallait-il les réunir ? Mon hypothèse pour l'instant est que si l'on s'intéresse à la circulation des stéréotypes interculturels, la variété de leurs modes d'inscription peut conférer au corpus un caractère probant.

Une troisième différence importante est constituée par le choix des entrées : les corpus fabriqués à partir de textes sont des sortes d'index, et se présentent sous la forme de tableaux incorporant à gauche une lexie (désignant, mot-thème, mot-phore, emprunt, néologisme, formule plus ou moins figée, etc.), à droite son contexte. Le corpus iconographique place chaque élément sous une étiquette regroupant quatre grandes rubriques : genre/émetteur/mot-clé/date. Dans les deux cas, c'est une intuition de fréquence, ou de représentativité, qui a guidé la sélection.

Enfin, comme tout un chacun, à côté de corpus pondérés quant à leurs masses, leurs sources et leurs durées, je fabrique ce que j'appellerai ici des « collections ». J'entends par « collection » un ensemble constitué de différents corpus orientés vers le même thème, mais comportant des éléments dont la compatibilité est encore problématique (en terme de chronologie, de méthodologie, etc.) ; ma collection « représentations du Japon » est faite de trois corpus différents (images, textes, corpus Kobé/San Francisco). Une collection peut aussi être constituée du corpus et de son environnement théorique (ouvrages, articles et tous documents afférents). On peut trouver des principes (chronologiques, génériques, statistiques, heuristiques) qui permettent de délimiter de façon cohérente un corpus. En revanche, je n'en vois pas qui délimitent une collection, ce qui rend la constitution et le maniement de celle-ci problématiques.

Partant de là, il m'est difficile de concevoir des « corpus-types » en AD, ou de prétendre fournir des principes valables pour un travail sur tout type de corpus. La méthode ainsi dégagée ne risquerait-elle pas, d'ailleurs, de se confondre avec les principes généraux de rigueur inhérents à toute recherche ? Faute de pouvoir prendre position sur ces ques-

2. Ma recherche a pour thème les représentations de l'altérité dans le discours de la presse contemporaine en France, d'où des articles sur le discours du Front National, les stéréotypes interculturels (discours et iconographie), le vocabulaire antisémite, les noms propres, etc.

tions générales³, je m'en tiendrai, comme annoncé, à une approche triviale et socialement répréhensible : « parler de mes problèmes », et spécialement de quelques-unes des affaires par lesquelles je passe quand j'entreprends un article pour *Mots*.

Problème de cohérence externe

On ne construit pas un corpus en fonction des conclusions que l'on souhaite en tirer. Mais on construit tout de même un corpus en fonction d'un projet générique (une thèse, un article...) et thématique. Cela oblige à vérifier qu'à toute étape de sa constitution et de son maniement, le corpus élaboré reste conforme à ce pour quoi on l'a produit.

Pour le numéro de *Mots* consacré aux noms propres en politique, mon premier projet était de travailler sur deux dictionnaires, ouvrages récents et que l'on se procure aisément en librairie ou en bibliothèque : d'une part, le *Dictionnaire des changements de noms*, dont l'auteur signe « L'archiviste Jérôme »⁴ ; d'autre part, l'*Encyclopédie des changements de noms*, parue en deux volumes (1995 et 1998), et signée « Emmanuel Ratier ». Pourquoi ces ouvrages-là ? Parce que si la littérature relative aux changements de noms est abondante, les nomenclatures systématiques sont rares : les deux dictionnaires en question occupent le terrain. On pouvait se demander s'ils ne constitueraient pas des inventaires objectifs et solides, propices à des recoupements, et dont les textes de présentation et d'accompagnement témoigneraient d'analyses et de jugements contemporains relatifs au changement de patronyme en général.

C'était ignorer que *l'archiviste Jérôme* et *Emmanuel Ratier* sont deux pseudonymes du même auteur : Henri (ou Henry) Coston. Celui-ci, pendant la guerre, dirigeait le Centre d'action et de documentation antimaçonniques. Puis il fit parler de lui dans des affaires de révisionnisme, et mit en circulation un certain nombre de textes dits antisémites. Le commentaire sur les changements de nom incorporé aux dictionnaires

3. Voir en revanche l'entretien de M. Tournier avec N. Arnold et F. Dougnac, « Zoom arrière. Passions et mesures » dans *Des mots en liberté. Mélanges Maurice Tournier*, 1998, ENS Éditions, tome 2.

4. *Dictionnaire des changements de noms de 1803 à 1956*, L'Archiviste Jérôme, chez l'auteur, (H. Coston), 1964. *Dictionnaire des changements de noms, 1957-1962*, L'archiviste Jérôme, chez l'auteur (H. Coston), 1964. *Encyclopédie des changements de noms*, E. Ratier, 2 vol., éd. Faits et documents, 1998.

en question (et à un troisième : le *Dictionnaire des pseudonymes*⁵) est donc orienté de façon bien particulière. Il n'est pas inintéressant sur le plan de l'AD : c'est un discours très manipulateur, caractérisé par le retour de la profération antisémite sous prétexte d'onomastique, par la tension entre habillage scientifique et stéréotypes judéophobes, par des phénomènes d'axiologisation et d'exemplification tendancieuses, de présupposition, de sous-entendus, etc.⁶

Mais d'un autre côté, sur le plan de la cohérence externe, quel était l'intérêt dans la perspective de ce numéro de *Mots* ? Pour contribuer à un travail collectif traitant des « Noms propres en politique », suffit-il de disposer, outre la nomenclature elle-même, d'une quinzaine de pages (préfaces, prières d'insérer, etc.) rédigées par un seul et même auteur – un marginal sulfureux – pour des ouvrages qui circulent et sont réédités, mais dont on apprécie mal le rayonnement, faute de savoir avec précision par qui et comment ils sont lus (ce qui serait le sujet d'une autre recherche) ? Écrire sur cette base ne revenait-il pas à créer un effet d'optique, en postulant la densité de la présence de ce locuteur-là dans les stratégies qui s'élaborent autour du nom propre au sein de notre société ?

Je m'étais donc trompé de corpus ; celui que j'avais préparé n'avait guère de cohérence avec le projet global. En revanche, il aurait pu fournir une chronique dans un numéro de *Mots* portant sur les inscriptions contemporaines du discours antisémite.

5. Henry Coston, chez l'auteur, 1961. Abondamment réédité, puis devenu *Encyclopédie des pseudonymes* en 1993, sous la signature de Ratier (avec une préface de Coston...).

6. Ce n'est pas, croyons-nous, faire un mauvais procès à l'auteur que de repérer dans l'« Avis au lecteur » du *Dictionnaire des changements de noms de 1803 à 1956* (édition de 1995) la remarque suivante : « Depuis le début du siècle, les modifications concernent principalement les noms portés par des israélites. *Qu'on le regrette ou qu'on s'en félicite*, cette tendance semble devoir s'accroître » (nous soulignons la notation modale qui, en imposant une alternative, constitue subrepticement l'évènement en problème). On peut méditer aussi le choix de certains des exemples par lesquels l'auteur illustre sa méthode : « Nous avons utilisé pour ce travail les publications officielles. Chaque nom modifié est suivi du nom précédent, placé entre parenthèses avec les prénoms, et de la référence. Par exemple, lorsque nous imprimons :

Franck (*Rosenstock-Franck*, Louis-Joseph Emile)

et ses enfants mineurs. D. 23-9-1955,

c'est que nous avons relevé dans le *Journal Officiel* du 1er octobre 1955, page 9643, le décret suivant : [...] ROSENSTOCK-FRANCK (Louis-Joseph-Emile, né le 28 avril 1906 à Jaffa (Palestine), demeurant à Neuilly-sur-Seine [...] et ses enfants mineurs: a) Guillaume, né [...] à Neuilly-sur-Seine [...]; b) Anne, née [...] à Neuilly-sur-Seine [...]; c) Doris, née [...] à Washington [...] sont autorisés à substituer à leur nom patronymique celui de "FRANCK", afin de s'appeler légalement à l'avenir [...] ... ».

Aucun exemple n'est choisi au hasard. Celui-ci, qui associe la constellation topique *Jaffa/Neuilly/Washington* à une entreprise de dévoilement du nom originel, nous paraît mettre intentionnellement en exergue la vocation réelle de l'ouvrage.

Problème de cohérence interne

J'entends par là toutes les difficultés qui touchent à la définition de l'objet sur lequel on s'interroge, et qui menacent l'homogénéité (toujours relative) du corpus. Par exemple, si l'on s'interroge sur les changements de nom, doit-on retenir le cas des francisations⁷ ? La question est importante, car en raison du nombre de celles-ci, la masse et les équilibres internes de l'échantillon sont modifiés, et certains mécanismes formels s'en trouvent privilégiés (la traduction, par exemple). Par ailleurs, on aura accès à une autre sociologie, à une autre symbolique – sans parler de phénomènes annexes, comme le rôle du prénom dans le processus. Il faut donc trouver un principe de cohérence interne (en l'espèce la modification de l'état-civil), lequel s'appuie nécessairement à une hypothèse (l'idée que les transformations formelles du nom nous situent efficacement à la croisée du linguistique et du social). Construire cette cohérence, tout en se gardant des hypothèses clandestines, n'est jamais simple ; mais tant que ces orientations ne sont pas formulées, on est dans la collection plutôt que dans le corpus.

Problème de représentativité

Dans l'approche empirique et intuitive qui est la mienne, c'est un problème redoutable. Qu'est-ce qu'un corpus représentatif ? Comment doit-il se présenter pour qu'on puisse supposer qu'il l'est ?

À propos du Japon, certains journaux ont des locuteurs experts (Ph. Pons pour *Le Monde*, B. Birolli pour *Le Nouvel observateur*...) : si l'on s'intéresse à la circulation des stéréotypes, comment faut-il traiter ces auteurs ? Comment faire le départ entre ce qui relève de « l'archive » que l'on s'efforce de saisir, et ce qui est plus idiosyncrasique ? Si l'on travaille sur tel ou tel de ces supports, doit-on ou non pondérer la place que prennent ces locuteurs-experts, et rétablir une sorte d'équilibre quantitatif qui n'existe pas dans les faits ? De la même façon, selon quel critère opérer le choix des différents supports (journaux et magazines) qui serviront de base au corpus « tremblement de terre de Kobé » ? Une solution un peu automatique consiste à s'appuyer sur les

7. Le changement de nom, au sens juridique de l'expression, ne peut concerner que des citoyens français. Ce que l'on appelle « francisation » est une procédure ouverte aux étrangers, en accessoire à leur naturalisation.

orientations politiques de ces périodiques : on essaie de les équilibrer plus ou moins. Mais est-ce bien le meilleur choix quand on tente de réfléchir aux représentations interculturelles ? Certes, au fur et à mesure de la constitution du corpus, d'autres choix se dessinent. Mais on ne commence jamais dans l'encyclopédie, ni dans la neutralité, et il existe une difficulté à ne pas se laisser définitivement conditionner, fût-ce inconsciemment, par les choix initiaux.

Le problème de la représentativité se pose sous beaucoup d'autres formes, notamment la définition de la période couverte par le corpus, et la constitution éventuelle, quand cette période est longue, d'un échantillon. En ce qui concerne le corpus « Noms propres », qui couvre la période 1949-1999, le nombre des demandes de changements de nom parues au *Journal officiel* est assez difficile à évaluer. Les décrets d'autorisation semblent mieux comptabilisés ; c'est sur eux, en tout cas, que s'appuient la plupart des études que l'on peut consulter. Les demandes, elles, relativement importantes en nombre après la guerre, sont allées en diminuant, avant que la courbe ne se redresse dans les années 1960. À la fin de la période examinée, l'ordre de grandeur était de 800 demandes par an. Il est hors de question de passer en revue, avec des moyens humains, l'ensemble des cas sur un demi-siècle, d'autant plus que pour les années 1990, par exemple, 800 demandes par an ne font pas 800 noms, mais au minimum 1600 patronymes qu'il faut relever, stocker et comparer. En fait, le total est encore bien supérieur, car si le nom de famille qu'on abandonne n'a qu'une seule forme, celui qu'on souhaite adopter en a souvent plusieurs : l'usage pour le requérant est de proposer un paradigme qui va de un à trois ou quatre noms, et l'administration fait son choix.

Comment, dans ces conditions, construire un échantillon qui soit représentatif du flux des demandes sur ces cinquante années ? Le parti retenu a consisté à faire des prélèvements de cinq ans en cinq ans, portant non sur une année entière, mais sur deux mois, toujours les mêmes : janvier et février. Pourquoi cinq ans ? Parce qu'il a semblé (est-ce bien justifié ?) que c'était le meilleur compromis entre le souci de produire un échantillon maniable, et celui de construire un maillage suffisamment fin pour adhérer, dans les grandes lignes, à l'histoire et à la sociologie de la période. Quant au choix des mois de janvier et février, il est purement arbitraire.

Il faut encore vérifier que l'échantillon ainsi constitué n'a rien d'aberrant sur le plan statistique, et qu'il est en conformité avec les informations chiffrées collectées par ailleurs. Je ne dispose, je l'avoue, d'aucune culture de la précaution statistique – instrument important mais dont

l'acquisition est toujours remise à plus tard – et n'ai guère pour outil que la règle de trois. Je me suis donc contenté de vérifier que le flux des demandes dont l'échantillon était extrait ne s'est pas interrompu dans le *JO* au cours de l'intervalle considéré ; que leur nombre total, rapporté à la masse de l'échantillon, présente une corrélation avec la durée totale de la période rapportée au nombre de mois retenus ; et, enfin, que les deux mois systématiquement et arbitrairement choisis (janvier et février) ne présentent rien d'exceptionnel par rapport aux autres⁸. Rien ne dit, cependant, que ces précautions soient suffisantes.

Le problème de la représentativité concerne aussi dans un tel cas le calibre de l'échantillon : quand on s'interroge sur les formes et les motivations du changement de nom, est-ce qu'un corpus qui représente moins d'un vingtième de l'ensemble peut être qualifié de fiable ? Je n'ai d'autre moyen de répondre à cette question que de mettre en avant l'homogénéité de l'objet, et suppose que dans ces conditions l'ultime ressource est d'expliquer dans une note comment le corpus a été fait.

Problème d'interprétabilité

À supposer qu'un corpus soit constitué de façon cohérente, on peut toujours en tirer un résultat ; mais peut-être peut-on dire, sans provocation exagérée, que les résultats les plus intéressants ne sont pas forcément à la portée de celui-là même qui a constitué le corpus.

En renonçant à un travail d'analyse du discours sur les dictionnaires mentionnés plus haut, j'avais gardé le projet d'une étude de la relation entre changement de nom et appartenance communautaire. Ayant formé le corpus dans les conditions que j'ai décrites, j'ai commencé le travail d'interprétation. Grosso modo, le corpus paraît montrer ou confirmer :

- l'évolution du nombre des demandes depuis la guerre ;
- la forte densité, puis le quasi-tarissement des sollicitations émanant de familles probablement juives ;
- l'impact de la reconstruction et celui de la guerre d'Algérie ;
- l'impact de la diversification démographique (avec en fin de période des demandes concernant des noms africains, asiatiques...) ;
- un mouvement de réappropriation identitaire assez net dans les années 1980 ;

8. En fait, un creux statistique peut être constaté en 1994. Il s'explique peut-être par le fait que les requérants sont dans l'attente du décret d'application consécutif à la loi du 8 janvier 1993, qui facilite la procédure.

- plus profondément, peut-être, une évolution de la représentation qu'on peut se faire, au sein de la population, de ce qu'est un « nom français ».

Mais au moment de la rédaction de l'article, beaucoup d'inconnues se sont révélées :

- Le caractère communautaire des noms n'est pas évident : Bensaïd ? Benchétrit ? Abensour ? Lehmann ? Tcheravatzky ?

- Les réarménisations ou les rejudéisations de patronymes sont intéressantes à souligner, mais jusqu'à quel point sont-elles représentatives d'un mouvement social profond dans les communautés en question ?

- La diversification de la société française met en circulation des patronymes dont l'origine linguistique est difficile à déterminer : *Fropo*, *Bob*, *Samb*, *Dunn*... En outre, les aires linguistiques et les aires ethniques ou nationales se chevauchent : dès lors comment traiter, faute d'une enquête de terrain, les notions d'« origine » ou de « communauté » ?

- Les demandes de changement de nom émanant des familles d'origine asiatique arrivent très tard. Que peut-on en conclure si l'on ne dispose pas d'informations démographiques (et d'une culture permettant de manier de telles informations) ? Le phénomène est-il lié, de façon mécanique et proportionnelle, à cette immigration ? Ou est-il lié à un phénomène interne à ce groupe social, à une transformation de sa culture du nom ?

D'où le problème de l'interprétabilité du corpus : qu'est-ce qu'un chercheur peut lui faire dire, dans la compétence qui est la sienne ?

On me fera observer que si c'était pour ne pas répondre à ces questions, je n'avais qu'à pas me les poser. Mais ce que je voudrais signaler ici, c'est l'acuité avec laquelle la recherche en AD me paraît exprimer une caractéristique des sciences humaines : leur inéluctable, excitante et dangereuse transversalité. Le corpus livre, dans une pseudo-évidence d'abord rassurante, un certain nombre de données qui comportent des ouvertures vers le social, l'historique, le politique, le démographique... Mais jusqu'où peut-on aller sans être sociologue, historien, politologue ou démographe ? On n'a pas toujours l'opportunité de pratiquer la rédaction collective et transdisciplinaire. Dans ces conditions, les corpus apparaissent comme des objets que l'on a intérêt à faire le plus riche possible, parce qu'en toute rigueur on ne peut se permettre d'en extraire qu'un nombre limité d'informations.

Finalement, sur cette question du nom propre, je n'avais pas beaucoup d'angles pour aborder la question. Il me restait celui de la morphologie lexicale. D'où le recentrage de l'article à partir de cette perspective :

« Aspects formels du changement de nom ». Les observations sociologiques, psychologiques, juridiques que j'ai tout de même proposées, et sans lesquelles je ne me serais pas situé en AD, m'ont donc fait sortir de ma spécialité, démarche qui ne va pas sans risque.

Problème de maniabilité

Un premier aspect de ce problème se manifeste lors de la constitution du corpus lui-même : il faut pouvoir manier les documents desquels il est appelé à se dégager. Il existe des documents auxquels on accède de façon directe, et d'autres auxquels on accède de façon indirecte, par l'intermédiaire de fonds, de collections privées, etc. Cela pose des problèmes de fiabilité qu'il convient de pallier.

Certains corpus sont particulièrement difficiles à construire : un corpus « images » peut exiger la numérisation soignée d'un grand nombre de documents, alors qu'aucune bibliothèque n'offre les facilités nécessaires. Dans un tel cas, il faut d'autant plus de temps pour réunir les données en question, et c'est ainsi qu'on finit par se heurter à l'obsolescence du matériel informatique : on découvre un jour que certains supports ne permettent plus d'exploiter commodément le stock constitué – par exemple, que tel ou tel périphérique ne permet pas d'extraire les images stockées sur différents disques. D'où parfois une course paradoxale au « vieux » matériel qui pourra sauver le résultat d'un long travail.

La sélection d'un échantillon rend maniable ce qui ne l'était pas, en tout cas par des moyens humains. Quelle forme matérielle celui-ci peut-il prendre ? Je ne travaille pas, je l'ai dit, sur des corpus informatisés à des fins lexicométriques, mais dépouille à la main, sélectionne et mets sur fiches. On se prive, en procédant ainsi, d'un certain nombre d'outils, notamment statistiques. Comment, dans ces conditions, apprécier en termes quantitatifs un phénomène – par exemple, la densité des métaphores guerrières ou relatives à la maladie dans le vocabulaire de Jean-Marie Le Pen ? Qu'est-ce que cela veut dire au juste quand on affirme que ces métaphores sont « nombreuses » ? L'appréciation qualitative des écarts, déjà délicate sur la base d'un corpus numérisé, devient franchement aléatoire quand celui-ci ne l'est pas. Pourtant, l'impression subjective de « fréquence » est bien là. Que faut-il en faire ?

Mais ajoutons qu'à défaut d'utiliser des moyens lexicographiques perfectionnés, on peut déjà effectuer des opérations utiles avec un logiciel

de traitement de texte (Word, par exemple). Cette démarche ne permet peut-être pas de tirer de conclusions plus acérées que celles auxquelles conduirait une méthode encore plus rudimentaire (fiches en carton et notes manuscrites), mais accélère énormément le travail. Tous mes corpus de textes ont été faits sous forme de tableaux, avec une entrée à gauche permettant le tri alphabétique. Dans le cas du corpus « noms propres », le tableau comporte quatre colonnes : à gauche, la date de la demande ; au centre, une colonne pour le nom abandonné, et une autre pour le / les noms sollicité(s) ; à droite, les éléments d'état-civil jugés pertinents pour l'analyse. Une fois entrées les 1562 demandes, deux ou trois manipulations simples ont démultiplié le corpus : classement chronologique des demandes, classement par fréquence des noms abandonnés, classement par fréquence des noms demandés. En outre, la fonction « recherche » a permis de repérer très vite le sort fait à certains noms ou à certains éléments constitutifs de noms : le morphème *ville*, la particule *de*, la forme *ben*, etc. Tout cela, sans aller bien loin, limite un peu la part de l'intuition, fait gagner du temps, et au bout du compte n'est pas inutile.

Problème de comparabilité

C'est un lourd travail de faire un corpus, et les conclusions que l'on en tire sont parfois limitées. Dans ces conditions, l'idéal est de préparer la suite, de faire des corpus qui se complètent les uns les autres et dont les conclusions peuvent être comparées.

En dépit d'un lien thématique (le discours sur l'altérité), il n'existe guère de comparabilité entre mon corpus Japon et mon corpus noms propres. C'est le genre de choix qui retentit sur le rythme des publications. Mais à supposer que se prolonge cette réflexion sur la patronymie, le principe de comparabilité est tout trouvé : après avoir travaillé sur les demandes parues au *JO*, il faudrait étudier, selon la même approche morphologique, les décrets autorisant le changement de nom, autrement dit la réponse que l'État apporte aux aspirations du citoyen. On dégagerait ainsi dans le détail les solutions formelles privilégiées par l'administration, et les principes qui, explicitement ou non, les inspirent.

Problème d'originalité

La littérature qui circule sur le patronyme est énorme. On trouve du reste des études formelles comme celle qui est évoquée ci-dessus, mais constituées, justement, à partir de corpus moins nettement circonscrits, écrites plutôt par des juristes ou par des historiens de la patronymie, et parfois sollicitées par le pittoresque. Au-delà des méthodes de l'analyse, comment mettre un peu de nouveauté dans l'approche de cette question, et apporter sa pierre à la réflexion collective ? En l'espèce, on a tenté de le faire dès la définition du projet et la constitution du corpus : d'où un travail sur les annonces du *JO* (c'est-à-dire sur la demande sociale directement exprimée, et parfois de façon surprenante) plutôt que sur les décrets (c'est-à-dire sur les réponses forcément positives de l'administration, puisque les refus, en la matière, ne se formulent que de façon officieuse). Il resterait beaucoup de choses à faire sur la base de ce corpus-là : décrire et interpréter, par exemple, l'évolution de la rubrique « Annonces » du *JO* depuis la guerre, c'est-à-dire le contexte formel des demandes de changement de nom. Ou encore, analyser l'évolution de la formulation de ces demandes (assez disparate après la guerre, elle devient de plus en plus technique et uniforme).

Cela dit, et en prenant les précautions nécessaires, on peut aussi exploiter un corpus qui existe déjà, et qu'on n'a pas soi-même constitué, simplement parce que sa richesse se prête à de nouvelles interrogations.

*

* *

Les lignes ci-dessus ne prétendaient pas passer en revue les interrogations théoriques relatives à la constitution des corpus. Rédigées au fil de la plume dans les conditions mentionnées en introduction, elles ne se voulaient que le témoignage d'une pratique et des interrogations qui l'accompagnent. Cette pratique est-elle représentative ? Au lecteur d'en juger. Ce qu'il y a de sûr, c'est qu'il existe un risque pour le chercheur à exposer ainsi les aspects les plus problématiques et les plus fragiles de son entreprise, et à révéler la part d'intuition qui le guide ainsi que les insuffisances qui en découlent. Mais il y a plus de risque encore à ne plus discuter de certaines questions, justement à cause de leur banalité. Peut-être est-il salubre de formuler pour soi, à nouveau, de temps en temps, en contrepoint d'un exercice quotidien tenté par les automatismes, quelques-unes des questions triviales que l'on s'était posé en débutant.