

Présentation

Marie-Anne PAVEAU et François PEREA

Depuis quelques années, les sciences du langage questionnent de manière renouvelée leurs corpus¹. Cette démarche est notamment ancrée dans l'évolution des technologies qui facilite de plus en plus les enregistrements audio et vidéo *in vivo*, la collecte de produits médiatiques ou de productions interpersonnelles médiatisées (en particulier sur les réseaux sociaux), le traitement des données à grande échelle ou encore le partage entre chercheurs.

Le sens de *sensible*

Si ces technologies (numériques notamment) permettent un accès croissant aux terrains, elles ne vont pas sans poser débat, notamment dans le champ éthique, puisqu'elles permettent la collecte, l'enregistrement, l'analyse et le partage de données pouvant être qualifiées de *sensibles*.

Dans ce contexte, *sensible* recoupe des phénomènes et des questionnements variés, très souvent intimement mêlés, qui ont en commun d'interpeller le chercheur dans sa démarche et ses rapports aux réalités et locuteurs observés. Qualifiant un corpus, l'adjectif peut renvoyer différemment ou de manière synthétique :

– à l'émotion du chercheur confronté à des thèmes anthropologiques forts, qui font écho en lui comme ils font écho chez tous les êtres humains : la mort, la sexualité, la liberté... et posent ainsi la question de l'engagement et de la distance (N. Elias, 1993 [1983]) par rapport à l'objet d'étude ;

– aux sujets parlants, lorsque l'approche refuse de les réduire à de simples instances locutrices désincarnées (« il y a quelqu'un », pour paraphraser R. Lafont), lorsque l'analyse se saisit de leurs identités, de leurs corps, de leurs sentiments... et commande au chercheur le plus impérieux des respects et la mesure des conséquences sociales et personnelles de sa recherche ;

– aux modalités méthodologiques de constitution de corpus sur des terrains peu évidents, obligeant à réfléchir aux cadres habituels et aux recettes toutes prêtes de la recherche, pour prendre en compte le biais que constitue alors le recueil des données et le risque qu'il comporte d'avoir des incidences sur les personnes.

D'une manière générale, *sensible* désigne ainsi toute recherche susceptible d'affecter – directement ou indirectement – les sujets et leurs rencontres dans le cadre de la recherche, et de conduire à des corpus affectés, *sensibilisés*. Le corpus sensible est donc celui qui ne laisse pas le chercheur indemne mais qui, au contraire, le touche, le modifie ou le mobilise.

Du côté des lois et règles

Une grande partie du débat est occupé par une réflexion sur les droits et les devoirs relatifs aux informations concernant les sujets enquêtés. Le chercheur dérangé par les données sensibles peut toujours changer de sujet de recherche. Par exemple, la Commission Nationale Informatique et Libertés (CNIL) distingue les « données sensibles » nécessitant un repérage et un traitement particulier :

¹ Voir à ce propos le numéro double des *Cahiers de praxématique*, 54-55 intitulé « Corpus, données, modèles », sous la direction de S. Azzopardi (2013).

Données sensibles : les « données à caractère personnel qui font apparaître, directement ou indirectement, les origines raciales ou ethniques, les opinions politiques, philosophiques ou religieuses ou l'appartenance syndicale des personnes, ou qui sont relatives à la santé ou à la vie sexuelle de celles-ci » (Art. 8 de la loi Informatique et libertés, dans CNIL, Guide : la sécurité des données personnelles).

Les journaux officiels de la République Française, emboîtant le pas de la CNIL, prévoient les sanctions financières voire des peines d'emprisonnement pour les contrevenants.

Côté recherche, les efforts se sont particulièrement concentrés sur le plan de la démarche, visant une conformité à la loi et à l'éthique. L'impératif de défense du droit des enquêtés, en partie inspiré par les instances éthiques biomédicales du continent nord-américain (A. Desclaux, 2008) a ainsi pu déboucher sur la proposition de « bonnes pratiques » largement commentées (cf. le monumental travail autour d'O. Baude, 2006), cherchant à allier impératifs pratiques et déontologiques.

Les choses semblent donc bien cadrées sous ces aspects.

Du côté des pratiques

Pourtant, il arrive que le chercheur se retrouve dans une situation aporique car le respect des enquêtés suppose aussi de rendre compte avec justesse de la réalité, de tendre vers « le respect [...] des données les plus exactes qu'il puisse atteindre » (J.-M. Eloy, 2000).

Que faire, alors, lorsque l'exactitude des données commande de déroger aux bonnes pratiques ? Ce paradoxe de l'observateur, pointé depuis longtemps par W. Labov, donne trop souvent lieu à des réponses techniques ou à des postures normatives qui, finalement, empêchent le débat de fond sur une approche scientifique des faits tels qu'ils sont et non tels que les enquêtés, par exemple, voudraient bien les montrer (A. Desclaux, 2008). Autrement dit, le chercheur peut-il s'abriter derrière « la Science » en majuscule ou doit-il se résoudre à s'éloigner des faits humains, délaissant l'éthique pour motifs... éthiques ?

De plus, les enquêtés ont-ils toujours intérêt à connaître les dessous de la recherche ? Par exemple lorsque l'annonce d'une enquête réalisée sur la dépendance, qui alimentera les démarches de soin et donc se situe dans le cadre d'une utilité sociale, conduit à stigmatiser et soumettre à une épreuve de réalité ceux qui se considèrent que comme de simples buveurs ou joueurs ? Ici, les trucages ne visent pas à tromper l'enquêté mais participent bien d'un principe de protection de ces sujets.

L'éthique confrontée au *sensible* s'adapte alors, se travaille, se justifie... sur la crête et dans le doute parfois.

Faute d'être réalisables dans le cadre des « bonnes pratiques », certaines constitutions de corpus demeurent difficilement faisables ou ont un statut particulier, définissant des objets de recherche qualifiés de « sensibles » voire de « scandaleux ». Ces objets d'études sont généralement ancrés dans des thèmes forts – parfois même tabous – d'un point de vue anthropologique : la maladie, le handicap, le sexe, la mort... qui posent, à côté de l'objet sensible (particulièrement lorsque « l'objet est aussi un sujet », F. Guay : 2009, p. 41), un sujet chercheur sensible, interpellé².

² Sauf à mettre une distance : la mort *dans la littérature* ou le sexe *dans la publicité*, pour ne donner que deux exemples, sont des sujets bien plus acceptables d'un point de vue institutionnel, notamment parce qu'ils ne supposent pas d'aller directement sur le terrain, au contact de ceux qui éprouvent.

La ligne de circonscription est alors commune. Le terrain « sensible » n'est plus alors défini par des positions éthiques légales mais en fonction d'une éthique morale, élaborée par la communauté scientifique instituée qui choisit les objets acceptables pour les chercheurs.

C. Lavigne (2007) en donne un exemple en racontant son sentiment de dépossession de son objet de recherche et de sa posture. La psychologue a longtemps travaillé sur le handicap avant de s'attacher à d'autres objets. Quelques années plus tard, après avoir eu une enfant sourde, elle décide de travailler sur les représentations de la surdité et de la parentalité (parent entendant/enfant sourd). La chercheuse raconte : « Des arguments dissuasifs ont été présentés par la majorité de mes collègues qui m'ont adressé conseils, mises en garde, voire interdictions méthodologiques et/ou déontologiques de mener une telle recherche » (2007, p. 31). Les arguments ont trait à l'abandon de la posture de chercheur à la faveur de celle de simple témoin, à la perte de crédit comme chercheuse et enseignante, au devoir d'aveu et à l'injonction à la transparence dans tous les écrits sur sa situation parentale ou encore l'impossible objectivité du fait du lien charnel.

P. Baudry (1997) fait, pour sa part, la pesée du regard de ses collègues, certains suspicieux (le chercheur doit sans doute participer à des « parties fines »), certains critiques (« ce n'est pas universitaire »), certains intéressés (a-t-il de bonnes adresses ?), lorsqu'il décide de travailler sur la pornographie.

Les deux exemples n'ont en commun que de présenter le rejet de la sensibilité (réelle ou supposée) du chercheur au contact de ses objets ou des sujets qu'il rencontre dans le cadre de ses investigations, alors que la démarche scientifique, objective voire positiviste, commanderait de l'évacuer.

Le parti pris du sensible

Au contraire, et certaines contributions du numéro s'y attachent, des chercheurs refusent cette forme de déni du sensible et posent la mesure de son interpellation comme processus dans la rigueur de la démarche.

Ils insistent alors sur la nécessité d'une réflexion approfondie, voire renouvelée, à porter sur les techniques d'enquêtes, ses méthodes de constitution et de reconstitution des corpus, les relations intersubjectives ou encore les modalités d'écriture de l'enquêté. Il ne s'agit plus de réfléchir au recueil de données « authentiques » (si d'aventure il existait un consensus sur cette qualification), mais d'observer la part sensible des enquêtés et de l'enquêteur telle qu'elle redéfinit les corpus d'une part, et les conceptions des rôles de chacun dans la démarche de recherche.

Comme toutes les recherches, parce que les réalités, les démarches et les terrains sont hétérogènes, les approches attachées à des « corpus sensibles » sont diverses et variées. Si elles ne se laissent pas saisir dans un cadre préétabli (quand bien même celui-ci existe, détermine les bonnes pratiques de manière nécessaire) parce qu'elles se situent sur le fil qui en délimite les contours, elles ont en commun de questionner nécessairement les rapports subjectifs entre les participants qui co-construisent, la réalité que se propose d'explorer la praxis scientifique.

Ce numéro n'apporte pas de solution sous forme d'un vade-mecum exportable.

Les contributions font écho à la complexité et à la diversité du questionnement en présentant plusieurs approches, expériences de terrain. Chacune aborde plusieurs aspects de ces corpus

sensibles et leurs incidences sur la recherche. Toutes se font écho par leur souci constant d'explicitier les postures et les démarches éthiques et épistémologiques, les impasses et les solutions trouvées, parfois différentes. Leurs auteurs – engagés dans le cadre général d'une linguistique empirique – exposent la singularité des situations de recherche dans lesquels ils se sont trouvés parfois, et l'ensemble fait émerger les lignes de force des différents rapports à ces corpus sensibles qui articulent :

- des aspects techniques liés aux corpus sensibles, dans une perspective critique ne se contentant pas, comme bien trop souvent sur le sujet, de décrire des dispositifs de prise de vue et de son et des difficultés d'encodage et de modélisation, mais interrogeant la démarche et l'objet même de collecte et d'enregistrement ;

- différentes postures éthiques, légales et morales (O. Baude, 2006 ; H. Doucet, 2010 ; M.-A. Paveau, 2013) qui prescrivent ou proscrivent des relations aux sujets et objets d'enquête résultant de la tension du chercheur à son travail et leur révélation dans l'espace public et scientifique ;

- une réflexion sur les états et mutations épistémologiques des sciences du langage, notamment influencés par les évolutions technologiques récentes, qui redessinent les observables, les objets, les discours disciplinaires.

Présentation des articles

A partir d'expériences d'enquête réalisées dans le cadre d'une recherche sur le français multiculturel parisien, F. Gadet et N. Kaci proposent de questionner l'interpellation de l'identité de l'enquêteur par l'enquêté. Ils soulignent comme les « identités fragiles » des personnes interrogées cherchent un écho à celle supposée de N. Kaci, sur la base d'origines supposées communes, pour faire sens et favoriser l'émergence de catégories que la recherche doit intégrer comme données sensibles.

Ce type d'interpellation est également présent dans le travail collectif de L. Alidières, B. Bonu, C. Charnet et L. Fauré. Les auteurs interrogent, dans le cadre d'observations réalisées en milieu carcéral (probablement inédites en sciences du langage, et commandant de réviser certains réflexes disciplinaires) les possibilités d'accéder à ce qui est par définition mis à l'écart voire caché. Ils exposent ainsi les difficultés liées à la constitution du corpus (accès, circulation très réglementée ; usage de la caméra en milieu carcéral) et les incidences sur les orientations participants dans l'interaction, notamment les effets de contamination.

Au secret de l'espace carcéral fait écho l'hyper-visibilité de l'espace numérique qu'arpente M.-A. Paveau, questionnant les « technodiscours » (c'est-à-dire les discours natifs en ligne) dans la perspective de l'éthique professionnelle d'une linguiste confrontée à de nouveaux objets discursifs. Dans cet espace clair-obscur que compose Internet, une attention particulière est notamment portée aux « scripteurs vulnérables » exposés entre privacité et publicité, et l'auteure propose une perspective écologique appuyé à l'épistémologie.

F. Perea poursuit l'investigation de l'espace numérique autour de travaux articulés à un sujet anthropologiquement fort : la sexualité en ligne (forums de masturbation, sites pornographiques, petites annonces de rencontre sans lendemain). Autour de deux préceptes majeurs des *bonnes pratiques* (l'anonymisation et le consentement éclairé), il montre comment les injonctions éthiques conduisent parfois le chercheur à déroger à la règle pour protéger les personnes observées et tenir une position morale.

Travaillant sur un thème tout aussi mobilisateur de réactions affectives, le deuil d'enfants, C. Ruchon, explore épitaphes, forums, livrets de condoléances et souligne les difficultés émotionnelles que soulève la démarche, auprès des interlocuteurs (le chercheur

« menaçant »), des collègues (le chercheur « morbide ») et questionne les postures du chercheur affecté par son corpus.

La réflexion sur les réactions du chercheur et son positionnement est également au cœur du travail de A. Arnold et L. Greco qui posent la question de leur légitimité à enquêter sur, pour et avec certains groupes situés aux marges d'un pouvoir dominant et marqués par une importante marginalisation et stigmatisation : familles homoparentales, drag king et trans. Ils investissent ainsi le champ de l'observation participante moins comme ressource méthodologique que comme objet d'analyse propre.

Leur questionnement croise ainsi celui de N. Marignier qui, travaillant sur les variations du développement sexuel, questionne l'« instabilité désignative » des sujets et le rôle du chercheur, comme celui des médias, dans l'inscription visible de ces personnes qui résistent aux tentatives de catégorisation et d'investigation. Son écriture même témoigne du souci du positionnement de la chercheuse dans son rapport à ces sujets, refusant de les objectiver et de les priver de leur capacité d'action.

L'ensemble des travaux réunis ici présente ainsi une diversité de terrain, de choix épistémologiques et méthodologiques, de réflexions et de parcours de chercheurs interpellés dans leur sensibilité, par des sujets ne se laissant pas réduire à des machines à parler, dans des contextes nécessitant de raisonner les réactions identitaires et affectives.

De cette hétérogénéité sourde des lignes de réflexions communes, aux conclusions parfois différentes mais toujours empreintes du souci de penser ce *sensible* s'imposant dans ces expériences de constitution et d'analyse des corpus. Toutes articulent le souci du respect des personnes et de la recherche dans des contextes pouvant leur porter préjudice, et la volonté de penser, au-delà des prescriptions, le métier de chercheur incarné.

BIBLIOGRAPHIE DE REFERENCE SUR LES CORPUS SENSIBLES

- AUROUX S., 1998, *La raison, le langage et les normes*, Paris, PUF.
- BACHELARD G., 1938, *La formation de l'esprit scientifique*, éd. Vrin.
- BARUS-MICHEL J., 1986, « Le chercheur, premier objet de la recherche », dans *Bulletin de psychologie*, tome 39 (16-18), n°377.
- BAUDE O. (coord.), 2006, *Corpus Oraux. Guide des bonnes pratiques*, éd. Presses universitaires Orléans / CNRS éditions.
- BAUDRY P., 1997, *La pornographie et ses images*, éd. Armand Colin, coll. « Agora / Pocket ».
- BOURDIEU P., CHAMBOREDON J.-C. et PASSERON J.-C., 1968, *Le métier de sociologue*, éd. Ecole Pratique des Hautes Etudes et Mouton Editeur.
- BOYER H., 2002, « Sociolinguistique : faire *corpus* de toute(s) voix ? », dans *Mots. Les langages du politique*, n°69. Mis en ligne et consulté de 30 juillet à l'adresse : <http://mots.revues.org/10553>.
- BRASSAC, (dir.), 2006, « Internalisme / Externalisme », *Intellectica* 43.
- BUTLER, J., 2008, *Le pouvoir des mots: discours de haine et politique du performatif*. Paris, France: Éditions Amsterdam.
- CAMERON D., FRAZER E., HARVEY P., RAMPTON B. et RICHARDSON K., 1992, *Researching Language. Issues of Power and Method*, éd. Routledge.
- CHARAUDEAU P., 2012, « Dis-moi ton corpus, je te dirai quelle est ta problématique », revue *Corpus*, n°8, Nice, 2009, consulté le 31 juillet 2012 sur le site de l'auteur : <http://www.patrick-charaudeau.com/Dis-moi-quel-est-ton-corpus-je-te.html>.
- CORI M. et DAVID S., 2008, « Les corpus fondent-ils une nouvelle linguistique ? », dans *Langage* n°171.

- COSTANTINI C., 2009, « Le chercheur : sujet-objet de sa recherche ? », *Cliopsy*, n°1.
- DERVIN F. & PAVEAU M.-A., 2012, “Quelle place pour les objets dans les sciences du langage et les sciences de la communication ?”, *Synergies pays de la Baltique* 9, revue du GERFLINT.
- DESCLAUX A., 2008, « L'éthique médicale appliquées aux sciences humaines et sociales : pertinence, limites, enjeux et ajustements nécessaires », dans *Bulletin de la Société de Pathologie Exotique*, vol. 101, n° 2.
- DEVEREUX G., 1967, *De l'angoisse à la méthode dans les sciences du comportement*, éd. Flammarion.
- DOUCET H., 2010, « De l'éthique de la recherche à l'éthique en recherche », dans *Ethique publique*, vol. 12, n°1.
- DURKHEIM E., 2009 (1ère éd. 1984), *Les Règles de la méthode sociologique*, éd. Payot, coll. « Petite Bibliothèque Payot ».
- ELIAS N., 1993 (1ère éd. 1983), *Engagement et distanciation*, éd. Fayard.
- ELOY J. -M., « Erkenntnistheorie und linguistische Ethik », *Grenzgänge* 14, pp. 116-130. Traduction par l'auteur : http://www.u-picardie.fr/LESCLaP/IMG/pdf/ethique_trad_cle0a16a6.pdf, consultée le 7 août 2012.
- GADET F., 2000, « Derrière les problèmes méthodologiques de recueil de données », dans *Les Cahiers de l'Université de Perpignan*, n° 31.
- GUAY F., 2009 ; « L'ethnographe et son objet : entre engagement et distanciation », dans les actes du colloque *Le chercheur et son objet : engagement et distanciation* ayant eu lieu le 1er mai 2009, à l'Université de Montréal : http://www.socio.umontreal.ca/ACSSUM/documents/Actes-colloque_volume-final.pdf, consulté le 10 août 2012.
- HARAWAY, D., 2009, *Des singes, des cyborgs et des femmes : la réinvention de la nature*. Paris: J. Chambon.
- HOLMES, M., 2009, *Critical Intersex*. Ashgate Publishing, Ltd.
- HUTCHINS E., 1994, *Cognition in the Wild*, Cambridge, MA: MIT Press.
- LABOV W., 1973, « Some principles of linguistic methodology », dans *Language and society*, n°1.
- LABOV W., 1982, « Objectivity and Commitment in Linguistic Science: the Case of the Black English Trial in Ann Arbor », *Language in Society*, n° 11.
- LAVIGNE C., 2007, « A qui appartient l'objet de la recherche ? Penser l'implication du chercheur dans son objet : le handicap (surdit ) », dans *Nouvelle revue de psychosociologie*, vol. 2, n°4, éd. Erès.
- MONDADA L., 2005, Rapport sur le projet « Pour une archive des langues parlées en interaction. Statuts juridiques, formats et standards, représentativité » financé par le *Programme Société de l'Information / Archivage et patrimoine documentaire*, mars 2005. http://icar.univ-lyon2.fr/projets/corinte/documents/Mondada_juridique_MARS05.pdf
- PAPERMAN P., 1992, « Les émotions et l'espace public », in *Quaderni* N°18 : 93-108.
- PAVEAU M.-A., 2009 : « Mais où est donc le sens ? Pour une linguistique symétrique », conférence invitée au deuxième colloque international *Res per nomen*, Reims, 30-31 mai, in *Actes* prépubliés, p. 21-31.
- PAVEAU M.-A., 2012 : “Ce que disent les objets. Sens, affordance, cognition“, dans *Synergies pays de la Baltique* 9, revue du GERFLINT.
- PAVEAU M.-A., 2012 (dir.) : “Texte, discours, interactions. Nouvelles épistémologies”, *Semen* 34.
- PAVEAU M.-A., 2012 : “Réalité et discursivité. D'autres dimensions pour la théorie du discours”, *Semen* 34.

- PAVEAU M.-A., 2013, *Langage et morale. Une éthique des vertus discursives*, Lyon, ENS ÉDITIONS.
- PAVEAU M.-A. 2013, « Genre de discours et technologie discursive. Tweet, twittécriture et twittérature », *Pratiques*, dossier coordonné par D. Ablali : « Modèles et théories du genre en confrontation ».
- PELLETIER M. L. et DEMERS M., 1994, « Recherche qualitative, recherche qualitative : expressions injustifiées », dans *Revue des sciences de l'éducation*, vol. 20, n°4, p. 757 à 771.
- PEREA F., 2012, « Les sites pornographiques par le menu : pornotypes linguistiques et procédés médiatiques », revue *Genre, sexualité et société*, n°7.
- PEREA F., 2013, « Les échanges dans les forums de masturbation internationaux : relations et scripts autour de l'acte corporel intime virtuel », dans *Les relations intimes internationales*, sous la direction de F. Dervin, éd. Archives contemporaines.
- PEREA F., 2014, « La figuration de soi dans les titres des petites annonces électroniques de rencontre sans lendemain », à paraître dans les actes de la journée d'étude « *Désirs et délices* ».
- PEREA F., 2013, « Fragmentation et saisissement pornographique : une approche par les légendes des sites web », dans la revue *Textualités numériques*, n°13, à paraître.
- RIGOT H., 2006, « (En)-jeux de corpus pour la recherche en SHS. Enoncés, textes et documents », dans F. RASTIER et BALLABRIGA M. (dir.), *Corpus en Lettres et Sciences sociales : des documents numériques à l'interprétation, Actes du colloque international d'Albi*, Paris, Texto, 2006. Disponible sur : <http://www.revue-texto.net/1996-2007/Parutions/Livres-E/Albi-2006/Rigot.pdf>.
- RICŒUR P., 1990, *Soi-même comme un autre*, éd. Seuil, coll. « Points ».
- SERIOT P., « Limites, bornes et normes : la délicate question de l'objet de connaissance en sciences humaines », dans *Colloque Géo Ponts*.
- SINCLAIR J., 1996, « Preliminary recommendations on Corpus Typology », rapport technique EAGLES (Expert Advisory Group on Language Engineering Standards), UE.
- TRAVERSO D., 2012, « Quelques repères juridiques pour les données à caractère personnel dans les banques de données de langue parlée en interaction », rapport de recherche par le groupe ICOR et P. Mollo, consulté le 2 août 2012 à l'adresse : <http://www.cil.cnrs.fr/CIL/spip.php?article1646>.
- WILLEMS D., 2000, « Diversité des domaines d'application » dans M. Blicher (éd.), *Corpus. Méthodologie et applications linguistiques*, éd. Honoré Champion.
- WILLIAMS G. (éd.), 2005, *La linguistique de corpus*, éd. Presses universitaires de Rennes.
- ARINO M., 2007, « Approche socio-sémiotique des logiques implicationnelles du chercheur en sciences de l'information et de la communication », dans *Esprit critique*, vol. 9, n°1.
- BILGER M. (éd.), 2000, *Corpus. Méthodologie et applications linguistiques*, éd. Honoré Champion.