Présentation

Josiane Boutet et Didier Demazière

La question des méthodes d'analyse des discours est de plus en plus abordée, discutée dans les sciences humaines et sociales. Ce mouvement résulte pour une part d'évolutions significatives du côté des corpus de données langagières : la croissance du volume des matériaux accessibles du fait du recours aux technologies d'enregistrement et de numérisation (Cappeau, Seijido, 2005) ; l'hétérogénéité des corpus accentuée par l'intérêt porté à des sources diversifiées élargissant le périmètre des matériaux considérés comme pertinents (Descamps, 2008) ; les enjeux et les potentialités de l'analyse secondaire de matériaux qualitatifs (Duchêne, à paraître). Parallèlement les méthodes d'analyse ont fait l'objet de réflexions renouvelées, orientées dans différentes directions : celle de l'automatisation des analyses textuelles et des difficultés croissantes pour se repérer dans le maquis des méthodes logicielles (Jenny, 1997) ; celle des usages des méthodes d'analyse des discours dans des cadres disciplinaires multiples (Bonnafous, Temmar, 2006) ; celle du statut épistémologique des méthodes et de leur irréductibles à des instruments externes aux opérations d'analyse (Demazière, 2006).

Un signe de cet intérêt croissant des chercheurs pour les discussions et réflexions méthodologiques nous a été apporté par nos auteurs eux-mêmes. En effet, la revue a reçu récemment plusieurs articles qui, de façon indépendante, abordaient chacun à leur façon différents aspects

du traitement des données, en sociologie comme en linguistique. Du fait de l'importance croissante des préoccupations méthodologiques dans la communauté scientifique, il nous a semblé intéressant de regrouper ces contributions dans un dossier thématique ; la revue *Langage et Société* ayant souvent publié des travaux en méthodologie mais n'ayant pas encore consacré de numéro spécifique à ces questions.

Les cinq articles ainsi rassemblés adoptent des positionnements méthodologiques distincts. En les réunissant dans un même dossier, nous souhaitons montrer au lecteur non seulement la diversité des méthodes actuellement disponibles, mais aussi la diversité de points de vue sur une même méthode.

Ainsi, le logiciel Alceste, très employé en sociologie, est ici discuté par Monique Dalud-Vincent d'un côté, et par Gérôme Truc de l'autre. L'une tente d'utiliser le logiciel pour analyser automatiquement un corpus d'entretiens semi-directifs très marqués par leur oralité : vingt entretiens portant sur la danse hip-hop auprès de jeunes artistes ou pratiquants. L'autre soumet au même logiciel un corpus de registres de condoléances : 60 000 messages électroniques rédigés à la suite des attentats du 11 mars 2004 à Madrid. Leurs articles ont l'intérêt d'argumenter des points de vue différenciés sur l'utilisation d'Alceste. M. Dalud-Vincent s'interroge sur la valeur ajoutée du traitement automatisé dans la mesure où les spécificités orales de son corpus sont mal ou pas prises en compte dans ce cadre. De plus, elle souligne qu'Alceste tend à « imposer son propre point de vue » au détriment de la problématique propre au chercheur. G. Truc souligne, à l'opposé, l'adéquation entre son corpus et le traitement par Alceste. Ce logiciel se révèle en effet tout à fait pertinent pour analyser un volume très important de textes, au caractère hautement répétitif, à la limite d'être des formules figées. La mise en regard de ces deux contributions montre clairement que la validité de techniques, comme le logiciel Alceste, est en lien étroit avec les propriétés, hautement variables, des corpus étudiés.

L'article de Wim Remysen ne se situe pas dans la problématique des analyses automatisées, mais dans le cadre théorique de l'Imaginaire linguistique, tel que proposé par la linguiste Anne-Marie Houdebine. Il s'agit d'un modèle d'analyse de données langagières diverses (textes écrits, entretiens, questionnaires) qui propose une grille de repérage de types de normes exprimées dans des discours. Le corpus analysé selon ce modèle est composé de 31 chroniques de langage, publiées dans la presse québécoise entre 1865 et 1996. Ces textes permettent à l'auteur d'avoir accès aux différentes normes régissant le français québécois durant cette période historique. L'auteur dégage différents types d'arguments utilisés

par les chroniqueurs pour caractériser le français du Québec. Il les associe ensuite aux catégories de normes proposées par le modèle de l'Imaginaire linguistique. En ce sens, cette méthode propose un instrument intellectuel reproductible d'aide à l'interprétation ; et non pas comme Alceste, un outillage à l'analyse automatisée des données.

Isabelle Hajek se pose en sociologue la question des relations entre cognition et langage. Son corpus est composé de 283 documents écrits de type divers (tracts, rapports, courriers, pétitions, etc.), produits dans le cadre d'une mobilisation contre un projet d'implantation d'un incinérateur de déchets en Provence. Son projet est d'y identifier des « grammaires langagières » qui sont mobilisées par une multiplicité d'acteurs afin d'exprimer des positionnements distincts dans le débat. Le cadre théorique mobilisé, celui de J.-M. Ferry, pose l'existence de quatre grammaires langagières (iconique, indiciaire, propositionnelle et discursive). C'est ce modèle théorique qui ici constitue, comme pour W. Remysen, une méthode d'interprétation outillant le classement des matériaux.

Dans le dernier article, Dominique Maingueneau met en débat la notion de formation discursive, qui fut centrale dans l'analyse de discours française. Il montre comment elle a évolué, en lien avec les transformations advenues dans le champ des études sur le discours. Ce faisant, il souligne l'instabilité actuelle de cette notion. Pour y pallier, l'auteur propose de distinguer entre deux grands types d'unités d'analyse : les unités topiques et les unités non-topiques. Il dessine ainsi un cadre méthodologique d'analyse des discours, de classement ou de catégorisation, qui ne s'appuie pas sur un découpage *a priori* des discours : ce travail reste de la responsabilité du chercheur, à la différence des méthodes automatisées d'analyse.

Nous présentons ainsi une sorte de continuum méthodologique, depuis des techniques automatisées dans lesquelles le travail réflexif du chercheur est pour l'essentiel repoussé après les analyses, et se situe au moment de l'interprétation des résultats quantitatifs, jusqu'à des cadres conceptuels conçus pour guider, d'une façon plus ou moins opératoire, le travail d'analyse et d'interprétation du chercheur.

Références bibliographiques

Bonnafous S. & Temmar M. (dirs), (2007), *Analyse du discours et sciences humaines et sociales,* Paris, Éditions Ophrys.

Cappeau P. & Seijido M. (2005), *Les corpus oraux en français (inventaire 2005 v. 1.0)*. Consulté le 15 janvier 2011 dhttp://www.culture.gouv.fr/culture/dglf/recherche/corpus_parole/Presentation_Inventaire.

Demazière D. (dir.), (2006), *Analyses textuelles en sociologie. Logiciels, méthodes, usages*, Rennes, Presses Universitaires de Rennes.

Descamps F. (dir.), (2006), *Les sources orales et l'histoire. Récits de vie, entretiens, témoignages oraux*, Paris, Bréal.

Duchêne S. (à paraître), « Développement de l'analyse secondaire et des méthodes d'analyse qualitative : une chance à saisir ? », dans Brugidou M., Dargentas M., Le-Roux D. & Salomon A.C. (dirs), *Analyse secondaire en recherche qualitative : enjeux pour les sciences humaines et sociales*, Paris, Éditions Lavoisier.

Jenny J. (1997), « Méthodes et pratiques formalisées d'analyse de contenu et de discours dans la recherche sociologique française contemporaine. État des lieux et essai de classification », *Bulletin de Méthodologie Sociologique (B.M.S.)*, 54, p. 64-112.