

L'histoire et l'historien face au quantitatif

Claire Lemerrier et Claire Zalc, chargées de recherche au CNRS (IHMC)

Claire.Lemerrier@ens.fr, Claire.Zalc@ens.fr

<http://lemercier.ouvaton.org/document.php?id=155>

Jeudi de 10h30 à 12h30, Bd Jourdan, salle 8 ou salle de réunion du bâtiment B.

26 avril 2007

Histoire des mots, histoire des textes

Des pionniers aux manuels

- Deux bilans historiographiques récents en ligne par Jacques Guilhaumou, très riches surtout sur le plan théorique : « Le corpus en analyse de discours : perspective historique », *Corpus*, novembre 2002, <http://corpus.revues.org/document.html?id=8> et « Où va l'analyse de discours ? Autour de la notion de formation discursive », *Texte !*, juin 2004, http://www.revue-texto.net/Inedits/Guilhaumou_AD.html
- Régine Robin, *Histoire et linguistique*, Paris, Armand Colin, 1973 [cela reste un très beau manuel avec une vraie réflexion de fond et quelques études de cas].
- Jacques Guilhaumou, Denise Maldidier, Antoine Prost, Régine Robin, *Langage et idéologies. Le Discours comme objet de l'Histoire*, Paris, Éditions ouvrières, 1974 [reprint d'un dossier du *Mouvement social*, recueil encore très recommandable]
- Antoine Prost, Louis Girard et Rémi Gossez, *Vocabulaire des proclamations électorales de 1881, 1885 et 1889*, Paris, PUF-Publications de la Sorbonne, 1974 [vraie réflexion sur la source et études fines, dont des analyses factorielles, axées sur la polarisation partisane et l'évolution des discours dans le temps : impressionnant, surtout pour l'époque !].
- Charles Muller, *Principes et méthodes de statistique lexicale*, Paris, H. Champion, 1977/1993 [clair et efface sur les définitions de linguistique].
- Marie-Anne Polo de Beaulieu, « Panorama de la lexicométrie », *Histoire & Mesure*, II-3/4, 1987, p. 173-197 [un état des lieux précis de la recherche à ce moment].
- Ludovic Lebart et André Salem, *Analyse statistique des données textuelles : questions ouvertes et lexicométrie*, Paris, Dunod, 1988 [axé sur l'analyse de questions ouvertes en sociologie, classique et clair : analyses factorielles, classifications, segments répétés].
- Didier Demazière, Claire Brossaud, Patrick Trabal et Karl Van Meter (dir.), *Analyses textuelles en sociologie. Logiciels, méthodes, usages*, Rennes, PUR, 2006 [large panorama de logiciels présentés de façon pratique et critique]

Quelques études de cas

- Pierre-Paul Zalio, « Territoires et activités économiques. Une approche par la sociologie des entrepreneurs », *Genèses*, n° 56, 2004, p. 4-27. [étude avec Alceste d'entretiens biographiques].
- « Mesurer le texte », *Histoire & Mesure*, XVIII-3/4, 2003, <http://histoiremesure.revues.org/sommaire1612.html> et voir l'index des articles de lexicométrie de la revue (prochainement tous en ligne) : <http://histoiremesure.revues.org/entree.html?id=195>
- Dominique Labbé et Denis Monière, *Le discours gouvernemental. Canada, Québec, France (1945-2000)*, Paris, H. Champion, 2003 [court, avec une vraie problématique, des techniques fines (arbres...) et clairement présentées].
- Jean-Pierre Courtial et Léna Kerneur, « La méthode des mots associés, outil d'analyse du changement social », *Histoire & Mesure*, XII-3/4, 1997, p. 251-270 [étude d'associations de mots-clés – cela suppose un pré-codage ; présentation du logiciel Leximappe].
- Carmen Beatriz Loza, « De la classification des Indiens à sa réfutation en justice (Yucay, Andes péruviennes, circa 1493-1574) », *Histoire & Mesure*, XII-3/4, 1997, p. 361-386 [réflexion sur la stratification sociale à partir d'une analyse quantifiée des discours]

L'histoire et l'historien face au quantitatif

Claire Lemerrier et Claire Zalc, chargées de recherche au CNRS (IHMC)

Claire.Lemerrier@ens.fr, Claire.Zalc@ens.fr

<http://lemercier.ouvaton.org/document.php?id=155>

Jeudi de 10h30 à 12h30, Bd Jourdan, salle 8 ou salle de réunion du bâtiment B.

- Philippe Dautrey, « Écrire sa guerre. Analyse d'un carnet de guerre », VII-3/4, 1992, p. 249-280 [à propos de la Première Guerre mondiale, un bon exemple d'étude surtout qualitative mais s'appuyant à bon escient sur des comptages : pronoms personnels, usages de « mais »...].
- Mark Olsen, « Gender representation and *histoire des mentalités*: Language and Power in the *Trésor de la langue française* », *Histoire & Mesure*, VI-3/4, 1991, p. 349-373 [étude très fine des contextes les plus fréquents de mots comme « femme », « sexe » ou « efféminé » au fil des siècles].
- Pour voir beaucoup d'études plutôt littéraires, aller voir sur le rayon de la bibliothèque d'Ulm correspondant à la cote LP fr 321 G/GA 8° la collection « Travaux de linguistique quantitative », puis « Lettres numériques » de l'éditeur H. Champion.
- La revue *Mots* s'est créée autour de l'étude du vocabulaire politique, consulter ses sommaires vaut la peine ; dans les numéros récents, on trouve toutefois peu d'études quantifiées.

Information plus « technique »

- Présentation de nombreux logiciels : <http://www.weblettrés.net/sommaire.php?entree=20&rubrique=75&sousrub=251>
- Un point technique intéressant et clair sur la lemmatisation, par Damon Mayaffre : <http://www.uottawa.ca/academic/arts/astrolabe/articles/art0048/Logometrie.htm>
- Éléments variés, en particulier bibliographiques, sur Alceste : http://www.melissa.ens-cachan.fr/rubrique.php3?id_rubrique=91 ; tutoriel (plus technique, un peu daté) par l'auteur du logiciel ici : http://www.univ-tlse2.fr/erc/Semin_Reinert_1.html.
- Présentation rapide de plusieurs logiciels et présentation plus précise de SPAD-T : http://www.melissa.ens-cachan.fr/article.php3?id_article=284
- Présentation/comparaison rapide et claire d'Hyperbase et Lexico 3 : Emmanuel Bonin et Alain Dallo, « Hyperbase et Lexico 3, outils lexicométriques pour l'historien », *Histoire & Mesure*, XVIII-3/4, 2003, p. 389-402. <http://histoiremesure.revues.org/document840.html>
- [ancien : le nouveau est moins riche !] site du laboratoire de Nice « Bases, corpus, langages », comprenant une présentation très claire du logiciel Hyperbase : <http://ancilla.unice.fr/default.html>
- Page web de Lexico 3 : <http://www.cavi.univ-paris3.fr/ilpga/ilpga/tal/lexicoWWW/lexico3.htm>
- Site du logiciel Prospero (exploration assistée de gros corpus de controverses contemporaines) : http://prospero.dyndns.org:9673/prospero/acces_public
- Présentation de Tropes, logiciel de recherche thématique (donc plus adapté aux textes récents) : <http://www.acetic.fr/tropes.htm>

Sites proposant des corpus historiques

- <http://www.atilf.fr/> : un peu de contenu en libre accès (dictionnaires anciens notamment), beaucoup en accès restreint (l'ENS est abonnée) : *L'Encyclopédie*, Frantext...
- <http://ancilla.unice.fr/~brunet/BALZAC/balzac.htm> : tout Balzac