

Bibliographie

- ABNEY, Steven. "Part-of-Speech Tagging and Partial Parsing" in Church, Ken, Young, Steve et Bloothoof, Gerrit (eds.), *Corpus-Based Methods in Language and Speech*. Dordrecht, Kluwer Academic Publishers, 1996.
- ACHARD, Pierre. *La sociologie du langage*. Paris, Presses Universitaires de France, 1995.
- ADAM, J.-M. *Les textes: types et prototypes. Récit, description, argumentation, explication et dialogue*. Paris, Nathan Université, 1992.
- ADAM, J.-M. *Linguistique textuelle. Des genres de discours aux textes*. Nathan, Paris, 1999.
- ADDA, Gilles et al. « L'action GRACE d'évaluation de l'assignation des Parties du Discours pour le français » in *Langues*, vol. 2, No 2, 119-129, juin 1999.
- ADDA, Gilles et al. « Métrique et premiers résultats de l'évaluation GRACE des étiqueteurs morphosyntaxiques pour le français » in *6ème conférence sur le Traitement Automatique du Langage Naturel (TALN99)*, Cargèse, France, juillet 1999.
- ALLISON, Desmond. "Book review (Academic writing in a second language: Essays on research & pedagogy, ed. D. Belcher & G. Braine; Norwood, NJ: Ablex, 1995, and Genre knowledge in disciplinary communication: Cognition/culture/power, by C. Berkenkotter & T.N. Huckin; Hillsdale, NJ: Lawrence Erlbaum, 1995)" in *SAAL Quarterly (Singapore Association for Applied Linguistics)*, 37: 10-14, 1997.
- AMORIM, Marilia. *Dialogisme et altérité dans les sciences humaines*. Editions l'Harmattan, Paris, 1996.
- ANDERSON, P. V., R. J. BROCKMANN, et C. R. MILLER (éds.). *New Essays in Technical and Scientific Communication: Research, Theory, and Practice*. Farmingdale: Baywood, 1983.
- ARISTOTE. *La poétique*. Seuil, Paris, 1980.
- ATWELL, Eric et al. "A comparative evaluation of modern English corpus grammatical annotation schemes" in *ICAME Journal*, volume 24, pages 7-23, International Computer Archive of Modern and medieval English, Bergen, 2000.
- AURAY, J.-P., DURU, G. et ZIGHED, A. *Analyse des données multidimensionnelles*. Vol. 1 à 4. Mathématiques Appliquées N°1. Editions Alexandre Lacassagne, Lyon, 1990-1.
- BAKER, Mona (ed.), FRANCIS, Gill (ed.) et TOGNINI-BONELLI, Elena (ed.). *Text and technology, in honour of John Sinclair*. John Benjamins Publishing Company, Philadelphia/Amsterdam, 1993.
- BAKHTIN, M.M. *Speech Genres and Other Late Essays*. Trad.: Vern. W. McGee. University of Texas Press, Austin; 1986.
- BAUMAN, Richard et Joel Sherzer (eds.). *Explorations in the Ethnography of Speaking*. Cambridge University Press, Cambridge, 1974.
- BAZERMAN, Charles. *Shaping written knowledge: The genre and activity of the experimental article in science*. University of Wisconsin Press, Madison, 1988.

- BEAUDOUIN, Valérie. *Statistique textuelle : une approche empirique du sens à base d'analyse distributionnelle*. Texto ! septembre 2000 [en ligne]. Disponible sur : <http://www.revue-texto.net/Inedits/Beaudouin_Statistique.html>. (Consultée le 26/01/06).
- BEAUVISAGE, Thomas. « Exploiter des données morphosyntaxiques pour l'étude statistique des genres textuels : application au roman policier » in TAL, Vol. 42, n°2/2001.
- BENVENISTE E. *Problèmes de linguistique générale*. Tomes I et II. Paris Gallimard, 1966/1974.
- BENZÉCRI, J.-P. « Introduction I », *Pratique de l'analyse des données, Linguistique et lexicologie*. Dunod, Paris, 1981.
- BENZÉCRI, J.-P. et coll. *Pratique de l'analyse des données*. Dunod, Paris, 1981.
- BERGMANN, Jorg R. et LUCKMANN, Thomas. "Reconstructive genres of everyday communication" in U. Quasthoff (ed.), *Aspects of Oral Communication*. De Gruyter, Berlin, 1995.
- BERKENKOTTER, Carol et HUCKIN, Thomas. *Genre knowledge in disciplinary communication : Cognition/culture/power*. N.J. Lawrence Erlbaum, Hillsdale, 1995.
- BERTHELOT, Jean-Michel. *Figures du texte scientifique*. PUF, Paris, 2003.
- BHATIA, Vijay K. *Analysing genre: language use in professional settings*. Longman, London, 1993.
- BIBER D., *Variation across Speech and Writing*. Cambridge University Press, Cambridge, 1988.
- BIBER, D., CONRAD, Susan et REPPEN, Randi. *Corpus linguistics. Investigating language structure and use*. Cambridge University Press, Cambridge, 1998-2002.
- BIGERT, Johnny, KNUTSSON, Ola and SJOBERGH, Jonas, "Automatic Evaluation of Robustness and Degradation in Tagging and Parsing" in RANLP 2003, Borovets, Bulgaria, 2003.
- BIZZELL, Patricia. *Academic discourse and Critical Consciousness*. University of Pittsburg Press, Pittsburg, 1992.
- BOURDIEU, Pierre, CHAMBOREDON, Jean-Claude et PASSERON, Jean-Claude. *Le Métier de sociologue*. Editions Mouton-Bordas, Paris-La Haye, 1968.
- BOURDIEU, Pierre. « Le Champ scientifique » in Actes de la recherche en sciences sociales, 1976, 2/3, pp. 88-104.
- BOURDIEU. *Langage et pouvoir symbolique*. Seuil, 2001.
- BOURE, Robert. « Produire une revue scientifique. Le cas de Sciences de la Société » in Renzetti, Françoise (coord.), *Stratégies informationnelles et valorisation de la recherche scientifique publique*. Editions ADBS, 1998, pp. 105-119.
- BOURION, Evelyne. *L'aide à l'interprétation des textes électroniques*. Th. : ling. : Nancy II : 2001 : 210 p.
- BOUVIER, Alain. « Un paradigme caché en sociologie de la connaissance scientifique: le paradigme Mill-Pareto? » in Busino, Giovanni (dir.). *Sciences et techniques dans la*

- société*. XIIIe colloque annuel du Groupe d'étude « Pratiques Sociales et Théories ». Revue européenne des sciences sociales. Tome XXXV, n°108, 1997, pp. 15-30.
- BRANCA-ROSOFF, Sonia. « Des innovations et des fonctionnements de langue rapportés à des genres » in *Langage et Société* n°87, Maison des Sciences de l'Homme, Paris, 1999.
- BRANCA-ROSOFF, Sonia. « Types, modes et genres : entre langue et discours » in *Langage et Société* n°87, Maison des Sciences de l'Homme, Paris, 1999.
- BRANTS, Thorsten. « TnT – A Statistical Part-of-Speech Tagger » in *Proceedings of the 6th Applied NLP Conference, ANLP-2000*, April 29-May 3, 2000, Seattle, WA.
- BREIVEGA, K.R., DAHL, T. & FLØTTUM, K. “Traces of self and others in research articles. A comparative pilot study of English, French and Norwegian research articles in medicine, economics and linguistics” in *Journal of Linguistics*, Vol. 12 (2), 2002.
- BRONCKART, Jean-Paul. *Activité langagière, textes et discours, Pour un interactionisme socio-discursif*. Editions Delachaux et Niestlé, Lausanne-Paris, 1996.
- BRONCKART, Jean-Paul. *Le fonctionnement des discours, Un modèle psychologique et une méthode d'analyse*. Editions Delachaux et Niestlé, Lausanne-Paris, 1994
- BRUNET E. & MULLER C. 1988. « La statistique résout-elle les problèmes d'attribution ? » in *Strumenti critici*, Sett., 367-387.
- BUGARSKI, Ranko. “Applied Linguistics as Linguistics Applied” in Miseska Tomic, Olga et Shuy, Roger W. *The Relation between Theoretical and Applied Linguistics*. Plenum Press, N.Y., 1987, pp. 3-19
- BUSINO, Giovanni (dir.). *Sciences et techniques dans la société*. XIIIe colloque annuel du Groupe d'étude « Pratiques Sociales et Théories ». Revue européenne des sciences sociales. Tome XXXV, n°108, 1997.
- CARLETTA, J et al. (2005). “A generic approach to software support for linguistic annotation using XML” in G. Sampson and D. McCarthy (eds.), *Readings in Corpus Linguistics*. London and NY: Continuum International.
- CHANOD, Jean-Pierre et TAPANAINEN, Pasi. “Tagging French – comparing a statistical and a constraint-based method” in *Proceedings of ACL Conference*, Dublin, 1995.
- CLARK, Stephen, CURRAN, James et OSBORNE, Mike. “Bootstrapping POS Taggers using Unlabelled Data” in Daelemans, Walter et Osborne, Miles (eds.), *Proceedings of CoNLL-2003*, pages 49--55. Edmonton, Canada, 2003
- CLYNE, Michael. “Cultural differences in the organization of academic texts: English and German” in Cheshire, J. et Trudgill, P. *The Sociolinguistics Reader*, Vol. 2, Arnold, 1998, pp. 315-347.
- COHEN W. et HIRSCH, H. “Joins that generalize: text classification using WHIRL” in AGRAWAL, R. et al. (eds.), *Proceeding of KDD-98, 4th International Conference on Knowledge Discovery and Data Mining*, AAAI Press, Menlo Park, New York, US, 1998, pp. 169-173.
- CONDAMINES, Anne (ed.). *Sémantique et corpus*. Coll. Cognition et Traitement de l'Information. Lavoisier, 2005.

- COPECK, Terry, BARKER, Ken, DELISLE, Sylvain et SZPAKOWICZ, Stan. "Automating the Measurement of Linguistic Features to Help Classify Texts as Technical" in *Conférence TALN 2000*, Lausanne, 16-18 octobre 2000.
- CUSIN-BERCHE, Fabienne. « Le statut de la référence dans les discours scientifiques et techniques » in *Les mots et leurs contextes*. Presse Sorbonne Nouvelle, Paris, 2003, pp. 93-103.
- CUSIN-BERCHE, Fabienne. « Qu'est-ce qu'un texte spécialisé ? » in *Les mots et leurs contextes*. Presses Sorbonne Nouvelle, Paris, 2003, pp. 105-123.
- CUTTING, Doug, KUPIEC, Julian et SIBUN, Penelope. "A practical Part-of-Speech Tagger" in *Proceedings of the Third Conference on Applied Natural Language Processing*, ACL, Trento, Italy, 1992.
- DÉJEAN Hervé (2000). "How to Evaluate and Compare Tagsets? A Proposal" in *Proceedings of the Second International Conference on Language Resources and Evaluations*, LREC'2000, Athens.
- DESROSIERES, Alain. *La politique des grands nombres. Histoire de la raison statistique*. La Découverte & Syros, Paris, 1993-2000.
- DHILLON I. S., MALLELA S., KUMAR R. « A divisive information theoretic feature clustering algorithm for text classification » in *Journal of Machine Learning Researches*, vol. 3, 2003, p. 1265-1287, MIT Press.
- DIDAY, E. et al. *Éléments d'analyse de données*. Dunod, Bordas, Paris, 1982.
- DOCHERTY M. et SMITH R. "The case for structuring the discussion of scientific papers" in *BMJ* 1999; 318: 1224-1225.
- DOLININE, Constantin. « Le problème des genres du discours quarante-cinq ans après Bakhtine » in *Langage et Société* n°87, Maison des Sciences de l'Homme, Paris, 1999.
- DUDLEY-EVANS, Tony. "The English for International Students Unit at the University of Birmingham, UK". Report by Les Kirkham on a presentation at the "LSP Forum '99", Charles University, Prague, 17-19 September, 1999.
- DUMAIS, S. "Using SVMs for text categorization" in *IEEE Systems*, 1998.
- DUMAIS S., PLATT J., HECKERMAN D., SAHAMI M. "Inductive learning algorithms and representations for text categorization" in *CIKM '98 : Proceedings of the seventh international conference on Information and knowledge management*, ACM Press, 1998, p. 148-155.
- DUSZAK, Anna (ed.). *Culture and Styles of Academic Discourse*. Mouton de Gruyter, Berlin, 1997.
- EAGLES. *Preliminary Recommendations on Text Typology*. EAGLES Document EAG-TCWG-TTYP/P, Version de Juin 1996.
- EAGLES. *Recommendations for the Morphosyntactic Annotation of Corpora*. EAGLES Document EAG-TCWG-MAC/R. Mars 1996. Disponible sur <<http://www.ilc.cnr.it/EAGLES96/annotate/annotate.html>> (consultée le 15/04/06).
- EGGS, Ekkehard. *Grammaire du discours argumentatif*. Editions Kimé, Paris, 1994.
- ESCOFIER, Brigitte et PAGES, Jérôme. *Analyses factorielles simples et multiples : Objectifs, méthodes et interprétation*. Dunod, 1998.

- EVANGELISTI ALLORI, Paola. *Academic Discourse in Europe, Thought Processes and Linguistic Realisations*. Proceedings of the Workshop held at the Terza Università degli Studi di Roma, 29-30 October, 1994.
- FLØTTUM, Kjersti. "Personal English, indefinite French and plural Norwegian scientific authors? Pronominal author manifestation in research articles" in *Norsk Lingvistisk Tidsskrift* 21 (1), 2003, pp. 21–55.
- FOLCH, Helka, HEIDEN, Serge et Habert, Benoît. "TypTex: Inductive typological text classification by multivariate statistical analysis for NLP systems tuning/evaluation" in *Second International Conference on Language Resources and Evaluation*, pp. 141-148. Ed: Maria Gavridiliou et al., volume 1, Athènes (Grèce), 31 mai–2 juin 2000.
- FREEDMAN, Aviva et MEDWAY, Peter (eds.). *Genre and the New Rhetoric*. Taylor & Francis, Londres, 1994.
- FUCHS, C. et HABERT, B. (eds.). *Traitement automatique et ressources numérisées pour le français*. *Le français moderne*, 72(1), 2004.
- FULLER, Steve. *The governance of science: ideology and the future of the open society*. Open University Press, Philadelphia, 2000.
- GARSIDE, R. "The CLAWS word-tagging system" in Garside R., Leech, F. and Sampson, G. (eds.), *The Computational Analysis of English*. Longman, 1987.
- GENETTE, G. *Théorie des genres*. Seuil, Paris, 1986.
- GILE, Daniel (2002). « L'interdisciplinarité en traductologie: une optique scientométrique » in colloque *Relations d'interdisciplinarité en traduction*.
- GLASER, Rosemarie. "Summarizing Texts as Genres of Academic Writing" in *Summarizing Text for Intelligent Communication*, Dagstuhl Seminar 13.12.-17.12.1993, Dagstuhl, Germany.
- GOBERT, Frédéric. *Glossaire bibliographique des sciences du langage*. Panormitis, 2001.
- GOPNIK, Myrna. *Linguistic Structures in Scientific Texts*. Editions Mouton, The Hague, 1972.
- GRAFTON, Anthony. *The Footnote: A curious history*. Harvard University Press, 1999.
- GRAÑA Jorge , BARCALA, Mario Fco., and VILARES, Jesús. "Formal Methods of Tokenization for Part-of-Speech Tagging" in Gelbukh, Alexander (ed.), *Computational Linguistics and Intelligent Text Processing*, volume 2276 of Lecture Notes in Computer Science, pp. 240-249, Springer-Verlag, Berlin-Heidelberg-New York, 2002.
- GROSSMANN F. & RINCK F. 2004. « La surénonciation comme norme du genre. L'exemple de l'article de recherche et du dictionnaire en linguistique » in Rabatel (ed.) *Effacement énonciatif et discours rapporté*. *Langages* 156. Paris : Larousse, 34-50.
- HABERT, Benoît. « Création de dictionnaires sémantiques et typologie des textes » in Tyvaert, Jean-Emmanuel (ed.). *Actes des journées scientifiques 1999 L'Imparfait - Philologie électronique et assistance à l'interprétation des textes*, CIRLEP, Reims, 2000, pp. 171-188.
- HABERT, Benoît. « Des corpus représentatifs : de quoi, pour quoi, comment ? » in Bilger (ed.), *Linguistique sur Corpus. Etudes et réflexions*, Cahiers de l'Université de Perpignan, numéro 31, pp. 11-58, 2000.

- HABERT, Benoît. « Linguistique et informatique : nouveaux défis » in *Revue française de linguistique appliquée*, IX-1, De Werelt, Amsterdam, 2004.
- HABERT, Benoît. « Portrait de linguiste(s) à l'instrument » in *Texto!* [en ligne], décembre 2005, vol. X, n°4. Disponible sur : <http://www.revue-texto.net/Corpus/Publications/Habert/Habert_Portrait.html>. (Consultée le 15/04/06).
- HABERT, Benoît. *Instruments et ressources électroniques pour le français*. Collection L'essentiel français, Ophrys, Gap/Paris, 2005.
- HABERT, B., FABRE, C. et ISSAC, Fabrice. *De l'écrit au numérique : constitution, annotation et analyse de corpus*. Masson, Interéditions, 1997.
- HABERT, B., ILLOUZ, G, LAFON, P. *et al.* « Profilage de textes : cadre de travail et expérience » in *JADT 2000*, 5es Journées Internationales d'Analyse Statistique des Données Textuelles.
- HABERT, Benoît, NAZARENO, Adeline et SALEM, André. *Les linguistiques de corpus*. Editions Armand Colin/Masson, Paris, 1997.
- HABERT, Benoît et SALEM, André. « L'utilisation de catégorisations multiples pour l'analyse quantitative de données textuelles » in *TAL*, vol. 36, n1-2, 1995, pp. 249-276. – *Traitements probabilistes et corpus*, Benoît Habert (resp.).
- HALLIDAY et HASAN. *Cohesion in English*. Longman, 1976.
- HALLIDAY, Michael A.K. *An introduction to Functional Grammar*. Seconde édition. Edward Arnold, Londres, 1994.
- HALLIDAY, Michael et HASAN, Ruqaiya. *Language, Context and Text: Aspects of Language in a Social-Semiotic Perspective*. Oxford University Press, Oxford, 1989.
- HAMBURGER, K. *La logique des genres*. Seuil, Paris, 1986.
- HARTMANN, R.R.K. *Contrastive Textology, Comparative Discourse Analysis in Applied Linguistics*. Julius Groos Verlag, Heidelberg, 1980.
- HERREMAN, Alain. L'analyse sémiotique des textes in *Éléments d'histoire sémiotique de l'homologie*. *Texto !* avril 1997 [en ligne]. Disponible sur : <http://www.revue-texto.net/Inedits/Herreman/Herreman_Elements.html>. (Consultée le 15/04/06).
- HEYLIGHEN, Francis et DEWAELE, Jean-Marc. “Variation in the contextuality of language: an empirical measure” in *Foundations of Science*. Ide, Nancy and Véronis, Jean. *Word sense disambiguation: the state of the art*, 1998.
- HOFMANN T. “Probabilistic Latent Semantic Indexing” in *Proceedings of the 22nd Annual ACM Conference on Research and Development in Information Retrieval*, Berkeley, California, August 1999, p. 50-57.
- HUCKIN, Thomas N. and OLSEN, Leslie A. *Technical Writing and Professional Communication. For Nonnative Speakers of English*. New York, McGraw Hill, Inc. Second edition, 1991.
- HYLAND, Ken. *Hedging in scientific research articles*. City University of Hong Kong, 1998.
- ILLOUZ, Gabriel. « Méta-Étiqueteur Adaptatif : Vers une utilisation pragmatique des ressources linguistiques » in *Actes TALN'99*.

- JEANNERET, Yves. *Ecrire la science, formes et enjeux de la vulgarisation*. Presses Universitaires de France, Paris, 1994.
- JOACHIMS T., “A Probabilistic Analysis of the Rocchio Algorithm with TFIDF for Text Categorization” in ICML '97, Proceedings of the Fourteenth International Conference on Machine Learning, Morgan Kaufmann Publishers Inc., 1997, p. 143–151.
- JORDAN, Michael. “The linguistic genre of abstracts”, in Angela Della Volpe (ed.), *The Seventeenth LACUS Forum*. Linguistics Association of Canada and the United States, 507-527, 1991.
- JUFER, Marcel. « Recherche scientifique et développement technique. Quelles priorités dans les choix ? » in *Revue européenne des sciences sociales*, Tome XXXV, N° 108, Librairie Droz, Genève-Paris, 1997, pp. 197-113
- KANDO, Noriko. “Text Structure Analysis as a Tool to Make Retrieved Documents Usable” in Proceedings of the 4th International Workshop on Information Retrieval with Asian Languages, Taipei, Taiwan, Nov. 11-12, 1999, p. 126-132.
- KARLGRÉN J., CUTTING D., “Recognizing text genres with simple metrics using discriminant analysis”, Proceedings of COLING 94, Kyoto, 1994.
- KARLGRÉN, J. “Text genre recognition using discriminant analysis“ in International Conference on Computational Linguistics, 1994.
- KERBRAT-ORECCHIONI, Catherine. *L'énonciation. De la subjectivité dans le langage*. Collection Linguistique. Editions Armand Colin, Paris, 1980.
- KESSLER, Brett, NUNBERG, Geoff and SCHÜTZE, Hinrich. “Automatic Detection of Genre” in Proceedings of the 35th Annual Meeting of the Association for Computational Linguistics and the 8th Meeting of the European Chapter of the Association for Computational Linguistics, pages 32-38, Morgan Kaufmann Publishers, San Francisco CA, 1997.
- KOHONEN, T. *Self-organizing maps*. Springer-Verlag, Berlin, 1995.
- KOURILOVA, M. “Communicative characteristics of reviews of scientific papers written by non-native users of English” in *Endocrine Regulations*, vol. 32, 107 . 114, 1998.
- L'HOMME, Marie-Claude. « Sélection de termes dans un dictionnaire d'informatique : comparaison de corpus et critères lexico-sémantiques » in Actes. Euralex 2004, Lorient (France), 6 au 10 juillet 2004, pp. 583-593.
- LAKOFF, G. “Hedges : a Study in Meaning Criteria and the Logic of Fuzzy Concepts” in *Papers from the 8th Regional Meeting*, Chicago Linguistics Society, 1972: 1983-228.
- LATOUR, Bruno & WOOLGAR, Steve. *La vie de laboratoire: la production des faits scientifiques*. Editions Sage, 1979, Editions La Découverte, Paris, 1996 (pour la traduction française).
- LATOUR, Bruno. *La science en action*. Editions La découverte/texte à l'appui, Paris, 1989.
- LATOUR, Bruno. *Nous n'avons jamais été modernes. Essai d'anthropologie symétrique*. Collection l'Armillaire. Editions La découverte, Paris, 1994.
- LEBART, Ludovic. « Validité des visualisations de données textuelles » in Actes des 7emes Journées internationales d'Analyse statistique des Données Textuelles, Louvain-la-Neuve, 10-12 mars 2004, pp.708-715.

- LEBART, L., MORINEAU, A et FENELON, J.-P. *Traitement des données statistiques. Méthodes et programmes*. Dunod, Paris, 1982.
- LEBART, Ludovic, MORINEAU, Alain et PIRON, Marie. *Statistique exploratoire multidimensionnelle*. 3^e édition. Dunod, Paris, 1998.
- LEBART, Ludovic, PIRON, Marie, STEINER, Jean-François. *La Sémiométrie. Essai de statistique structurale*. Dunod, Paris, 2003.
- LEBART, Ludovic et SALEM, André. *Statistique textuelle*. Dunod, Paris, 1994.
- LECOMTE Josette et PAROUBEK Patrick (1996). *Le catégoriseur d'Eric BRILL (University of Pennsylvania), mise en œuvre de la version entraînée pour l'INaLF*, rapport interne, INaLF-CNRS, Nancy.
- LECOMTE, Josette, LUCAS, Nadine et RAJMAN, Martin. "Linguistic Issues in Grace (Evaluation of Part-of-Speech Tagging for French)" in LREC, *Linguistic Resources, Evaluation and Corpora*, Grenade, mai 1998.
- LEE Y.-B., MYAENG S. H. "Text genre classification with genre-revealing and subject-revealing features" in SIGIR '02, Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval, ACM Press, 2002, p. 145–150.
- LEKI, Iliana. *Academic writing : exploring processes and strategies*. St. Martin's press, NY, 1995
- LERAT, Pierre. *Les langues spécialisées*. PUF, Paris, 1995.
- LEWIS D. D. RINGUETTE M. "A comparison of two learning algorithms for text categorization" in Proceedings of SDAIR-94, 3rd Annual Symposium on Document Analysis and Information Retrieval, Las Vegas, US, 1994, p. 81–93.
- LOISEAU, Sylvain. "Philosophical discourse from autonomy to engagement: Deleuze commentator of Spinoza" in Fløttum et Rastier (eds.), *Academic discourse — Multidisciplinary Approaches*, Novus, Oslo, 2003, pp. 36-54.
- LOISEAU S. « Thématique et sémantique contextuelle d'un concept philosophique », in Williams G. (dir) *La linguistique de corpus*, Presse Universitaires de Rennes, Rennes, 2005, pp. 129-140.
- LOISEAU, S.. *Sémantique du discours philosophique: théorie et descriptions*. Th : ling. : Paris X-Nanterre : 2006 (à venir).
- LOISEAU, Sylvain, POUDAT, Céline et ABLALI, Driss. « Exploration contrastive de trois corpus de sciences humaines ». JADT 2006, 19-21 avril 2006, Besançon.
- LUNDQUIST, L. "Coherence in Scientific Texts", in *Connexity and Coherence. Research in Text Theory*, vol. 12, W. Heydrich et al. éds, Berlin, De Gruyter, 1989.
- MACAULAY, M. et BRICE, C. "Gentlemen prefer blondes: A Study of gender bias in example sentences" in Bucholtz, Liang, Sutton et Hines (eds.), *Cultural preferences: Proceedings of the third Berkeley women and language conference*. Berkeley Women and Language Group, 1994, pp. 449-61.
- MACAULAY, M. et BRICE, C. "Don't touch my projectile: Gender bias and stereotyping in syntactic examples" in *Language*, 73 (4), 1997, 798-825.
- MAINGUENEAU, Dominique. *Nouvelles tendances en analyse du discours*. Collection Langue, Linguistique, Communication. Editions Hachette, Paris, 1987.

- MALRIEU, Denise et RASTIER, François. « Genres et variations morphosyntaxiques » in TAL, Vol. 42, n°2/2001.
- MARCHAND, Pascal. *L'Analyse du Discours Assistée par Ordinateur*. Editions Armand Colin, Paris, 1998.
- MARIANI, J. et PAROUBEK, P. "Human Language Technologies Evaluation in the European Framework" in Proceedings of the DARPA Broadcast News Workshop, Washington, Morgan Kaufman Publishers, February 1999, pp 237-242.
- MAYAFFRE, Damon. « Rôle et place des corpus en linguistique : réflexions introductives » in *Texto!* [en ligne], décembre 2005, vol. X, n°4. Disponible sur : <http://www.revue-texto.net/Reperes/Themes/Mayaffre_Corpus.html>. (Consultée le 15/04/06).
- MEZAILLE T. 2003. 'Thématiques littéraires : enseignement des textes numériques' *Texto !* [en ligne]. Disponible sur http://www.revue-texto.net/Inedits/Mezaille_Thematiques.html (Consultée le 25/04/2005).
- MILLER, Carolyn R. "Rhetorical community: the cultural basis of genre" in Freedman et Medways (eds.), *Genre and the New Rhetoric*. Taylor and Francis, Londres, 1994.
- MISESKA TOMIC, Olga et SHUY, Roger W. *The Relation between Theoretical and Applied Linguistics*. Plenum Press, N.Y., 1987.
- MOIRAND, Sophie. *Enseigner à communiquer en langue étrangère*. Hachette éducation, Paris, 1992.
- MOIRAND, Sophie. *Une histoire de discours... une analyse des discours de la revue LE FRANCAIS DANS LE MONDE 1961-1981*. Hachette, Paris, 1988.
- MOLINIÉ, G. and CAHNÉ, P.. *Qu'est-ce que le style ?* Presses Universitaires de France, Paris, 1994.
- MOLLA, Diego et HUTCHINSON, Ben. "Intrinsic versus Extrinsic Evaluations of Parsing Systems" in Proceedings European Association for Computational Linguistics (EACL), workshop on Evaluation Initiatives in Natural Language Processing. Budapest, 14 April 2003.
- MULKAY, Michael. *Science and the sociology of knowledge*. Gregg Revivals, Hampshire, 1992.
- NEUMANN, Stella and HANSEN-SCHIRRA, Silvia. *Proceedings of the workshop on Multilingual corpora: Linguistic Requirements and Technical Perspectives*. Corpus Linguistics 2003. Lancaster, March 2003.
- NUCHEZE V. de. « Approche pragmatique-énonciative du discours de recherche (à l'usage des apprentis-chercheurs) » in Dabène & Reuter (eds) *Pratiques de l'écrit et modes d'accès au savoir dans l'enseignement supérieur*. Lidil 17. Grenoble : Université Stendhal, 1998.
- OOSTDIJK, Nelleke et DE HAAN, Pieter (eds.). *Corpus-based research into language*. Editions Rodopi, Amsterdam-Atlanta, 1994.
- PADRO, Lluís et MARQUEZ, Lluís. "On the Evaluation and Comparison of Taggers: the Effect of Noise in Testing Corpora" in *Coling-ACL*, 1998, pp. 997-1002.

- PALTRIDGE, Brian. *Genres, frames and writing in research settings*. John Benjamins B.V., 1997.
- PHILIPPE G.. « L'appareil formel de l'effacement énonciatif et la pragmatique des textes sans locuteur » in Amossy (ed.) *Pragmatique et analyse des textes* 17-34. Université de Tel-Aviv, 2002.
- PIETRI, Etienne (ed.). *Problèmes théoriques et méthodologiques de l'analyse contrastive*. Actes du Colloque 29-30-31 Octobre 1986. CRELIC, Service des Publications de la Sorbonne Nouvelle.
- PIGNON, D. « Débuts et fins des sciences ; à propos des représentations modernes des sciences » in *Revue européenne des sciences sociales*, Tome XXXV, N° 108, Librairie Droz, Genève-Paris, 1997, pp. 31-45.
- PONTILLE D. *La signature scientifique. Une sociologie pragmatique de l'attribution*. CNRS Editions, Paris, 2004.
- POUDAT Céline. "Characterization of French linguistic research papers with morphosyntactic variables" in Fløttum & Rastier (ed.) *Academic discourse — Multidisciplinary Approaches*. Novus, Oslo, 2003, pp. 77-96.
- POUDAT, Céline. Rapport final d'un contrat de bourse Marie Curie à Bergen Advanced Training Site in Multilingual Tools (BATMULT) – Octobre-décembre 2003. Disponible sur <<http://helmer.aksis.uib.no/batmult/Celine-compte-rendu.htm>>. Consultée le 15/04/06.
- POUDAT, Céline. "Evaluation des manifestations pronominales de l'auteur dans l'article de revue linguistique en français et en anglais" in *Akademisk Prosa*, n° 2, Skrifter fra KIAP, Romansk institutt – universitetet i Bergen, Bergen, 2004, pp. 69-86.
- POUDAT, Céline. « Une annotation de corpus dédiée à la caractérisation du genre de l'article scientifique » in Workshop TCAN « La construction du savoir scientifique dans la langue », Maison des Sciences Humaines – Alpes, Grenoble, 20 octobre 2004.
- POUDAT, Céline. « Recension et présentation comparative d'étiqueteurs pour le français et l'anglais » in *Texto!* [en ligne], décembre 2004, vol. IX, n°4. Disponible sur : <http://www.revue-texto.net/Corpus/Poudat_Taggers.html>. (Consultée le 15/04/06).
- POUDAT, Céline. « Genres et sous-domaines scientifiques dans un corpus d'articles de linguistique » in Journée scientifique Corpus de sciences sociales – Etablissement, numérisation, analyses sémantiques, INALCO, Paris, 08/06/05. Disponible sur <http://www.revue-texto.net/Archives/Corpus_SHS/Poudat_08juin05.ppt> (Consultée le 15/04/06).
- POUDAT, Céline. « Typologie des concepts de linguistique : évaluation et élaboration en corpus de critères discriminants » in *Corpus en lettres et sciences sociales : des documents numériques à l'interprétation*, Colloque Langages et signification, Albi, 10-14 juillet 2006 (à venir).
- POUDAT, Céline et CLEUZIQU, Guillaume. "Genre and Domain processing in an Information Retrieval perspective" in 3rd International Conference on Web Engineering, ICWE 2003, Oviedo, Espagne, Juillet 2003, Lecture Notes in Computer Science, pp. 399-402.
- POUDAT, Céline, CLEUZIQU, Guillaume et CLAVIER, Viviane. « Catégorisation de textes en domaines et genres : complémentarité des indexations » in *Document numérique* vol. 9 n°1, Hermès, Editions Lavoisier, 2005, pp. 61-76.

- POUDAT, Céline et LOISEAU, Sylvain. "Authorial presence in academic genres" in Gabriella Del Lungo & Elena Tognini Bonelli (eds.), *Strategies in academic discourse*. Coll. Studies in Corpus Linguistics, Vol. 19. John Benjamins, Amsterdam/Philadelphie, 2005, pp. 51-68.
- POUDAT, Céline et RINCK, Fanny. « Genres scientifiques et style d'auteur : des variations stylistiques de l'article de revue linguistique » in 4e Journées Internationales de la Linguistique de Corpus, Lorient, 15-17 septembre 2005 (à paraître).
- POUDAT, Céline et RINCK, Fanny. « Contrastes internes et variations stylistiques du genre de l'article scientifique en linguistique ». JADT 2006, 19-21 avril 2006, Besançon.
- PORTER M. F. "An algorithm for suffix stripping" in Program 14, 1980, p. 130-137.
- PRIAM, Rodolphe. *Méthodes de carte auto-organisatrice par mélange de lois contraintes. Application à l'exploration dans les tableaux de contingence textuels*. Th : info. : Rennes 1 : 2003 : 130 p.
- PRYTZ, Klas. "Evaluation of the Syntactic Parsing Performed by the ENGCG Parser" in The 11th Nordic Conference on Computational Linguistics, January 28-30, Copenhagen. University of Copenhagen, 1998.
- QUINLAN J. R., C4.5 : Programs for Machine Learning, Morgan Kaufmann, 1993.
- RASTIER, François. *Sémantique interprétative*. Collection Formes sémiotiques. Presses Universitaires de France, Paris, 1987.
- RASTIER, François. *Sens et textualité*. Hachette, Paris, 1989.
- RASTIER, François, CAVAZZA, Marc et ABEILLE, Anne. *Sémantique pour l'analyse, de la linguistique à l'informatique*. Masson, Paris, 1994.
- RASTIER, François. « Le problème du style pour une sémantique du texte » in Molinié, G. et Cahné, P. (dir.), *Qu'est-ce que le style ?*, PUF, Paris, 1994, pp.263-82.
- RASTIER, François (dir.). *L'analyse thématique des données textuelles. L'exemple des sentiments*. Didier Erudition, Paris, 1995.
- RASTIER, François. « Pour une sémantique des textes – questions d'épistémologie » in RASTIER, F. (ed.). *Sens et Textes*, Paris, Didier, 1996.
- RASTIER, François. *Arts et Sciences du texte*. Collection Formes sémiotiques. Presses Universitaires de France, Paris, 2001.
- RASTIER, François. « Ontologie(s) » in *Revue d'Intelligence artificielle*, 2004, vol. 18, n°1, p. 15-40.
- RASTIER, François. "Enjeux épistémologiques de la linguistique de corpus" in *Texto!* [en ligne], juin 2004. Rubrique Dits et inédits. Disponible sur : <http://www.revue-texto.net/Inedits/Rastier/Rastier_Enjeux.html>. (Consultée le 15/04/06).
- RASTIER, François. « Pour une sémantique des textes théoriques » in *Revue de sémantique et de pragmatique*, 2005, n°17, p. 151-180. Disponible sur <http://www.revue-texto.net/Inedits/Rastier/Rastier_Textes.html>. (Consultée le 15/04/06).
- RAUBER, Andreas et MÜLLER-KÖGLER Alexander. "Integrating Automatic Genre Analysis into Digital Libraries" in Fox, E.A., and Borgman, C.L. (eds.), *Proceedings of*

- the ACM/IEEE Joint Conference on Digital Libraries 2001, June 24 - 28 2001, Roanoke, VA, pp.1-10, ACM, 2001.
- RAYSON, Paul et GARSIDE, Roger. "Comparing Corpora using Frequency Profiling" *in* Proceedings of the Workshop Comparing Corpora, 7th October 2000.
- REUTER, Y. « De quelques obstacles à l'écriture de recherche » *in* Lidil, 17, Grenoble, 1998, pp. 11-24.
- RIEZLER et al. "Parsing the Wall Street Journal using a Lexical-Functional Grammar and Discriminative Estimation Techniques" *in* Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics, 2002, Philadelphia.
- RINCK, F.. "Images of scientific activity through the research article : a comparison between linguistics and literary studies" *in* Fløttum and Korsnes (eds), *Academisk Prosa*, n° 3, Department of Romance Studies, University of Bergen, 2005.
- RINCK, F.. *L'article de recherche en sciences du langage et en lettres : figure de l'auteur et identité disciplinaire du genre*. Th. : ling. : Stendhal Grenoble 3: 2006 (à paraître).
- ROUANET, Henry et LE ROUX, Brigitte. *Analyse des données multidimensionnelles, statistique en sciences humaines*. Editions Dunod, Paris, 1993.
- ROWLEY-JOLIVET, Elizabeth. "Visual discourse in scientific conference papers. A genre-based study" *in* *English for Specific Purposes*, Vol. 21 (1), 2002, pp. 19-40.
- SAMPSON, G. *English for the Computer: The Susanne Corpus and Analytic Scheme*. Clarendon Press, Oxford, 1995.
- SANJUAN E., IBEKWE-SANJUAN F. « Terminologie et classification automatique des textes » *in* Actes 6ème Journées internationales d'Analyse statistique des données textuelles (JADT' 2002). Saint-Malo, 13-15 March 2002, pp. 677-688.
- SANTORINI, B. *Bracketing Guidelines for the Penn Treebank Project. Technical report*. Department of Computer and Information Science, University of Pennsylvania, 1991.
- SANTOS, Diana & GASPERIN, Caroline. "Evaluation of parsed corpora: experiments in user-transparent and user-visible evaluation", *in* Manuel González Rodríguez & Carmen Paz Suárez Araujo (eds.), Proceedings of LREC 2002, the Third International Conference on Language Resources and Evaluation (Las Palmas de Gran Canaria, Spain, 29-31 May 2002), ELRA, 2002, pp.597-604.
- SAVILLE-TROIKE, Muriel. *The Ethnography of Communication : An Introduction*. Blackwell, Oxford, 1982.
- SCARLETT, Elizabeth. *An Evaluation of a Rule-Based Parser of English Sentences*. Thèse, Ontario, Canada, Mars 2000. Disponible sur <<http://www.site.uottawa.ca/~scarlett/thesis/toc.html>> (consultée le 15/04/06)
- SCHMID, Helmut. "Improvements in Part-Of-Speech Tagging With an Application to German" *in* EACL SIGDAT workshop, 1995, Dublin. [revised version]
- SCHMID, Helmut. "Probabilistic Part-Of-Speech Tagging Using Decision Trees" *in* International Conference on New Methods in Language Processing, 1994, Manchester, UK. [revised version]
- SEBASTIANI F. "Machine learning in automated text categorization" *in* ACM Comput. Surv., vol. 34, no 1, 2002, p. 1-47, ACM Press.

- SELDEN, Roman et WIDDOWSON, Peter. *A reader's Guide to Contemporary Literary Theory*. Third edition. Harvester Wheatsheaf, Londres, 1993.
- SERPOLLET, Noëlle et POUDAT, Céline. « Exploring the generic structure of scientific articles in a contrastive and corpus-based perspective ». *Digital Humanities*, 5-9 juillet 2006, Paris-Sorbonne. (à venir).
- SIMONIN et GRUMBACH J. « Pour une typologie des discours » in Kristeva, Milner, Ruwet (eds) *Langue, discours, société: pour Émile Benveniste*, Seuil, Paris, 1975, 85-121.
- SJOBERGH, Jonas. "Combining POS-taggers for improved accuracy on Swedish text" in *Proceeding of NoDaLiDa 2003* (Reykjavik, Iceland 2003)
- SKELTON, John R. et EDWARDS, Sarah J. L. "The function of the discussion section in academic medical writing" in *BMJ*. 2000; 320: 1269-70
- SRPOVÁ, Milena. « D'une langue à l'autre, quelles méthodologies pour la description des textes » in *Parcours linguistiques de discours spécialisés*. Colloque en Sorbonne les 23-24-25 septembre 1992. collection Sciences pour la communication. Editions Peter Lang, Berne 1994.
- STUBBS, Michael. *Text and Corpus Analysis*. Blackwell publishers, Oxford-Malden, 1996.
- SWALES, John et FEAK, Christine. *Academic writing for graduate students: Essential tasks and skills*. University of Michigan Press, Ann Arbor, 1994.
- SWALES, John et NAJJAR, Hazem. « The writing of research article introductions » in *Written Communication* 4: 175-192, 1995.
- SWALES, John. *Genre Analysis : English in Academic and research settings*. Cambridge University Press, Cambridge, 1990.
- TEUFEL, Simone et al. *EAGLES, Study of the relation between Tagsets and Taggers*. EAG-CLWG-TAGS/V. Mai 1996.
- TEI: Text Encoding Initiative. *Les Cahiers Gutenberg*. Cahier 24, juin 1996. Disponible sur <<http://www.gutenberg.eu.org/publications/cahiers/50-cahiers24.html>> (Consultée le 15/04/06)
- TODOROV, Tzvetan. *Les genres du discours*. Collection poétique. Editions du seuil, Paris, 1978.
- TODOROV, Tzvetan. *Mikhaïl Bakhtine, le principe dialogique*. Le Seuil, Paris, 1981.
- TOGNINI-BONELLI, E. *Corpus Linguistics at Work*. John Benjamin's Publishing, Amsterdam, 2001.
- TOUTANOVA, Kristina et al. "Text classification in a hierarchical mixture model for small training sets", in *Proceedings of the ACM Conference on Information and Knowledge Management (CIKM)*, 2001.
- TRIMBLE, Louis. *English for Science and Technology. A discourse approach*. Cambridge University Press. Cambridge - London - New York - New Rochelle - Melbourne - Sydney, 1985.
- TROSBORG, Anna. *Analysing professional genres*. John Benjamins Publishing Company, Amsterdam, 2000.

- VALETTE, Mathieu. "Conceptualisation and Evolution of Concepts. The example of French Linguist Gustave Guillaume" in Fløttum & Rastier (ed.) *Academic discourse — Multidisciplinary Approaches*. Novus, Oslo, 2003, pp. 55-74.
- VAN DIJK (A.), Teun. *Text and context, Explorations in the Semantics and Pragmatics of Discourse*. Longman, 1977.
- VAPNIK, V. *The nature of statistical theory*. Springer Verlag, 1995.
- VERGNE, J. et GUIGUET, E., « Regards théoriques sur le Tagging » in Actes de la Conférence sur le Traitement Automatique du Langage Naturel (TALN'98), Paris, France, juin 1998.
- VIGNAUX, Georges. *L'argumentation, essai d'une logique discursive*. Librairie Droz, Genève-Paris, 1976.
- VION, Robert. « Pour une approche relationnelle des interactions verbales et des discours » in *Langage et Société* n°87, Maison des Sciences de l'Homme, Paris, 1999.
- WILLIAMS, Geoffrey (dir.). *La linguistique de corpus*. Coll. Rivages linguistiques. Presses Universitaires de Rennes, Rennes, 2005.
- WOLTERS, Maria and KIRSTEN, Mathias. "Exploring the Use of Linguistic Features in Domain and Genre Classification" in Ninth Conference of the European Chapter of the Association for Computational Linguistics, 1999.
- ZAVREL, Jakub et DAELEMANS, Walter. "Bootstrapping a tagged corpus through combination of existing heterogeneous taggers" in Conference on Language Resources and Evaluation (LREC 2000), pages 17-20.
- ZAVREL, Jakub et DAELEMANS, Walter. "Recent Advances in Memory-Based Part-of-Speech Tagging" in *Actas del VI Simposio Internacional de Comunicacion Social*, Santiago de Cuba, pp. 590-597, 1999. ILK pub: ILK-9903.