

## Références bibliographiques

- Abi Farah A. (1988) - Reconnaissance de l'auteur d'un texte d'après les caractères utilisés, *Les Cahiers de l'Analyse des Données*, XIII, n°1, p 95-96.
- Achard P. (1993) - *La sociologie du langage*, Que-sais-je ? PUF, Paris.
- Adams L. L.(1975) - *A statistical analysis of the book of Isaiah in relation to the Isaiah problem*, Brigham Young University, Provo.
- Aitchison J., Aitken C. G. G. (1976) - Multivariate binary discrimination by the kernel method, *Biometrika*, 63, p 413-420.
- Akuto H. (Ed.) (1992) - *International Comparison of Dietary Cultures*, Nihon Keizai Shimbun, Tokyo.
- Akuto H., Lebart L. (1992) - Le repas idéal. Analyse de réponses libres en anglais, français, japonais, *Les Cahiers de l'Analyse des Données*, vol XVII, n°3, Dunod, Paris, p 327-352.
- Aluja Banet T., Lebart L. (1984) - Local and partial principal component analysis and correspondence analysis, *COMPSTAT, Proceedings in Computational Statistics*, Physica Verlag, Vienna, p 113-118.
- ASU (1992) - *La qualité de l'inform* - *Latent semantic indexing is an optimal special case of multidimensional scaling*, *Proceedings of the 15th Int. ACM-SIGIR Conf. on Res. and Dev. in Information Retrieval*, Belkin N and al. Ed., p 161-167, ACM Press, New York.
- Barthélémy J.P., Luong X. (1987) - Sur la typologie d'un arbre phylogénétique : Aspects théoriques, algorithmes et applications à l'analyse de données textuelles, *Math. Sci. Hum.* n°100, p 57-80.
- Baudelot C. (1988) - Confiance dans l'avenir et vie réussie (Réponse à un questionnaire), *Recherches économiques : études en hommage à Edmond Malinvaud*, Economica et EHESS, Paris.
- Bécue M. (1988) - Characteristic repeated segments and chains in textual data analysis, *COMPSTAT, 8th Symposium on Computational Statistics*, Physica Verlag, Vienna.



- Boeswillwald E. (1992) - L'expérience du CESP en matière de qualité des mesures d'audience, in : *La qualité de l'information dans les enquêtes*, ASU, Dunod, Paris.
- Bolasco S. (1992) - Sur différentes stratégies dans une analyse des formes textuelles : Une expérimentation à partir de données d'enquête, *Jornades Internacionals d'Analisi de Dades Textuals*, UPC, Barcelona, p 69-88.
- Bonnafous S. (1991) - *L'immigration prise aux mots. Les immigrés dans la presse au tournant des années quatre-vingt*, Kimé, Paris.
- Bonnet A. (1984) - *L'intelligence artificielle, promesses et réalités*, InterEditions, Paris.
- Boscher F. (1984) - *Le système d'enquêtes sur les conditions de vie et aspirations des Français*, Phase 5, Thème Transport, Rapport Credoc.
- Bouchaffra D., Rouault J. (1992) - Solving the morphological ambiguities using a nonstationary hidden Markov model, *AAAI Fall Symposium on Probabilistic Approach to Natural Language processing*, Cambridge, Massachusetts.
- Bouroche J.M., Curvalle B. (1974) - La recherche documentaire par voisinage, *R.A.I.R.O.*, V-1, p 65-96.
- Bradburn N., Sudman S., and Associates (1979) - *Improving Interview Method and Questionnaire Design*, Jossey Bass. San Francisco.
- Brainerd B. (1974) - *Weighing Evidence in Language and Literature. A Statistical Approach*, University of Toronto Press.
- Brunet E. (1981) - *Le vocabulaire français de 1789 à nos jours, d'après les données du Trésor de la langue française*, Slatkine-Champion, Genève-Paris.
- Burt C. (1950) - The factorial analysis of qualitative data, *British J. of Stat. Psychol.*, vol 3, n°3, p 166-185.
- Burtschy B., Lebart L. (1991) - Contiguity analysis and projection pursuit, in : *Applied Stochastic Models and Data Analysis*, R. Gutierrez and M.J.M. Valderrama, Eds, World scientific, Singapore, p 117-128.
- Callon M., Courtial J.-P., Turner W. (1991) - La méthode Leximappe, in : *Gestion de la recherche, nouveaux problèmes, nouveaux outils*. Vinck D. (Ed.), De Boeck-Wesmael, Bruxelles.
- Carré R., Dégremont J.F., Gross M., Pierrel J.M., Sabah G. (1991) - *Langage humain et machine*, Presses du CNRS, Paris.
- Celeux G., Hébrail G., Mkhadri A., Suchard M. (1991) - Reduction of a large scale and ill-conditioned statistical problem on textual data, in : *Applied Stochastic Models and Data Analysis, Proceedings of the 5th Symposium in ASMDA*, Gutierrez R. and Valderrama M.J. Eds, World Scientific, p 129-137.
- Chartron G. (1988) - *Analyse des corpus de données textuelles, sondage de flux d'Information*, Thèse. Univ. Paris 7.
- Church K.W., Hanks P. (1990) - Word association norms, mutual information and lexicography, *Computational Linguistics*, Vol 16, p 22-29.
- Cibois P. (1992) - Eclairer le vocabulaire des questions ouvertes par les questions fermées : le tableau lexical des questions, *Bull. de Method. Sociol.*, 26, p 24-54.

- Cohen M. (1950) - Sur la statistique linguistique, *Conférence de l'institut de linguistique de l'université de Paris*, Vol IX, Année 1949, Klincksieck, Paris.
- Cooper W., Gey F.C., Dabney D.P. (1992) - Probabilistic retrieval based on staged logistic regression, *Proceedings of the 15th Int. ACM-SIGIR Conf. on Res. and Dev. in Information Retrieval*, Belkin N and al., Eds, ACM Press, New York, p 198-209.
- Coulon D., Kayser D. (1986) - Informatique et langage naturel : Présentation générale des méthodes d'interprétation des textes écrits, *Techniques et sciences informatiques*, Vol.5, n°2, p 103-128.
- Cutting D. R., Karger D.R., Pedersen J.O., Tukey J.W. (1992) - Scatter / Gather : A cluster based approach to browsing large document collections. *Proceedings of the 15th Int. ACM-SIGIR Conf. on Res. and Dev., in Information Retrieval*, Belkin N. and al., Ed, ACM Press, New York, p 318-329.
- Deerwester S., Dumais S.T., Furnas G.W., Landauer T.K., Harshman R. (1990) - Indexing by latent semantic analysis, *J. of the Amer. Soc. for Information Science*, 41 (6), p 391-407.
- Demonet M., Geffroy A., Gouaze J., Lafon P., Mouillaud M., Tournier M. (1975) - *Des tracts en Mai 68. Mesures de vocabulaire et de contenu*, Armand Colin et Presses de la Fondation Nat. des Sc. Pol., Paris.
- Dendien J. (1986) - La Base de données de l'Institut National de la Langue Française, *Actes du colloque international CNRS*, Nice, juin 1985, 2 vol., Slatkine-Champion Genève, Paris.
- Desval H., Dou H. (1992) - *La veille technologique. L'information scientifique, technique et industrielle*, Dunod, Paris.
- Diday E. (1971) - La méthode des nuées dynamiques, *Revue Stat. Appl.*, vol.19, n°2, p 19-34.
- Diday E. (1992) - From data to knowledge : Probabilist objects for a symbolic data analysis, in : *Computational Statistics*, Dodge Y., Whittaker J. (Eds), Physica Verlag, Heidelberg, p 193-214.
- Efron B. (1982) - *The Jackknife, the Bootstrap and other Resampling Plans*, SIAM, Philadelphia.
- Efron B., Thisted R. (1976) - Estimating the number of unseen species : how many words did Shakespeare know ?, *Biometrika*, 63, p 435-437.
- Ellegard A. (1962) - A statistical method for determining authorship : the Junius Letters, 1769-1772, *Gothenburg Studies in English*, n°13, University of Gothenburg.
- Escoufier B. (1978) - Analyse factorielle et distances répondant au principe d'équivalence distributionnelle, *Revue de Statist. Appl.*, vol. 26, n°4, p 29-37.
- Escoufier B. [Cordier] (1965) - *Analyse des correspondances*, Thèse, Faculté des Sciences de Rennes.
- Escoufier Y. (1985) - L'Analyse des correspondances, ses propriétés, ses extensions. *Bull. of the Int. Stat. Inst.*, 4, p 28-2.

- Estoup J. B. (1916) - *Gammes sténographiques*, 4<sup>ème</sup> Edition, Paris. (cf. aussi Mandelbrot, 1961).
- Fiala P., Habert B., Lafon P., Pineira C. (1987) - Des mots aux syntagmes. Figements et variations dans la résolution générale du congrès de la CGT de 1978, *Mots*, N° 14, p 47-87.
- Fisher R. A. (1936) - The use of multiple measurements in taxonomic problems, *Annals of Eugenics*, 7, p 179-188.
- Fowler R.H., Fowler W.A.L., Wilson B.A. (1991) - Integrating query, thesaurus, and documents through a common visual representation, *Proceedings of the 14th Int. ACM Conf. on Res. and Dev. in Information Retrieval*, Bookstein A. and al., Ed, ACM Press, New York, p 142-151.
- Friedman J.H. (1989) - Regularized discriminant analysis, *Journal of the American Statistical Association*, 84, p 165-175.
- Froeschl K.A. (1992) - Semantic metadata : query processing and data aggregation, in : *Computational Statistics*, Dodge Y., Whittaker J. (Eds), Physica Verlag, Heidelberg, p 357-362.
- Fuchs W. (1952) - On the mathematical analysis of style, *Biometrika*, 39, p 122-129.
- Fuhr N., Pfeifer U. (1991) - Combining model-oriented and description-oriented approaches for probabilistic indexing, *Proceedings of the 14th Int. ACM Conf. on Res. and Dev. in Information Retrieval*, Bookstein A. and al., Ed, ACM Press, New York, p 46-56.
- Furnas G. W., Deerwester S., Dumais S.T., Landauer T.K., Harshman R. A., Streeter L.A., Lochbaum K.E. (1988) - Information retrieval using a singular value decomposition model of latent semantic structure, *Proceedings of the 14th Int. ACM Conf. on Res. and Dev. in Information Retrieval*, p 465-480.
- Geary R.C. (1954) - The contiguity ratio and statistical mapping, *The Incorporated Statistician*, Volume 5, N° 3, p 115-145.
- Geffroy A., Lafon P., Tournier M. (1974) - L'indexation minimale, Plaidoyer pour une non-lemmatisation, Colloque sur l'analyse des corpus linguistiques : "*Problèmes et méthodes de l'indexation minimale*", Strasbourg 21-23 mai 1973.
- Geisser S. (1975) - The predictive sample reuse method with applications, *J. Amer. Statist. Assoc.* 70, p 320-328.
- Gifi A. (1990) - *Non Linear Multivariate Analysis*, Wiley, Chichester.
- Gobin C., Deroubaix J. C. (1987) - Du progrès, de la réforme de l'Etat, de l'austérité. Déclarations gouvernementales en Belgique, *Mots*, n°15, p 137-170.
- Goldstein M., Dillon W. R. (1978) - *Discrete Discriminant Analysis*, Wiley, Chichester.
- Good I. J., Toulmin G.H. (1956) - The number of new species and the increase in population coverage when a sample is increased, *Biometrika*, 43, p 45-63.
- Greenacre M. (1984) - *Theory and Applications of Correspondence Analysis*, Academic Press, London.

- Greenacre M. (1993) - *Correspondence Analysis in Practice*, Academic Press, London.
- Gruaz C. (1987) - *Le mot français, cet inconnu. Précis de morphographémologie*, Publ. de l'Univ. de Rouen, n°138, Rouen.
- Guilbaud G.-Th. (1980) - Zipf et les fréquences, *Mots N° 1*, p 97-126.
- Guilhaumou J. (1986) - L'historien du discours et la lexicométrie. Etude d'une série chronologique : Le père Duchesne de Hébert, juillet 1793- mars 1794, *Histoire & Mesure*, Vol. I, n° 3-4.
- Guiraud P. (1954) - *Les caractères statistiques du vocabulaire*, P.U.F., Paris.
- Guiraud P. (1960) - *Problèmes et méthodes de la statistique linguistique*, P.U.F., Paris.
- Guttman L. (1941) - The quantification of a class of attributes: a theory and method of a scale construction, in *The prediction of personal adjustment* (P. Horst, ed.), SSCR New York, p 251 -264.
- Habbema. D. F., Hermans J., van Den Broek K. (1974) - A stepwise discriminant analysis program using density estimation, in *COMPSTAT 1974*, Bruckman G.(ed), Physica Verlag, Vienna.
- Habert B., Tournier M. (1987) - La tradition chrétienne du syndicalisme français aux prises avec le temps. Evolution comparée des résolutions confédérales (1945 - 1985), *Mots*, n°14.
- Hand D. J. (1981 and 1986) - *Discrimination and Classification*, Wiley, New-York.
- Hand D. J. (1992) - Microdata, macrodata, and metadata, *Computational Statistics*, Dodge Y., Whittaker J. (Eds), Physica Verlag, Heidelberg, p 325-340.
- Haton J. P. (1985) - Intelligence artificielle en compréhension automatique de la parole : état des recherches et comparaison avec la vision par ordinateur, *Techniques et sciences informatiques*, 4 (3), p 265-287.
- Hayashi C. (1956) - Theory and examples of quantification (II), *Proc. of the Institute of Stat. Math.* 4 (2) p 19-30.
- Hayashi C. (1987) - Statistical Study of Japanese National Character, *J. Japan Statistical Soc.*, Special Issue, p 71-95.
- Hayashi C., Suzuki T., Sasaki M. (1992) - *Data Analysis for Social Comparative research : International Perspective*. North-Holland, Amsterdam.
- Hébrail G., Marsais J. (1993) - Experiment in textual data analysis, *EDF, Centre de Recherche*, Clamart, France.
- Hébrail G., Suchard M. (1990) - Classifying documents : a discriminant analysis and an expert system work together, *COMPSTAT 90*, (Momirovic K. and Midner, eds), Physica Verlag, p 63-68.
- Herdan G. (1964) - *Quantitative Linguistics*, Londres, Butterworths.
- Hirschfeld H.D. (1935) - A Connection between correlation and contingency, *Proc. Camb. Phil. Soc.* 31, p 520-524.
- Holmes D.I. (1985) - The analysis of literary style - A Review, *J. R. Statist. Soc.*, 148, Part 4, p 328-341.

- Holmes D.I. (1992) - A Stylometric analysis of mormon scripture and related texts. *J. R. Statist.Soc.*, 155, Part 1, p 91-120.
- Jambu M. (1978) - *Classification automatique pour l'analyse des données, Tome 1 : méthodes et algorithmes*, Dunod, Paris.
- Juan S. (1986) - L'ouvert et le fermé dans la pratique du questionnaire. *R. Fr. Socio.* XXVII, Ed. du CNRS, Paris, p 301-306.
- Kasher A. (1972) - The book of Isaiah : Characterization of Authors by Morphological Data Processing, *Revue (R.E.L.O) LASLA N°3*, Liège, p 1- 62.
- Labbé D. (1990) - *Le vocabulaire de François Mitterrand*, Presses de la Fond. Nat. des Sciences Politiques, Paris.
- Labbé D. (1983) - *François Mitterrand - Essai sur le discours*, La pensée sauvage, Grenoble.
- Labbé D. (1990) - *Normes de dépouillement et procédures d'analyse des textes politiques*, CERAT, Grenoble.
- Labbé D., Thoiron P., Serant D. (Ed.) (1988) - *Etudes sur la richesse et la structure lexicales*, Slatkine-Champion, Paris-Genève.
- Lachenbruch P.A., Mickey M.R. (1968) - Estimation of error rate in discriminant analysis, *Technometrics*, 10, p 1-11.
- Lafon P. (1980) - Sur la variabilité de la fréquence des formes dans un corpus, *Mots N°1* , p 127-165.
- Lafon P. (1981) - Analyse lexicométrique et recherche des cooccurrences, *Mots N°3* , p 95-148.
- Lafon P. (1981) - *Dépouillements et statistiques en lexicométrie*, Slatkine-Champion, 1984, Paris.
- Lafon P., Salem A. (1983) - L'Inventaire des segments répétés d'un texte, *Mots N°6*, p 161-177.
- Lafon P., Salem A., Tournier M. (1985) - Lexicométrie et associations syntagmatiques (Analyse des segments répétés et des cooccurrences appliquée à un corpus de textes syndicaux). *Colloque de l'ALLC*, Metz -1983, Slatkine-Champion, Genève, Paris, p 59-72.
- Lazarsfeld P.E. (1944) - The controversy over detailed interviews - an offer for negotiation, *Public Opinion Quat.* n°8, p 38-60.
- Lebart L. (1969) - L'Analyse statistique de la contiguïté, *Publications de l'ISUP*, XVIII- p 81 - 112.
- Lebart L., Houzel van Effenterre Y. (1980) - Le système d'enquête sur les aspirations des français, une brève présentation, *Consommation* n°1, 1980, Dunod, p 3-25.
- Lebart L. (1987) - Conditions de vie et aspirations des français, évolution et structure des opinions de 1978 à 1986, *Futuribles* , sept 1987, p 25-56.
- Lebart L. (1982a) - Exploratory analysis of large sparse matrices, with application to textual data, *COMPSTAT*, Physica Verlag, p 67-76.

- Lebart L. (1982b) - L'Analyse statistique des réponses libres dans les enquêtes socio-économiques, *Consommation*, n°1, Dunod, p 39-62.
- Lebart L. (1992a) - Discrimination through the regularized nearest cluster method, *in : Computational Statistics*, (Y. Dodge, J. Whittaker, Eds) Physica Verlag, Heidelberg, p 103-118.
- Lebart L. (1992b) - Assessing and comparing patterns in multivariate analysis, *Second Japanese French Seminar on Data Science*, in *Data Science and application*, Hayashi et al. Eds, HBJ, Tokyo, Japan..
- Lebart L., Memmi D. (1984) - Analisi dei Dati Testuali : Applicazione al Discorso Politico. Atti delle XXXII Riunione Scient. (Soc. Ital. Statis.), p 27-41.
- Lebart L., Morineau A., Bécue M. (1989) - *SPAD.T, Système Portable pour l'Analyse des Données Textuelles, Manuel de Référence*, CISIA , Paris.
- Lebart L., Morineau A., Warwick K. (1984) *Multivariate Descriptive Statistical Analysis*, Wiley, New York.
- Lebart L., Salem A. (1988) - *Analyse statistique des données textuelles*, Dunod, Paris.
- Lebart L., Salem A., Berry E. (1991) - Recent development in the statistical processing of textual data, *Applied Stoch. Model and Data Analysis*, 7, p 47-62.
- Lelu A., Rozenblatt D. (1986) - Représentation et parcours d'un espace documentaire. Analyse des données, réseaux neuronaux et banques d'images, *Les Cahiers de l'Analyse des Données*, Vol XI - n° 4, p 453-470.
- Lelu A. (1991) - From data analysis to neural networks : new prospects for efficient browsing through databases, *Journal of Information Science*, vol. 17, p 1-12.
- Lewis D. D., Croft W. B. (1990) - Term clustering of syntactic phrases, *Proceedings of the 13th Int. ACM Conf. on Res. and Dev. in Information Retrieval*, Vidick J.L., Ed, A.C.M.Press, New York, p 385-395.
- Lewis D.D. (1992) - An evaluation of phrasal and clustered representation on a text categorization task, *Proceedings of the 15th Int. ACM-SIGIR Conf. on Res. and Dev. in Information Retrieval*, Belkin N. and al., Eds, ACM Press, New York, p 37-50.
- Luong X. (Ed.) (1989) - *Analyse arborée des données textuelles*, INALF, CUMFID/CNRS, Nice.
- Mahalanobis P.C. (1936) - On the generalized distance in statistics, *Proc. Nat. Inst. Sci. India*, 12, p 49-55.
- Mandelbrot B. (1961) - On the theory of words frequency and on related Markovian models of discourses, *Structure of Language and its Mathematical Aspects*, Am. Math. Soc., Providence, R.I., p 190-219.
- Mandelbrot B. (1968) - Les constantes chiffrées du discours, *Le Langage*, Encyclopédie de la Pléiade, vol XXV, Gallimard, Paris.
- Margulis E.L. (1992) - N-Poisson document modelling, *Proceedings of the 15th Int. ACM-SIGIR Conf. on Res. and Dev. in Information Retrieval*, Belkin N and al., Ed, ACM Press, New York, p 177-189.



- McKevitt P., Partridge D., Wilks Y. (1992) - Approaches to natural language discourse processing, *Artificial Intelligence Review*, 6, p 333-364.
- McKinnon A., Webster R. (1971) - A method of "author" identification, *The Comput. in Liter. and Linguist. Res.*, R.A. Wisbey, Ed., Cambridge Univ. Press.
- McLachlan G.J. (1992) - *Discriminant Analysis and Statistical Pattern Recognition*, Wiley, New-York.
- Menard N. (1983) - *Mesure de la richesse lexicale, théorie et vérifications expérimentales*, Slatkine-Champion, Paris.
- Michelet B. (1988) - *La logique d'association*, Thèse, Université Paris 7.
- Moran P.A.P. (1954) Notes on continuous stochastic phenomena, *Biometrika*, 37, p 17-23.
- Morton A.Q. (1963) - The authorship of the Pauline corpus, *The New Testament and Contemporary Perspective*, W. Barclay, ed., Oxford.
- Mosteller F., Wallace D. (1964) - *Inference and disputed Authorship : The Federalists*. Addison-Wesley, Reading, Mass.
- Mosteller F., Wallace D.L. (1984) - *Applied Bayesian and Classical Inference, the Case of the Federalist Papers*, Springer Verlag, New York.
- Muller C. (1964) - *Essai de statistique lexicale : L'illusion comique de P. Corneille*, Klincksieck, Paris.
- Muller C. (1968) - *Initiation à la statistique linguistique*, Larousse, Paris.
- Muller C. (1977) - *Principes et méthodes de statistique lexicale*, Hachette, Paris.
- Muller C. (1967) - *Etude de statistique lexicale. Le vocabulaire du théâtre de Pierre Corneille*, Paris, Larousse.
- Nishisato S. (1980) - *Analysis of Categorical Data. Dual Scaling and its Application*, Univ. of Toronto Press.
- Palermo D., Jenkins J. (1964) - *Word Association Norm*, University of Minnesota Press, Minneapolis.
- Pêcheux M. (1969) - *Analyse automatique du discours*, Dunod, Paris.
- Peschanski D. (1988) - *Et pourtant, ils tournent. Vocabulaire et stratégie du PCF (1934 - 1936)*, Klincksieck, Paris.
- Petruszewycz M. (1973) - L'histoire de la loi d'Estoup-Zipf, *Math. Sciences Hum.*, n°44.
- Pitrat J. (1985) - *Textes, ordinateurs, et compréhension*, Eyrolles, Paris.
- Plante P. (1985) - La structure des données et des algorithmes en Dérédec, *Revue Québécoise de Linguistique*, 14:2.
- Radday Y. T. (1974) - "And " in Isaiah, *Revue (R.E.L.O) LASLA N°2*, Liège, p 25-41.
- Rao C.R. (1989) - *Statistics and Truth*, International Cooperative Publishing House, Fairland, USA.
- Reinert M. (1983) - Une méthode de classification descendante hiérarchique : Application à l'analyse lexicale par contexte, *Les Cahiers de l'Analyse des Données*, 3, Dunod, p 187-198.

- Reinert M. (1986) - Un Logiciel d'analyse lexicale, *Les Cahiers de l'Analyse des Données*, 4, Dunod, p 471-484.
- Reinert M. (1990) - Alceste, Une méthodologie d'analyse des données textuelles et une Application : Aurélia de Gérard de Nerval, *Bull. de Method. Sociol.* n°26, p 24-54.
- Romeu L. (1992) - *Approche du discours éditorial de Ya et Arriba (1939 - 1945)*, Thèse Paris 3.
- Rugg D. (1941) - Experiments in wording questions, *Public Opinion Quarterly*, 5, p 91-92.
- Salem A. (1979) - *Contribution à une méthodologie de la validation en analyse des données textuelles*, Thèse, Université Paris 6.
- Salem A. (1982) - Analyse factorielle et lexicométrie. Synthèse de quelques expériences, *Mots N°4*, p 147-168.
- Salem A. (1984) - La typologie des segments répétés dans un corpus, fondée sur l'analyse d'un tableau croisant mots et textes, *Les Cahiers de l'Analyse des Données*, Vol IX, n° 4, p 489-500.
- Salem A. (1986) - Segments répétés et analyse statistique des données textuelles, Etude quantitative à propos du père Duchesne de Hébert, *Histoire & Mesure*, Vol. I- n° 2, Paris, Ed. du CNRS.
- Salem A. (1987) - *Pratique des segments répétés, Essai de statistique textuelle*, Klincksieck, Paris.
- Salem A. (1993) - *Méthodes de la statistique textuelle*, Thèse d'Etat, Université Sorbonne Nouvelle (Paris 3).
- Salton G. (1988) - *Automatic Text Processing : the Transformation, Analysis and Retrieval of Information by Computer*, Addison-Wesley, New York.
- Salton G., Mc Gill M.J. (1983) - *Introduction to Modern Information Retrieval*, International Student Edition.
- Sasaki M., Suzuki T. (1989) - New directions in the study of general social attitudes : trends and cross-national perspectives, *Behaviormetrika*, 26, p 9-30.
- Saussure F. de (1915) - *Cours de linguistique générale*, Payot, Genève (rééd. 1986).
- Schank, R. C., (Ed.) - *Conceptual Information Processing*, North-Holland, Amsterdam, 1975.
- Schuman H. (1966) - The random probe : a technique for evaluating the validity of closed questions, *Amer. Socio. Rev.* n°21, p 218-222.
- Schuman H., Presser F. (1981) - *Question and Answers in Attitude Surveys*, Academic Press, New York.
- Sekhroui M. (1981) - *La saisie des textes et le traitement des mots : Problèmes posés, essai de solution*, Mémoire, Ecole des hautes études en sciences sociales, Paris.
- Shannon C.E. (1948) - A mathematical theory of communication, *B.S.T.J.*, n° 27.
- Simpson E.H. (1949) - Measurement of diversity, *Nature*, 163, p 688.

- Smith M. W. A. (1983) - Recent experience and new developments of methods for the determination of authorship, *Ass. for Lit. and Linguist. Comput. Bull.*, 11, p 73-82.
- Somers H. H. (1966) - Statistical methods in literary analysis, *The Computer and Literary Style*, (J. Leed, Eds), Kent State University Press, Kent, Ohio.
- Spevack M. (1968) - *A Complete and Systematic Concordance of the Work of Shakespeare*, George Olms, Hildesheim.
- Stone C.J. (1974) - Cross-validatory choice and assessment of statistical predictions. *J.R. Statist. Soc., B* 36, p 111-147.
- Sudman S., Bradburn N. (1974) - *Response Effects in Survey*, Aldine, Chicago.
- Suzuki T. (1989) - Cultural link analysis : its application to social attitudes - A study among five nations, *Bull. of the Int. Stat. Inst.*, 47<sup>o</sup> Session, vol.1, p 363-379.
- Tabard (1974) - *Besoins et aspirations des familles et des jeunes*, Coll. Etudes CAF, n<sup>o</sup> 16, CNAF, Paris.
- Tabard N. (1975) - Refus et approbations systématiques dans les enquêtes par sondage, *Consommation*, n<sup>o</sup>4, Dunod, p 59-76.
- Tenenhaus M., Young F. W. (1985) - An analysis and synthesis of multiple correspondence analysis, optimal scaling, dual scaling, homogeneity analysis and other methods for quantifying categorical multivariate data, *Psychometrika*, vol 50, p 91-119.
- Thisted R., Efron B. (1987) - Did Shakespeare write a newly discovered poem?, *Biometrika*, 74, p 445-455.
- Tournier M. (1985a) - Sur quoi pouvons-nous compter ? Hommage à Hélène Nais, *Verbum*.
- Tournier M. (1985b) - Texte propagandiste et cooccurrences. Hypothèses et méthodes pour l'étude de la sloganisation, *Mots N<sup>o</sup>11*, p 155-187.
- Tournier M. (1980) - D'ou viennent les fréquences de vocabulaire?, *Mots N<sup>o</sup>1*, p 189-212.
- Van Rijckevorsel J. (1987) - *The application of fuzzy coding and horseshoes in multiple correspondances analysis*, DSWO Press, Leyde.
- Van Rijsbergen C.J. (1980) - *Information Retrieval*, 2nd Ed., Butterworths, London.
- Von Neumann J. (1941) - Distribution of the ratio of the mean square successive difference to the variance, *Annals of Math. Stat.*, vol.12.
- Warnesson I., Parisot P., Bedecarrax C., Huot C. (1993) - Traitements linguistiques et analyse des données pour une exploitation systématique des banques de données, *Revue Française de bibliométrie*, i 21.
- Weber R.P. (1985) - *Basic Content Analysis*, Sage, Beverly Hills.
- Weil G.E., Salem A., Serfaty M. (1976) - Le livre d'Isaïe et l'analyse critique des sources textuelles, *Revue (R.E.L.O) LASLA*, N<sup>o</sup>2, Liège.

- Wilks Y., Fass D., Guo C., McDonald J. E., Plate T., Slator B. M. (1991) - Providing machine tractable dictionary tools, *Machine translation*, p 99-154.
- Wold S. (1976) - Pattern recognition by means of disjoint principal component models, *Pattern Recognition*, 8, p 127-139.
- Yule G.U. (1944) - *The Statistical Study of Literary Vocabulary*, Cambridge University Press, Reprinted in 1968 by Archon Books, Hamden, Connecticut.
- Zipf G. K. (1935) - *The Psychobiology of Language, an Introduction to Dynamic Philology*, Boston, Houghton-Mifflin.

## Bibliographie complémentaire

On trouvera ci-dessous une liste (non-exhaustive) d'*ouvrages généraux* ou de *thèses*, pour la plupart en langue française, qui complètent les références précédente du point de vue des domaines de recherche abordés et des domaines connexes (Statistique et analyse des données, linguistique et analyse de discours, méthodes d'enquêtes, applications).

- Andreewsky A., Fluhr C. (1973) - *Apprentissage, analyse automatique du langage, application à la documentation*, Documents de linguistique quantitative, n° 21, Dunod, Paris.
- Antoine J. (1982) - *Le Sondage, outil du marketing*, Dunod, Paris.
- Balpe J. P. (1990) - *Hyperdocuments, hypertextes, hypermedias*, Eyrolles, Paris.
- Bar-Hillel Y. (1964) - *Language and Information*, Addison-Wesley,
- Baudelot C., Establet R. (1989) - *Le niveau monte*, Seuil, Paris.
- Bourdieu P. (1982) - *Ce que parler veut dire, l'économie des échanges linguistiques*, Fayard, Paris.
- Bouroche J.M., Saporta G. (1980) - *L'analyse des données*, coll."Que sais-je", n°1854, PUF, Paris .
- Bourques G., Duchastel J. (1988) - *Restons traditionnels et progressifs, pour une nouvelle analyse du discours politique*, Boréal, Montréal.
- Brian E. (1984) - *Analyse des données lexicométriques*, Rapport Credoc / D.G.T.
- Caillez F., Pagès J.P. (1976) - *Introduction à l'analyse des données*, S.M.A.S.H., Paris.
- Cibois P. (1984) - *Analyse des données en sociologie*, P.U.F. Paris.
- Cliff A.D. and Ord J.K. (1981) - *Spatial Processes : Models and Applications*, Pion, London.

- Daoust F. (1990) - *SATO: Système de base d'analyse de textes par ordinateurs*, ATO-UQAM, Montréal.
- Deroo M., Dussaix A. M. (1980) - *Pratique et analyse des enquêtes par sondage*, P.U.F., Paris.
- Dodge Y. (1993) - *Statistique. Dictionnaire encyclopédique*, Dunod, Paris.
- Ducrot O., Todorov T. (1972) - *Dictionnaire encyclopédique des sciences du langage*, Ed. du Seuil, Paris.
- Eissfeldt O. (1965) - *The Old Testament. An introduction*, transl. by P.R. Ackroyd, Oxford.
- Fénelon J.P. (1981) - *Qu'est-ce-que l'analyse des données?*, Lefonen, Paris.
- Fuchs C. (1993) - *Linguistique et traitements automatiques des langues*, Hachette, Paris.
- Gardin J-C. et coll. (1981) - *La logique du plausible : essai d'epistémologie pratique*, Ed. de la Maison des sciences de l'homme, Paris.
- Gardin J-C. (1971) - *Les analyses de discours*, Delachaux & Niestlé, Neuchâtel.
- Georgel A. (1992) - *Classification statistique et réseaux de neurones formels pour la représentation des banques de données documentaires*, Thèse - Université Paris-7.
- Ghiglione R. (1986) - *L'homme communicant*, Armand Colin, Paris.
- Ghiglione R., Matalon B. (1991) - *Les enquêtes sociologiques*, Armand Colin, Paris.
- Grangé D., Lebart L. (Ed.) (1993) - *Traitements statistiques des enquêtes*, Dunod, Paris.
- Groves R.M. (1989) - *Survey Errors and Survey Costs*, J. Wiley, New-York.
- Lafon P., Lefevre J., Salem A., Tournier M. (1985) - *Le Machinal, Principes d'enregistrement informatique des textes*, Publications de l'INaLF, Klincksieck, Paris.
- Lebart L., Morineau A., Fénelon J.P. (1979) - *Traitement des données statistiques*. Dunod, Paris.
- Lebart L., Morineau A., Tabard N. (1977) - *Technique de la description statistique*, Dunod, Paris.
- Lelu A. (1993) - *Modèles neuronaux pour l'analyse documentaire et textuelle*, Thèse, Univ. Paris 6.
- Maingueneau D. (1976) - *Initiation aux méthodes de l'analyse du discours*, Hachette, Paris.
- Marchand P. (1993) - *Engagement politique et rationalisation - Analyse psychosociale du discours militant*, Thèse de l'Université de Toulouse II.
- Meynaud H., Duclos D. (1985) - *Les Sondages d'opinions*, La Découverte, Paris.
- Muller P., (1991) - *Vocabulaire et rhétorique dans Etudes socialistes de Jean Jaurès*, Thèse, Paris 3.

- 
- Prost A. (1974) - *Vocabulaire des proclamations électorales de 1881,1885 et 1889*, PUF, Paris.
- Rey A. (1970) - *La Lexicologie*, Klincksieck, Paris.
- Ripley B.D. (1981) - *Spatial Statistics*, Wiley, New York.
- Robin R. (1973) - *Histoire et Linguistique*, Armand Colin, Paris.
- Rouanet H., Le Roux B. (1993) - *Analyse des données multidimensionnelles*, Dunod, Paris.
- Rouault J. (1986) - *Linguistique automatique : Applications documentaires*, Berne P. Lang.
- Roux M. (1985) - *Algorithmes de classification*, Masson, Paris.
- Saporta G. (1990) - *Probabilité, analyse des données et statistique*, Technip, Paris.
- Sarrazin M. (1992) - *Le débat parlementaire sur la laïcité*, Thèse, Paris 3.
- Searle, J.R. (1979) - *Sens et expression : étude de la théorie des actes de langages*. Editions de Minuit, Paris.
- Tournier M. (1975) - *Un Vocabulaire ouvrier en 1848, essai de lexicométrie*, Thèse d'état, multigr., ENS de Saint-Cloud.
- Volle M. (1980) - *Analyse des données*, Economica, Paris.
- Wagner R.-L. (1970) - *Les vocabulaires français*, tomes 1 et 2, Didier, Paris.
- Wisbey R. A. (Ed) (1971) - *The computer in Literary and Linguistic Research*, Cambridge Univ. Press, Cambridge.