

André Clas
Université de Montréal

Philippe Thoiron
Université Lumière Lyon-2

Henri Béjoint
Université Lumière Lyon-2

LEXICOMATIQUE
ET
DICTIONNAIRIQUES

Actes du Colloque de Lyon
1995

Ouvrage publié
avec l'aide
de
Rank Xerox
Research Centre
Meylan - France



AS

actualité scientifique

**LEXICOMATIQUE
ET
DICTIONNAIRIQUES**

ISBN 2-92 00 21 70 2

*Tous droits de reproduction, de traduction
et d'adaptation réservés © 1996*

FMA

Bibliothèque nationale du Québec
Bibliothèque nationale du Canada
Bibliothèque nationale de France
Imprimé au Liban

LEXICOMATIQUE ET DICTIONNAIRIQUES

IV^{es} Journées scientifiques du réseau thématique
«Lexicologie, Terminologie, Traduction»
Lyon, France, 28, 29, 30 septembre 1995

Sous la direction de :

André CLAS, Université de Montréal, Canada
Philippe THOIRON, Université Lumière Lyon-2
Henri BÉJOINT, Université Lumière Lyon-2

1996

FMA
Beyrouth

AUPELF • UREF
B.P 400, succ. Côte-des-Neiges
Montréal (Québec) Canada
H3S 2S5

Avant-propos

La diffusion de l'information scientifique et technique est un facteur essentiel du développement. Aussi dès 1988, l'Agence francophone pour l'enseignement supérieur et la recherche (AUPELF-UREF), mandatée par les Sommets francophones pour produire et diffuser revues et livres scientifiques, a créé la collection Universités francophones.

Lieu d'expression de la communauté scientifique de langue française, Universités francophones vise à instaurer une collaboration entre enseignants et chercheurs francophones en publiant des ouvrages, coédités avec des éditeurs francophones, et largement diffusés dans les pays du Sud, grâce à une politique tarifaire préférentielle.

Quatre séries composent la collection :

– Les manuels : cette série didactique est le cœur de la collection. Elle s'adresse à un public de deuxième et troisième cycles universitaires et vise à constituer une bibliothèque de référence couvrant les principales disciplines enseignées à l'université.

– Actualité scientifique : dans cette série, à laquelle appartient le présent ouvrage, sont publiés les actes des journées scientifiques organisées par les réseaux thématiques de recherche de l'UREF. A ce jour, 16 réseaux thématiques, rassemblant plusieurs milliers de chercheurs de toute la francophonie, sont constitués au sein de l'UREF.

– Prospectives francophones : s'inscrivent dans cette série des ouvrages de réflexion donnant l'éclairage de la francophonie sur les grandes questions contemporaines.

– Savoir plus Université : cette nouvelle série se compose d'ouvrages de synthèse qui font un point précis sur des sujets scientifiques d'actualité.

Notre collection, en proposant une approche plurielle et singulière de la science, adaptée aux réalités multiples de la francophonie, contribue efficacement à promouvoir la recherche dans l'espace francophone et le plurilinguisme dans la recherche internationale.

Professeur Michel Guillou
Directeur général de l'AUPELF
Recteur de l'UREF

Sommaire

Liste des auteurs	XI
Membres du Comité de réseau «LTT»	XIV
Préface André Clas (Coordonnateur du Réseau LTT)	XV
Enseignement du lexique assisté par ordinateur Jurij D. Apresjan (Académie des sciences, Moscou, Russie)	1
L'hypertexte <i>Hyperbase</i> Étienne Brunet (INaLF, UPR 68 (CNRS), Nice, France)	11
Réseaux sémantiques et dictionnaires bilingues électroniques Thierry Fontenelle (Université de Liège, Belgique)	31
Une maquette de base lexicale multilingue à pivot lexical (« acceptions multilingues ») : PARAX Étienne Blanc (GETA-CLIPS, Institut IMAG, Grenoble, France)	43
Élaboration d'un dictionnaire informatisé pour le traitement automatique de la langue Lorne H. Bouchard et Louise Emirkanian (Université du Québec à Montréal, Canada)	59
Génération de dictionnaires-machines multilingues pour la traduction automatique de diagnostics médicaux Guy Deville et Emmanuel Herbigniaux (École des Langues Vivantes, Faculté universitaire de Namur, Belgique)	77
IDAREX : description formelle des expressions à mots multiples en français et en allemand dans le cadre de la technologie des états finis Frédérique Segond et Elisabeth Breidt (Rank Xerox Research Centre, Meylan, France et Université de Tübingen, Allemagne)	93

Lexicographie bilingue informatisée au quotidien : témoignage du rédacteur face à l'écran Thomas Szende et Dominique Radanyi (INALCO, Paris et CIEH, Université de Paris III, France)	105
Orientation de combinants dans les langues de spécialité : comparaison entre l'anglais et le français Patricia Thomas et Frank Knowles (Terminologie indépendante, Cranleigh et Aston University, Birmingham, Grande-Bretagne)	115
ACABIT : une maquette d'aide à la construction automatique de banques terminologiques Béatrice Daille (Université de Nantes, IRIN, Nantes, France)	123
Conception et exploitation d'un logiciel d'extraction de termes : problèmes théoriques et méthodologiques Didier Bourigault (Centre d'Analyse et de Mathématiques Sociales (Unité Mixte EHESS-CNRS-Paris Sorbonne) et EDF, Direction des Études et Recherches, Clamart, France)	137
Amélioration automatique incrémentale de dictionnaires bilingues utilisant un corpus monolingue Kumiko Tanaka et Violaine Prince (Université de Tokyo, Japon et LIMSI-CNRS, Paris, France)	147
Conception d'un dictionnaire terminologique et phraséologique trilingue anglais/français-arabe dans le domaine de l'optique Xavier Lelubre (Université Lumière Lyon-2, France)	163
Génération automatique de néologismes arabes à partir des règles de formation de mots Hussein Habaili et Mohamed Ben Ahmed (Laboratoire de Recherche en Informatique Arabisée et Documentique Intégrée, (RIADI), Tunis, Tunisie)	173
Lexicographie berbère. Construction des formes de mot et classification des entrées lexicales Miloud Taifi (Université de Fès, Maroc)	189
Informatisation du <i>Dictionnaire explicatif et combinatoire</i> : le projet Nadia-DEC Gilles Sérasset (GETA-CLIPS-IMAG (UJF & CNRS), Grenoble, France)	205
La formalisation des collocations pour le traitement automatique du langage naturel : le modèle des fonctions lexicales syntagmatiques Agnès Tutin (URA SILEX, Université de Lille III, Villeneuve d'Ascq, France)	217
Description lexicographique des collocatifs dans un <i>Dictionnaire explicatif et combinatoire</i> : articles de dictionnaire autonomes ? Margarita Alonso Ramos et Suzanne Mantha (Universidade da Coruña, Espagne et Université de Montréal, Canada)	233

Vers un nouvel outil interactif d'aide à la conception de dictionnaires électroniques spécialisés Christophe Jouis et Widad Mustafa-Elhadi (UFR Information, Documentation, Information Scientifique et Technique, Université Charles De Gaulle Lille III et Centre d'Analyse et de Mathématiques Sociales, Unité Mixte CNRS-EHESS, Université Paris-Sorbonne, France)	255
La construction de dictionnaires à partir de l'analyse informatisée de corpus bruts : un outil pour le langagier Sylvain Delisle (Université du Québec à Trois-Rivières, Canada)	267
Réseau notionnel, intelligence artificielle et équivalence en terminologie multilingue : essai de modélisation Marc Van Campenhoudt (Centre de recherche TERMISTI, Institut supérieur de traducteurs et interprètes, Bruxelles, Belgique)	281
Les mots-clés métalinguistiques comme outil d'interrogation structurante des dictionnaires anciens Russon Wooldridge et Isabelle Leroy-Turcan (Université de Toronto, Canada et Université Jean Moulin Lyon-3, France)	307
Représentation de la polysémie dans un dictionnaire électronique Michel Mathieu-Colas (LLI, Université Paris XIII-CNRS-INaLF, Villetaneuse, France)	317
Une base de données lexicale multilingue interactive Catherine Walther et Éric Wehrli (LATL, Université de Genève, Suisse)	327
Acquisition semi-automatique du lexique Évelyne Viegas et Sergei Nirenburg (Computing Research Laboratory, New Mexico State University, Las Cruces, États-Unis)	337
Pistes de description sémantique : le cas de Biolex, dictionnaire des bio-industries François Gaudin et Myriam Bouveret (URA CNRS 1164, Université de Rouen et Praxiling, Université de Montpellier, France)	349
Le lexique génératif : une alternative au traitement de la polysémie Pierrette Bouillon (ISSCO, Université de Genève, Suisse)	359
Quand l'informatique tutoie le dictionnaire des difficultés de la langue française... Daniel Blampain (Institut supérieur des traducteurs et interprètes, Bruxelles, Belgique)	371
Choix de grammaire et organisation du lexique Philippe Barbaud (Université du Québec à Montréal, Canada)	379

Outil d'intégration de bases de connaissances lexicales aux analyseurs syntaxiques	
Philippe Blache et Mireille Delpui (2LC-CNRS, Sophia-Antipolis, France)	397
Un réseau lexico-sémantique des verbes construit à partir du dictionnaire pour le traitement automatique du français	
Karim Chibout et Nicolas Masson (Groupe Langage et Cognition, LIMSI-CNRS, Orsay, France)	405
Références	421

Liste des auteurs

- Alonso Ramos, Margarita**, Département de linguistique générale, Université de La Corunha, Campus da Zapateira s/n, 15071 La Corunha, Espagne
- Apresjan, Jurij D.**, Russie 101 447, Moscou YSP 4, rue Ermolova 19, IPPI RAN, Laboratoire 15
- Barbaud, Philippe**, Département de linguistique, UQAM C.P. 8888, succ. Centre-ville, Montréal (Québec), H3C 3P8, Canada
- Ben Ahmed, Mohamed**, Laboratoire de recherche en informatique arabisée et documentaire intégrée (RIADI), Tunis, Tunisie
- Blache, Philippe**, 2LC CNRS, 1361 route des Lucioles, Sophia Antipolis, 06560 Valbonne, France
- Blampain, Daniel**, Département de linguistique française, Groupe de recherche TERMISTI, Institut supérieur de traducteurs et interprètes, 34 rue Joseph Hazard, 1180 Bruxelles, Belgique
- Blanc, Étienne**, GETA CLIPS, Institut IMAG, Université Joseph Fourier Grenoble I, B.P. 53, 150 rue de la Chimie, 38041 Grenoble Cedex 9, France
- Bouchard, Lorne**, Département d'informatique, Université du Québec à Montréal, C.P. 8888, succ. Centre-ville, Montréal (Québec), H3C 3P8, Canada
- Bouillon, Pierrette**, ISSCO (Institut pour les études sémantiques et cognitives), Université de Genève, 54 route des Acacias, 1227 Genève, Suisse
- Bourigault, Didier**, Centre d'analyse et de mathématiques sociales, Unité mixte EHESS CNRS Paris Sorbonne, EDF Direction des études et recherches, 1 avenue du Général de Gaulle, 92141 Clamart Cedex, France
- Bouveret, Myriam**, Praxiling, Université de Montpellier, France
- Breidt, Elisabeth**, Seminar für Sprachwissenschaft, Universität Tübingen, Wilhelmstr. 113, D 72074 Tübingen, Allemagne

- Brunet, Étienne**, Institut National de la langue française, UPR 6861, UFR Lettres, Arts et Sciences Humaines, 98 bd Herriot, B.P. 209, 06204 Nice Cedex 3, France
- Campenhoudt, Marc Van**, Département de linguistique française, Centre de recherche TERMISTI, Institut Supérieur de traducteurs et interprètes, 34 rue Joseph Hazard, 1180 Bruxelles, Belgique
- Chibout, Karim**, Groupe Langage et Cognition, LIMSI CNRS, B.P. 133, 91403 Orsay Cedex, France
- Daille, Béatrice**, IRIN, Université de Nantes, 2 rue de la Houssinière, 44072 Nantes Cedex 03, France
- Delisle, Sylvain**, Département de mathématiques et d'informatique, Université du Québec à Trois-Rivières, 3351 bd des Forges, Trois Rivières (Québec), G9A 5H7, Canada
- Delpui, Mireille**, 2LC CNRS, 1361 route des Lucioles, Sophia Antipolis, 06560 Valbonne, France
- Deville, Guy**, École de Langues vivantes, Facultés universitaires de Namur, 61 rue de Bruxelles, B 5000 Namur, Belgique
- Emirkanian, Louisette**, Département de linguistique, Université du Québec à Montréal, C.P. 8888, succ. Centre-ville, Montréal (Québec), H3C 3P8, Canada
- Fontenelle, Thierry**, Département d'anglais, Université de Liège, 3 Place Cockerill, B 4000 Liège, Belgique
- Gaudin, François**, Université de Rouen, CNRS URA 1164 IRED, 7 rue Thomas Becket, 76130 Mont Saint Aignan, France
- Habaili, Hussein**, 2 rue du 3 septembre 1934, Tunis 1007, Tunisie
- Herbigniaux, Emmanuel**, École de Langues vivantes, Facultés universitaires de Namur, 61 rue de Bruxelles, B 5000 Namur, Belgique
- Jouis, Christophe**, UFR DIST/CREDO, Université Charles de Gaulle Lille III, B.P. 149, 59653 Ville d'Ascq, France
- Knowles, Frank**, Language Studies Unit, Aston University, Birmingham, B4 7ET, Royaume-Uni
- Lelubre, Xavier**, 5 rue Nicolaï, 69007 Lyon, France
- Leroy-Turcan, Isabelle**, Université Jean Moulin Lyon 3, 18 rue Chevreul, 69007 Lyon, France
- Mantha, Suzanne**, Université de Montréal, Département de linguistique, C.P. 6128 succ. Centre-ville, Montréal (Québec), H3C 3J7, Canada

- Masson, Nicolas**, Groupe Langage et Cognition, LIMSI CNRS, B.P. 133, 91403 Orsay Cedex, France
- Mathieu-Colas, Michel**, Laboratoire de linguistique informatique, Université Paris XIII CNRS INaLF, Villetaneuse, France
- Mustafa-Elhadi, Widad**, UFR DIST/CREDO, Université Charles de Gaulle Lille III, B.P. 149, 59653 Ville d'Ascq, France
- Nirenburg, Sergei**, Computing Research Laboratory, New Mexico State University, Box 30001, Las Cruces, New Mexico 88003 8001, États-Unis
- Prince, Violaine**, LIMSI CNRS, B.P. 133, 91403 Orsay Cedex, France
- Radanyi, Dominique**, CIEH, Université de Paris III, France
- Segond, Frédérique**, Rank Xerox Research Centre, 6 chemin de Maupertuis, 38240 Meylan, France
- Sérasset, Gilles**, GETA CLIPS IMAG, Université Joseph Fourier Grenoble I, B.P. 53, 150 rue de la Chimie, 38041 Grenoble Cedex 9, France
- Szende, Thomas**, Centre interuniversitaire d'études hongroises, Université Sorbonne Nouvelle Paris III, 1 rue Censier, 75005 Paris, France
- Taifi, Miloud**, 20 rue Abdallahben Maskoud, Quartier Prince Héritier, Fès, Maroc
- Tanaka, Kumiko**, Takeichi Laboratory, Information Engineering Course, Graduate School of Engineering, The University of Tokyo, 1-3-7 Hongo Bunkyo-ku Tokyo, 113, Japon
- Thomas, Patricia**, Unit 54, Smithbrook Kilns, Cranleigh, GU6 8JJ, Royaume-Uni
- Tutin, Agnès**, URA SILEX, Université Lille III, B.P. 149, 59653 Villeneuve d'Ascq Cedex, France
- Viegas, Évelyne**, Computing Research Laboratory, New Mexico State University, Box 30001, Las Cruces, New Mexico 88003 8001, États-Unis
- Walther, Catherine**, Laboratoire d'analyse et de technologie du langage, Département de linguistique, Université de Genève, 1211 Genève 4, Suisse
- Wehrli, Éric**, Laboratoire d'analyse et de technologie du langage, Département de linguistique, Université de Genève, 1211 Genève 4, Suisse
- Wooldridge, Russon**, University of Toronto, Toronto (Ontario), Canada

Membres du Comité de réseau « LTT »

Chad, Mohammed, Professeur, doyen de la Faculté des lettres, Université Sidi Mohamed Ben Adballah, Fès, Maroc

Clas, André, Coordonnateur du réseau, professeur, directeur du GRESLET, Université de Montréal, Montréal, Canada

Ouoba, Benoît Bendi, Professeur, Université de Ouagadougou, Ouagadougou, Burkina Faso

Thoiron, Philippe, Professeur, directeur du CRTT, Université Lumière Lyon-2, Lyon, France

Goffin, Roger, Professeur, Université Libre de Bruxelles, Bruxelles, Belgique

Préface

André CLAS

Coordonnateur du réseau LTT, Université de Montréal, Canada

Lexicomatique et dictionnaires

Le réseau lexicologie, terminologie, traduction (LTT), créé en 1988, a tenu à l'Université Lumière-Lyon 2, du 28 au 30 septembre 1995, ses Quatrièmes Journées scientifiques sur le thème général **LEXICOMATIQUE et DICTIONNAIRES**.

Après Fès en 1989 (*Visages du français. Variétés lexicales de l'espace francophone*), il y a eu, en 1991, Mons (*L'environnement traductionnel. La station de travail du traducteur de l'an 2001*), puis en 1993, Montréal (*TA-TAO : recherches de pointe et applications immédiates*). Ainsi après avoir exploré des questions de variations lexicales en Francophonie (signalons que l'on trouve dans les *Actes de Fès*, en plus des contributions variées de divers spécialistes de la lexicographie une excellente *bibliographie scientifique concernant la langue française en Afrique noire*, et une présentation d'une proposition de *Trésor informatisé des vocabulaires francophones*), le réseau a abordé les autres domaines de sa compétence et notamment le rôle et l'importance de la traduction, le thème général se situait dans le secteur des postes de travail de traducteurs et l'automatisation de certaines tâches permettait d'aborder dans des perspectives résolument futuristes la question des stations de travail. On examine dans les *Actes de Mons* toute la problématique de l'utilisation de l'ordinateur et de son importance grandissante dans tout le secteur de ce que l'on peut regrouper sous l'étiquette de traitement des langues. Il était normal que les Journées suivantes se consacrent aux questions de la traduction assistée et automatique : un bilan des recherches en cours et des applications en vue s'imposait. On comprend donc ainsi mieux le thème retenu pour les Journées de Lyon. Toutes ces rencontres, on nous permettra de le dire, car nous avons des preuves éclatantes et des témoignages indubitables, ont été extrêmement fructueuses à de très nombreux égards. En plus de permettre de faire connaissance, de tisser de nouveaux liens, d'opérer certains rapprochements, de conclure des ententes pour des travaux en commun ou de mieux cibler certains travaux, elles ont été l'occasion pour tous de mieux savoir ce qui se faisait, de comprendre les progrès faits, d'examiner les nouvelles voies dans la recherche, bref d'aller de l'avant, de mieux savoir ce qui se faisait et de mieux connaître ceux qui le

faisaient. Ainsi de nouvelles recherches ont été lancées, de nouvelles orientations ont été trouvées et indubitablement des idées neuves sont nées. Il me semble donc pouvoir affirmer que le bilan a été, à notre humble avis, nettement positif. Il convenait aussi d'explorer de nouvelles avenues dans la recherche en génie linguistique et les transformations technologiques créant ce qu'on appelle les industries de la langue, il fallait examiner les domaines de la lexicologie, de la lexicographie, de la terminologie et de la terminographie. En un mot la préparation de nouveaux outils, d'outils linguistiques électroniques, d'outils qui puissent donner accès à des bases documentaires, à des réseaux des autoroutes de l'information, à des prolégomènes du cyberspace. Le thème de ces Journées devait permettre de mettre en perspective les différents outils destinés aux « langagiers », la variété des outils imposant en plus une réflexion sur les méthodes et les méthodologies de la recherche lexicographique. La *lexicomatique* se définissant comme l'ensemble des méthodes, des techniques et des pratiques qui utilisent l'informatique pour les explorations sémantiques, grammatico-lexicales du lexique d'une langue. On inclut ainsi facilement les débordements vers l'étude de corpus informatisés, de codages de corpus, de gestion de corpus bilingues et même vers la préparation de nouveaux outils pour la traduction assistée. La *dictionnaire* concerne toutes les questions techniques qui servent à élaborer de nouveaux dictionnaires et plus particulièrement des dictionnaires électroniques. On a ainsi pu montrer que la lexicologie et la lexicographie venaient d'entrer dans une nouvelle ère.

Si l'on examine quelque peu l'histoire de la lexicographie, on peut découvrir diverses étapes. Celle de la pragmatique d'abord : les lexicographes sont avant tout des praticiens, même si des relents théoriques percent ici et là, ce sont des dépouilleurs d'œuvres, des collectionneurs d'attestations, surtout littéraires. Peu à peu les sciences interviennent dans les dictionnaires, le dictionnaire a de plus en plus une vocation sociale, ce sont les Larousse et les Littré qui donnent la mesure. Aujourd'hui, c'est l'informatique qui apporte sa contribution à l'élaboration de dictionnaires. L'ordinateur a bouleversé la lexicographie, la lexicologie, et le dictionnaire même. Comme on le sait, le dictionnaire s'appuie de plus en plus largement sur des bases de données, donc des textes aux milliers d'occurrences de lexies, bien plus le dictionnaire est devenu lui-même base de données, où l'on peut retrouver les attestations les plus diverses. Les nouveaux supports informatiques permettent un emmagasinage gigantesque et un questionnement multiple et varié. La répartition entre dictionnaire de langue générale et dictionnaire spécialisé semble s'amenuiser, le nouveau type de dictionnaire étant en mesure de répondre en fonction des intérêts de l'utilisateur. C'est l'ère du dictionnaire à la carte qui s'ouvre.

Les textes dans ces *Actes* abordent toutes ces questions, présentent les diverses possibilités et explorent avec clarté les voies suivies ou à suivre. Ils fournissent dans maints cas des cheminements à entreprendre, des orientations à développer. Les méthodologies sont plurielles, les propositions sont multiples, et comme dans toute recherche scientifique on trouve des constructions de schémas ou des créations de modèles qui méritent d'être exploitées et réalisées.

C'est ainsi que l'on pourra mesurer tout l'intérêt que présente la proposition de **Jurij D. Apresjan** pour un *Enseignement du lexique assisté par ordinateur*, où l'on voit percer toute une nouvelle pédagogie stimulant le désir de savoir par des exercices programmés mais aussi par des jeux. **Étienne Brunet** avec sa communication sur son logiciel *Hyperbase* montre tout ce qu'on peut tirer d'un logiciel convivial et remet la

recherche philologique et littéraire à l'ordre du jour en intégrant les différentes versions d'un même texte sur ordinateur et surtout donne accès sur Internet à celui qui a illustré le mieux l'invention verbale, la saveur et la richesse linguistique : Rabelais. **Thierry Fontenelle** explore les *Réseaux sémantiques et les dictionnaires bilingues électroniques* en s'appuyant sur la théorie Sens \leftrightarrow Texte. L'utilisateur peut ainsi moduler à son gré des requêtes sur le réseau lexico-sémantique des fonctions lexicales mel'čukiennes. Avec **Étienne Blanc**, on aborde la présentation d'*Une maquette de base lexicale multilingue à pivot lexical : PARAX* qui est structurée pour l'expérimentation d'acceptions ou sens de mots et de leurs correspondants dans les autres langues et démontre son utilité notamment dans les questions qui appartiennent à la traduction automatique. **Lorne H. Bouchard** et **Louissette Ermikianian** présentent l'*Élaboration d'un dictionnaire informatisé pour le traitement automatique de la langue* dont les informations sont de nature morphologique et qui doit permettre le développement d'une grammaire computationnelle du français. Leur exposé est illustré par l'extraction de connaissances liées aux verbes de mouvement. **Guy Deville** et **Emmanuel Herbigniaux** décrivent leur projet *ANTHEM* dont l'objectif est le développement d'un prototype portable d'interface visant, par la création d'un modèle de représentation sémantique, la traduction et l'encodage automatiques de diagnostics médicaux. L'article de **Frédérique Segond** et **Élisabeth Breidt** présente *IDAREX* une *description formelle des expressions à mots multiples en français et en allemand dans le cadre de la technologie des états finis*, c'est-à-dire une expérimentation de compréhension des quasi-phrasèmes et des phrasèmes ainsi que des collocations. On voit tout l'intérêt que présente une telle étude pour un traitement de la langue tant en traduction automatique qu'en indexation automatique, par exemple. **Thomas Szende** et **Dominique Radanyi** traitent d'*aspects informatiques et pratiques dans la cadre de la lexicographie bilingue*. Leur expérience s'appuie sur la préparation du dictionnaire bilingue hongrois-français et montre, si besoin en était, que le rédacteur doit rester maître de son entreprise. **Patricia Thomas** et **Frank Knowles** s'attachent à déterminer l'orientation, ou mieux l'identification des bases et des collocataires dans les collocations en langues de spécialité. **Béatrice Daille** décrit *ACABIT : une maquette d'aide à la construction automatique de banques terminologiques* qui facilite la tâche des chercheurs en proposant un format et des structures morphosyntaxiques, grâce à des données linguistiques et statistiques, pour déterminer le recensement de termes spécialisés. **Didier Bourigault** décrit les problèmes théoriques et méthodologiques que présente la *Conception et l'exploitation d'un logiciel d'extraction de termes*. **Kumiko Tanaka** et **Violaine Prince** proposent un algorithme pour l'*Amélioration automatique incrémentale de dictionnaires bilingues utilisant un corpus monolingue* en s'appuyant sur des heuristiques telles que des informations morphologiques, la connaissance des synonymes et des valeurs de cooccurrence. **Xavier Lelubre** décrit sa *Conception d'un dictionnaire terminologique et phraséologique trilingue anglais/français-arabe dans le domaine de l'optique*. **Hussein Habaili** et **Mohamed Ben Ahmed** dans leur étude *Génération automatique de néologismes arabes* spécifient les règles lexicales qui permettent de dériver automatiquement des mots à partir des mots de base. **Miloud Taifi** dans son article *Construction des formes de mot et classification des entrées lexicales* présente de façon conventionnelle les problèmes particuliers de la lexicographie berbère. Les études de **Gilles Sérasset**, *Informatisation du Dictionnaire explicatif et combinatoire : le projet Nadia-DEC*, d'**Agnès Tutin**, *La formalisation des collocations pour le traitement automatique du langage naturel : le modèle des fonctions lexicales syntagmatiques*, et de **Margarita Alonso Ramos** et **Suzanne Mantha**, *Description lexicographique des collocatifs dans un dictionnaire*

explicatif et combinatoire : articles de dictionnaire autonome ?, appartiennent toutes trois, sous des aspects différents, informatisation et théorisation, à la même théorie linguistique, celle du Sens \Leftrightarrow Texte d'Igor Mel'čuk. **Christophe Jouis** et **Widad Mustafa-Elhadi** proposent un *nouvel outil interactif d'aide à la conception de dictionnaires électroniques spécialisés*. **Sylvain Delisle** dans son article *La construction de dictionnaires à partir de l'analyse informatisée de corpus bruts* vise à faciliter la rédaction d'un dictionnaire spécialisé, plus spécifiquement à préciser les propriétés syntaxiques et sémantiques des verbes. **Marc Van Campenhoudt** démontre dans son article *Réseau notionnel, intelligence artificielle et équivalence en terminologie multilingue : essai de modélisation* l'usage du réseau notionnel pour gérer divers problèmes d'équivalence dans une base de données multilingue. **Russon Wooldridge** et **Isabelle Leroy-Turcan** analysent dans *Les mots-clés métalinguistiques comme outil d'interrogation structurante des dictionnaires anciens* les difficultés qu'il faut envisager pour une exploitation optimale informatisée des dictionnaires anciens. **Michel Mathieu-Colas** dans son article *Représentation de la polysémie dans un dictionnaire électronique* propose le dégroupement maximal comme préalable à la description des liens qui lient les mots polysèmes. **Catherine Walther** et **Éric Wehrli** dans leur communication *Une base de données lexicale multilingue interactive* montrent tout l'intérêt qu'il y a à explorer les formes morphologiques associées à un lexème lorsqu'elles sont exprimées sous forme d'un ensemble de relations entre entrées indépendantes. **Évelyne Viegas** et **Sergei Nirenburg** démontrent dans leur article *Acquisition semi-automatique du lexique* le processus d'acquisition lexicale (Spanlex) dans un système de traduction automatique (Mikrokosmos). **François Gaudin** et **Myriam Bouveret** décrivent les *Pistes de description sémantique* telles que présentées dans *Biolex, dictionnaire des bio-industries*. **Pierrette Bouillon** dans son article *Le lexique génératif : une alternative au traitement de la polysémie* décrit le traitement de la polysémie des adjectifs qui dénotent un état mental dans le cadre de la théorie du lexique génératif. **Daniel Blampain** tantôt linguiste, tantôt usager de la langue décortique toute la problématique qu'il faut envisager *Quand l'informatique tutoie le dictionnaire des difficultés de la langue française*. **Philippe Barbaud** dans *Choix de grammaire et organisation du lexique* remet à jour la question du choix d'une grammaire et propose un réaménagement, plus spécifiquement une syntaxe dérivationnelle, pour le traitement informatisé des mots composés. **Philippe Blache** et **Mireille Delpui** décrivent un *Outil d'intégration de bases de connaissances lexicales aux analyseurs syntaxiques* et plaident pour une construction de lexiques adaptés à un formalisme donné. **Karim Chibout** et **Nicolas Masson** présentent comme thème d'étude *Un réseau lexico-sémantique des verbes construit à partir du dictionnaire pour le traitement automatique du français*.

On trouvera dans *Lexicomatique et dictionnaires* une richesse expérimentale lexicologique et lexicographique tout à fait remarquable. Il y a là les modèles les plus variés, les applications les plus profondes, les réflexions les plus prolifiques en lexicologie, en lexicographie, en sémantique pour un développement fantastique des industries de la langue. On voit que le travail artisanal de la lexicographie a donné naissance à un savoir technique imprégné de connaissance scientifique. C'est en effet la science qui exige une réduction à des schémas qui permettent des substitutions. On voit facilement la complémentarité qui existe entre les données linguistiques et l'informatique. Il y a là des rapports étroits, un appui mutuel, si l'on peut dire, qui décrit bien la diversité des méthodes mais l'unicité de l'esprit scientifique qui s'appuie sur la réalité en vue de décrire, d'expliquer et de valider. Les schémas, les modèles, s'ils sont nécessairement abstraits doivent cependant subir la procédure de vérification, en

Préface

linguistique comme dans les autres sciences. La langue étant une des plus extraordinaires créations de l'Homme.

Pour de simples raisons financières, il n'a pas été possible d'inclure dans ce volume toutes les excellentes communications présentées aux IV^{es} Journées scientifiques de Lyon. Nous le regrettons sincèrement et nous nous en excusons auprès des collègues. Ces communications seront cependant publiées dans des revues spécialisées reconnues internationalement.

