

INFORMATIQUE ET LITTÉRATURE COMPARÉE

Alain VUILLEMIN

Parce qu'elle se situe par définition au carrefour des études littéraires françaises et étrangères et aussi à la rencontre des disciplines de lettres, de langues et de sciences humaines, puisque sa vocation principale porte surtout sur l'histoire des relations littéraires et intellectuelles internationales, la littérature comparée est en train de devenir en cette dernière décennie du XX^e siècle, dans l'Université française, l'une des premières disciplines de lettres à être vraiment concernée par l'essor spectaculaire de l'édition électronique depuis 1992.

De documentaires, les disques optiques compacts, les CD-ROM ou DC-MEF¹, sont devenus éditoriaux et multimédias. Le progrès de l'informatique et des technologies de la documentation et de l'information ont fait que les textes, les images, les sons se sont intégrés et que les disques compacts se sont transformés en de nouveaux « vecteurs » et en de nouveaux « supports transportables » de l'édition. Annoncé dès 1989², amorcé en 1992, le phénomène s'est accéléré en 1993 et en 1994, et devrait « exploser » en 1995, mettant ainsi les études de lettres, toutes disciplines et spécialités confondues, au défi de les utiliser.

L'apparition de ces nouvelles sources d'information littéraire qui était réservées jusqu'en 1990 à la consultation de rares spécialistes³ ne pourra pas demeurer indéfiniment méconnue par la grande masse des étudiants et des enseignants. Les titres des disques compacts disponibles sur des sujets littéraires se comptent désormais par centaines et par milliers. Et, pour s'en tenir au seul domaine de la littérature comparée, cette situation ne saurait manquer d'avoir des répercussions sur les pratiques de l'enseignement et de la recherche. Bien des idées préconçues risquent d'ailleurs de se trouver remises en question à brefs délais car ce

1 « CD-ROM » pour « Compact Disc-Read Only Memory », ou « DC-MEF » pour « Disque Compact à MEmoire Fixée ».

2 voir Vuillemin (Alain) : "L'Avenir du disque compact", in Vuillemin (Alain) et alii : *Les Marchés de l'information documentaire*, Paris, ADBS, 1991, p. 143-153.

3 Voir Vuillemin (Alain) : *Informatique et littérature*, Genève-Paris, Slatkine-Champion, 1990, p. 131 et sq.

sont, pour les comparatistes, de nouveaux matériaux de travail et de nouveaux territoires de recherche qui surgissent.

I. DE NOUVEAUX MATERIAUX DE TRAVAIL

Ce sont en effet de nouvelles sources d'information et de nouveaux matériaux de travail qui apparaissent sur les études littéraires, depuis la consultation des banques de données bibliographiques jusqu'à l'exploitation de « corpora », de banques de données textuelles, en texte intégral, sur les principales littératures.

Le recours aux banques de données bibliographiques tend à devenir une pratique courante, dans la recherche au moins, s'il n'est pas encore pratiqué partout dans l'enseignement universitaire français. Pour les comparatistes, qui sont contraints de travailler sur plusieurs littératures, les banques de données bibliographiques qui sont les plus utiles sont étrangères et américaines. C'est la *MLA Bibliography*, qui comporte toujours une édition imprimée mais qui est diffusée en ligne ⁴ et qui a commencé aussi à être distribuée depuis 1992 sur des disques compacts documentaires. Ce sont aussi les bases *Arts and Humanities et Research in Progress Data base* ⁵, cette dernière sur les recherches en cours. En France, en comparaison, les bases disponibles, *Téléthèse*, *Francis*, *Electre*, le *Catalogue Collectif National* (C.C.N.), des publications en série, ne sont que des banques de données « de second ordre » ⁶. Seule la multiplication de travaux bibliographiques analogues au projet « Orphée » de l'AUPELF-UREF ⁷, qui a abouti en 1993 à la réalisation d'un premier disque compact documentaire en langue française, *Orphée. Volume 1*, sur les littératures francophones, permettrait de pallier cette carence.

Ce serait une urgence. Que ce soit à des fins documentaires ou éditoriales, la production de disques compacts n'était qu'une virtualité ⁸.

4 pour autant que l'on eût réussi à trouver en France, en 1994, une bibliothèque universitaire ou un centre de documentation qui y ait souscrit un abonnement !

5 voir Panijel (Claire) : "Des banques de données sur les études littéraires francophones", in Vuillemin (Alain) et alii : *Les Banques de données littéraires, comparatistes et francophones*, Limoges, Pulim, 1993, p. 33-43.

6 voir Faule (Jacques) : "Téléinformatique et littérature française", *ibid*, p. 45-47.

7 AUPELF-UREF : Association des Universités Partiellement ou Entièrement de Langue Française - Université des Réseaux d'Expression Française.

8 voir Vuillemin (Alain) : "L'Avenir des disques compacts", in Vuillemin (Alain) : *Les Marchés de l'information documentaire*, Paris, ADBS, 1991, p. 143-153.

En 1992, en forçant sur le sens du terme « littéraire », on ne dénombrait guère qu'une dizaine de disques littéraires, qui ne suscitaient qu'un mince intérêt d'estime de la part des spécialistes. En 1993, le principal répertoire international qui existait, *The Cd-rom directory* de TFPL Publishing, en identifiait déjà plus d'une centaine diffusés dans sa rubrique « Literature ». Il recensait des bibliographies certes mais aussi des anthologies et des éditions de textes sur des sujets variés, de la poésie au roman et au théâtre, avec les disques *Living Poetry*, *Poetry in Motion*, *Poem Finder on Disc*, *The English Poetry Full-Text Database*, des anthologies comme *Survey of English Literature* sur la littérature anglaise, *CD-Littérature*, *Great Literature* et *Library of the future* sur la littérature mondiale, et des éditions en texte intégral telles que *The Brontës : The Complete Novels and Selected Minor Works with Illustrations*, *Jane Austen : The Complete Illustrated Novels* et, déjà, pas moins de cinq éditions différentes du théâtre de Shakespeare : *Shakespeare* par Creative Multimedia Corporation, *Shakespeare's Twelfth Night or What You Will* par Anglia Polytechnic University, *Shakespeare* par New Media Schoolhouse, *Illustrated Shakespeare* et *Illustrated Shakespeare (CDTV)* par Animated Pixels Ltd. D'autres disques portaient sur les sources de la littérature grecque, comme *Perseus 1.0*, la littérature latine chrétienne, la littérature australienne, les écrivains japonais et les auteurs américains et britanniques.

En 1994, la nouvelle édition du *CD-rom directory* de TFPL Publishing répertoriait près de 160 disques disponibles dans sa rubrique « Literature », auxquels il conviendrait d'ajouter sans doute un bon nombre de titres inventoriés sous les rubriques « Arts & Humanities » (i.e. : "Lettres et sciences humaines") et « Languages » (i.e. : "Langages"). Les frontières des études de lettres ne sont pas identiques en effet entre le monde anglo-saxon et le monde francophone. Toutefois, si imparfaite que soit cette esquisse de recensement, de nouvelles tendances se manifestaient, en particulier vers une certaine diversification des productions. Les bibliographies restent prépondérantes. Des index de citations apparaissent, comme *Arts & Humanities Citation Index* et *Zitatenhandbuch Dr. Puntsh (Dictionary of Quotation) EB*, de même que des index biographiques, *Biography Index*, des dictionnaires de la littérature, *The Dictionary of cultural Literacy (EB)*, des histoires littéraires en anglais, *A Multimedia History of American Literature*, et en italien, *I Cataloghi Impossibili - vol 7 (The History of Literature)*, des éditions multimédias, telles que *A Christmas Carol by Charles Dickens*, *A Multimedia Storybook*, des éditions spécifiques, *Call of the Wild by Jack London (EB)*, l'édition illustrée par G. Doré de *La Divine Comédie* de LA REVUE DE L'EPI INFORMATIQUE ET LITTÉRAURE COMPARÉE

Dante, Dante Alighieri, and Gustave Doré - Electronic Version, des éditions de texte, soit sélectives, de saint Thomas d'Aquin à Aristote et Platon en passant par Hobbes, Kierkegaard, Wittgenstein et par les philosophes rationalistes, et de Mark Twain ou de Charles Dickens à Lewis Carrol ou Philip Marlowe. Il n'est pas jusqu'aux analyses et aux commentaires critiques des oeuvres, si contestables soient-elles, les *Baron's Book Notes* et les *Monarch Notes* qui ne connaissent désormais des éditions électroniques sur des disques compacts.

Le résultat sera jugé hétéroclite. Des éditions analogues, *Liz* sur la littérature italienne et en langue italienne, *Letras* sur la littérature espagnole et en langue espagnole, étaient annoncées pour le courant de l'année 1994. L'important est de relever que chacun de ces titres, quel qu'en soit le sujet, inaugure un phénomène nouveau, dont l'extension s'annonce vertigineuse. C'est par milliers, à partir de 1995, prévoient les spécialistes, que les disques compacts seront produits. Leur existence ne pourra plus être ignorée.

II. DE NOUVEAUX TERRITOIRES DE RECHERCHE

Les processus précédents devraient s'accélérer encore pour d'autres raisons. En 1991, l'Université de Georgetown a publié un premier inventaire de 329 organismes qui s'étaient déjà engagés dans des programmes de numérisation de textes littéraires⁹, dont 136 aux Etats-Unis, 36 en Grande-Bretagne, 20 au Canada, 26 en Allemagne, 20 en France. L'« Oxford Text Archive » avait déjà recueilli à cette date, auprès de l'Université d'Oxford en Grande-Bretagne, plus de mille textes sur le patrimoine littéraire anglais. Un projet analogue, étendu à la littérature européenne, « The European Corpus Initiative », s'est mis en place depuis, en regroupant les Universités d'Edimbourg, de Pise, de Pennsylvanie, de Münster et de Genève. Les grandes bibliothèques, de la Bibliothèque du Congrès de Washington aux Etats-Unis à la Bibliothèque Nationale de France, à Paris, ont mis enfin en chantier la constitution d'immenses collections de textes numérisés. En 1993, la Bibliothèque du Congrès disposait déjà d'un fonds documentaire de 6 000 disques compacts et une bibliothèque immatérielle de plus de 25 000 documents sur la seule « Mémoire de l'Amérique ». En 1996, à son inauguration, la Bibliothèque Nationale de France prévoit de disposer à son tour, à cette échéance, d'une « bibliothèque » analogue (y compris en littérature

⁹ voir Richard (Michel) : "Le Programme de numérisation de la Bibliothèque de France", in *Bulletin Bibliographique de France*, Paris, 1993, t38, n°3, p. 53-63.

comparée) de 100 000 à 300 000 volumes selon les chiffres avancés. Ce seront aussi, alors, d'autres instruments de travail et surtout d'autres territoires de recherche qui s'ouvriront.

De premières tentatives comme la création de la collection « CAP'AGREG » en littérature française, aux éditions Nizet, riche de quatre numéros sur Huysmans, Renan, Molière et Céline, commencent à montrer comment le recours aux nouvelles technologies permettrait de mieux étudier un texte, une oeuvre ou un auteur. Le principe mériterait d'ailleurs d'être étendu aux études de littérature comparée au fur et à mesure que de grands « corpora » de données nationaux ou internationaux, vont se constituer. Shakespeare, Dante, Goethe, Claudel sont désormais numérisés. Il en sera de même, en des délais assez brefs, des principaux auteurs de chaque grande littérature. Mais des éditions de textes informatisés ne s'appréhendent pas de la même manière que des livres imprimés. La conception en est différente. Les méthodes d'approche sont autres. Les contenus en sont plus complexes. Les démarches critiques en sont quelquefois moins assurées. Les résultats que l'on en obtient exigent d'être examinés avec rigueur. Les cinq éditions sur disques compacts des oeuvres de Shakespeare que l'on dénombrerait au début de 1994 reposent sur des principes distincts, que ce soit sur un plan technique ou littéraire. A mesure que vont se multiplier les éditions en langue originale et, on peut le prévoir, les traductions sur des disques ou sur tout autre support informatique concevable, le « regard » critique des comparatistes sur ces nouvelles formes de « productions » littéraires deviendrait très nécessaire. Ce sont aussi d'autres objets d'étude qui vont se constituer et ce sont sans doute les recherches sur les aspects méthodologiques et épistémologiques de la discipline qui s'en trouveraient renouvelés.

D'autres réalisations, encore ponctuelles, suggèrent également ce qui pourrait être envisagé à une échelle plus vaste, sur plusieurs littératures. La *Banque de données d'histoire littéraire* (BDHL) de l'Université de La Sorbonne Nouvelle (Paris III)¹⁰ préfigure, sur la littérature française, ce que de futures banques de données semblables pourraient apporter sur l'histoire littéraire générale. L'apparition de premières tentatives d'édition de textes littéraires bilingues, comme la banque de données *El Hadj* de l'Université de Paris Sorbonne (Paris

10 voir Bernard (Michel) : "La banque de données d'histoire littéraire", in Vuillemin (Alain) et alii : *Les Banques de données littéraires, comparatistes et francophones*, Limoges, Pulim, 1993, p. 253-269.

IV) ¹¹ sur laquelle des modules de traduction automatique en français et en anglais avaient été greffés ouvre de nouvelles perspectives aux travaux et aux réflexions sur l'acte de traduction. Les progrès de l'accès multilingue aux banques de données informationnelles auront peut-être, à terme, dans un avenir indéterminé, des retombées sur la pratique de la traduction littéraire. L'élaboration de banques de « faits » littéraires de toute nature, linguistiques, stylistiques, historiques, critiques, devient aussi concevable avec l'expérimentation en cours depuis 1992 en différentes universités françaises des nouvelles « stations de lecture assistée par ordinateur » dont la Bibliothèque Nationale de France prévoit de s'équiper et dont les bibliothèques universitaires devraient disposer aussi. Là encore, ce sont de nouvelles perspectives qui apparaissent et, si les bibliothèques que l'on appelle « immatérielles » ou « électroniques » ¹² n'en étaient, en 1994, qu'à leurs tous premiers balbutiements, les ressources qu'elles proposeront aux chercheurs ne feront qu'aller en s'accroissant et en se multipliant.

L'informatique est aussi un instrument de création littéraire ¹³ et il n'est pas jusqu'à la « littérature générée par ordinateur » qui ne devienne un autre domaine de recherche pour la littérature comparée. Un premier colloque international : « Nord Poésie et Ordinateur », à l'Université de Lille III, l'a fait découvrir le 12 mai 1993 ¹⁴. Le congrès international qui s'est tenu depuis en Sorbonne sur la littérature et l'informatique, du 19 au 24 avril 1994, sous les auspices de l'Association for Literary and Linguistic Computing l'a confirmé. Trois journées d'études internationales, organisées sur ce même sujet du 20 au 22 avril 1994 par l'association ITEM-Sup, l'Université canadienne de Montréal et les universités françaises Denis Diderot (Paris VII) et d'Artois ont permis d'en approfondir les implications. Le phénomène est mondial. De la Norvège à la Nouvelle Zélande en passant par le Brésil, le Canada, les Etats-Unis, la Grande-Bretagne, l'Italie, l'Espagne, Israël, les expérimentations se multiplient. Deux thèses, l'une en portugais à

11 voir Vuillemin (Alain) : « El Hadj » : une maquette de banque de données littéraires, éditoriale et bilingue, en littérature comparée", *ibid*, p. 221-227.

12 voir Belbenoit-Avich (Pierre-Marie) : "La Bibliothèque électronique : bibliothèque de demain ou d'aujourd'hui ?", in *Bulletin Bibliographique de France*, Paris, 1993, t38, n°6, p. 60-65.

13 voir Vuillemin (Alain) : *Informatique et Littérature*, Genève-Paris, Slatkine-Champion, 1990, p. 209 et sq.

14 voir Bootz (Philippe) et alii : A : \ *Littérature*, Lille-Villeneuve d'Ascq, CIRCAV-GERICO-MOTS-VOIR, 1994, 146p.

Lisbonne, en 1992 de P. Barbosa sur *Criacão Literaria e Computador*¹⁵ et la seconde, en espagnol, à Madrid, en 1991, d'O. Carreño Rodriguez-Maribona sur *Nuevas Tecnologias de la Informacion y creacion Literaria*¹⁶ ont déjà été soutenues sur les apports de l'informatique et des nouvelles technologies à la création littéraire et se trouvaient en cours de traduction en français en 1994. Ces nouvelles formes de création concernent tous les domaines, de la poésie sous toutes ses formes. Une nouvelle littérature est en train d'émerger. Le phénomène est mondial.

L'intégration de l'informatique et des nouvelles technologies était très loin d'être acquise en France, en 1994, dans les études de Littérature comparée. En 1983, dans *Qu'est-ce que la littérature comparée ?* P. Brunel notait déjà, à propos des perspectives que la traduction automatique laissait espérer à cette époque, que « les progrès réalisés en vingt ans [entre 1960 et 1980] ont été tels que des tâches plus subtiles pourraient bien un jour lui [à l'ordinateur] être confiées »¹⁷. Dix ans après, en 1992, dans le guide de *L'Étudiant chercheur en littérature*, Y. Chevrel, tout en laissant encore de côté "le problème de la « création de textes »¹⁸", observait néanmoins que l'informatique était désormais devenue un instrument de travail précieux et que l'on se trouvait « effectivement [à cette date] aux débuts de métamorphoses importantes, accès aux documents, stockage de données, recherche d'informations, traitement de texte, etc., autant de domaines dans lesquels une révolution [était] en train de s'opérer »¹⁹. Cette « révolution » annoncée, parfaitement prévisible et prévue au demeurant²⁰, s'est produite en effet.

Quelques dates et quelques chiffres aussi donneront une idée de la brusque accélération de cette évolution. Le principe de la conception des disques compacts dérive de celui de la technologie des vidéodisques et il est connu depuis 1926. Les premiers disques prototypes ont été fabriqués entre 1970 et 1980. En 1987, on dénombrait, dans le monde, 48 disques compacts environ, toutes catégories confondues. En 1988, la société américaine TFPL Publishing en recensaient 189, 390 en 1989, 817 en

15 Barbosa (Pedro) : *Criacão Literaria e Computador*, Lisboa, Faculdade de Sciencia sociais e Humanes, 1992, 420 p.

16 Carreño Rodriguez-Maribona (Orlando) : *Nuevas Tecnologias de la Informacion y Creacion Literaria*, Madrid, Universidad Complutense, Facultad de Filologia, 1991, 2vol : 698 p.

17 Brunel (Pierre) : *Qu'est-ce que la littérature comparée ?*, Paris, Colin, 1983, p. 146.

18 Chevrel (Yves) : *L'Étudiant chercheur en littérature*, Paris, Hachette, 1992, p. 118.

19 Ibid, p. 117.

20 voir Vuillemin (Alain) : "L'avenir du disque compact", in Vuillemin (Alain) et alii : *Les Marchés de l'information documentaire*, Paris, ADBS, 1991, p. 143-153.

1990, 1 522 en 1991, 2 212 en 1992, 3 597 en 1993 et 5 503 enfin en 1994. La proportion des disques littéraires est encore minime mais l'on sait depuis la publication en 1989 du livre de R. Lillet, *Pour une Europe du livre*²¹ que les applications éditoriales de cette technologie nouvelle sont illimitées.

Le disque se substitue ainsi au livre et l'informatique commence aussi par ce biais à pénétrer dans les études de lettres. Ce sont en effet d'autres sources d'information, d'autres catégories de textes, de matériaux et de territoires qui s'ouvrent pour la recherche. La connaissance de l'histoire littéraire, la perception des relations entre les littératures, l'étude des échanges entre la littérature française et les littératures étrangères en seront certainement affectées. Mais ce processus sera long. L'expérience prouve que « l'aptitude des humanités à tirer parti de la technologie »²² n'est pas démontrée. Les hésitations, les réticences en entraveront peut-être encore longtemps l'intégration dans la recherche littéraire.

Alain VUILLEMIN
 Université d'Artois
 Centre d'Etudes et de Recherches
 sur les Textes Electroniques Littéraires
 Membre du Bureau national de l'EPI

²¹ voir Lillet (René) : *Pour une Europe du livre*, Paris, La Documentation Française, 1990, p. 14-sq.

²² Hirscheim (Rudy), Smithson (Steve), Horwood (Ellis), in Vuillemin (Alain) : "L'informatique dans les humanités en Grande-Bretagne", Paris, *Bulletin* de l'EPI, 1992, n°67, p. 202-203.