

L'INFORMATIQUE DANS LES ETUDES LITTERAIRES

Alain VUILLEMIN

Balbutiante jusqu'en 1990, date à laquelle apparaît un premier ouvrage de référence, *Informatique et littérature (1950-1990)*¹, l'utilisation de l'informatique a commencé à se faire connaître dans les études littéraires, au cours de la dernière période, avec la constitution de premières équipes de recherche dans les milieux universitaires, en littérature française depuis 1992 autour de l'Université de La Sorbonne Nouvelle (Paris III), d'H. Béhar et du collectif Hubert de Phalèze, et, en littérature comparée, à l'Université d'Artois, avec la création du Centre d'Études et de Recherches sur les Textes Electroniques Littéraires (CERTEL)². Ce processus pourrait s'accélérer. Le monde du livre était en France, en 1995, à la veille de connaître de grands changements et l'essor prévu de l'édition électronique, à partir de 1997, pourrait transformer d'une manière radicale les conceptions admises jusqu'à présent du livre, de la lecture et de la littérature. De fait, avec la création de la collection « Cap'Agreg » à Paris, aux éditions Nizet³, en 1992, avec aussi la multiplication des numéros spéciaux des revues consacrées à l'informatique depuis le n°95 d'*Action Poétique* dès 1984 jusqu'au n°96 de la revue

¹ Vuillemin (Alain) : *Informatique et littérature (1950-1990)*, Genève-Paris, Slatkine-Champion, 1990.

² Le Centre d'Études et de Recherches sur les Textes Electroniques Littéraires (CERTEL) de l'Université d'Artois a pour vocation d'explorer les ressources que l'informatique et que les nouvelles technologies de l'information, de la communication et de la documentation proposent désormais à la recherche dans les études de lettres, en littérature générale et comparée, sur les relations littéraires et intellectuelles internationales. Le C.E.R.T.E.L. est membre associé de la Fédération ITEM-Sup (Intégration des technologies éducatives modernes dans l'enseignement supérieur).

³ A savoir : Phalèze (Hubert de) : *Compte "A Rebours". L'oeuvre de Huysmans à travers les nouvelles technologies, Renant tout comptes faits. Souvenirs d'enfance et de jeunesse à travers les nouvelles technologies, Les Mots de Molière, Guide de "Voyage au bout de la nuit", Voltaire portatif. Le "Dictionnaire philosophique" à travers les nouvelles technologies. Dictionnaire des "Misérables". Dictionnaire encyclopédique du roman de Victor Hugo réalisé à l'aide des nouvelles technologies. Les Voix de la Condition humaine. "La Condition Humaine" d'A. Malraux à travers les nouvelles technologies.* Paris, Nizet, 1991-1995.

Littérature en 1994, l'informatique est entrée depuis longtemps dans les études de lettres. Mais les recherches en cours, bien que de plus en plus nombreuses, que ce soit sur les littératures anciennes, la littérature médiévale ou les littératures modernes, n'ont pas encore abouti. Aucun inventaire systématique n'en existe non plus. Il est difficile pour cette raison d'apprécier où et comment les utilisations littéraires s'y développent même si l'on peut en esquisser une typologie rapide, et essayer d'en décrire les territoires nouveaux, autrement dit les principaux domaines d'application, de découverte et de prospective.

I. LES DOMAINES D'APPLICATION

Les domaines d'applications de l'informatique ne cessent de se multiplier dans les études de lettres, même si la pratique n'en est pas encore partout courante. Sur ce plan, toutefois, il est des truismes qu'il est nécessaire de rappeler : les littéraires travaillent sur des livres. Il leur faut donc les identifier, ne serait-ce que pour les lire. Leurs réflexions aussi donnent lieu à leur tour à l'élaboration d'autres livres. Et, même si c'est encore d'une manière éclatée et auxiliaire, l'usage de l'informatique et des ordinateurs commence à apparaître comme une ressource privilégiée dans la recherche bibliographique, dans l'édition savante et dans les relations internationales.

Avec le développement des banques de données et l'extension du réseau de télécommunication « Internet », la recherche bibliographique ne peut plus ignorer l'apport des répertoires informatisés des thèses dont on ne citera que deux : *Docthèses* (dérivé de *Téléthèses*) en langue française, qui inventoriat en 1995 230 000 thèses soutenues en France depuis 1972 et, en langue anglaise, *Dissertation Abstract Online*, qui recense 90 % des thèses américaines et, depuis 1988, un grand nombre de thèses canadiennes et européennes. En ce qui concerne les ouvrages imprimés, les principales ressources sur la littérature française sont désormais représentées par la *Bibliographie Nationale Française* de la Bibliothèque Nationale de France, qui comptait en 1995 environ 1 300 000 notices bibliographiques, et par les catalogues collectifs *Sibil* et *Pancatalogue* auxquels participent presque toutes les bibliothèques universitaires françaises et enfin, pour les revues, le *Catalogue Collectif National* des publications en série, qui catalogue et qui localise 190 000 titres de périodiques auprès de 2 700 bibliothèques et centres de documentation environ, situés en France. La consultation de la banque de données *Electre* du Cercle de la Librairie permet enfin de savoir quels

sont les ouvrages immédiatement disponibles dans toutes les maisons d'édition françaises. Sur ce plan néanmoins, dans les études littéraires, le principal instrument de références reste la *MLA Bibliography*, en langue anglaise, riche en 1995 de plus de 12 000 000 de références recueillies en près de 40 langues sur les littératures nord-américaines, anglaises, européennes, asiatiques, africaines et latino-américaines. L'accès à « Internet » permet enfin de consulter immédiatement - et directement - l'ensemble des bibliographies ou des banques de données analogues dans les principaux pays industrialisés, aux Etats-Unis ou en Grande-Bretagne, mais aussi au Canada, en Belgique, en Allemagne, en Espagne, dans les pays scandinaves, etc.

Dans un tout autre secteur, l'édition savante commence à être affectée par les transformations de l'édition électronique. Déjà, de nombreuses revues scientifiques n'existent plus que sous une forme électronique dans les sciences exactes. Dans les études littéraires, la revue comparatiste *Surfaces*⁴, en langue française, du département de littérature comparée de l'Université de Montréal préfigure ce que les revues littéraires savantes risquent de devenir à brefs délais sur « Internet ». On peut aussi signaler l'existence sur le réseau « Internet » de la revue littéraire canadienne *Axe* en langue française, spécialisée dans la littérature québécoise et francophone. Ailleurs, dans un proche avenir, à Paris, à la Bibliothèque Nationale de France, ce sont des milliers d'ouvrages de référence sur la langue et la littérature française, des dictionnaires, des encyclopédies, des grammaires, des ouvrages de rhétorique et les oeuvres complètes des principaux écrivains de langue française qui deviendront accessibles sous une forme numérisée à partir de 1997. A cette date, prévoit-on, un premier fonds documentaire numérisé de 100 000 volumes au moins, qui avaient déjà été saisis dès la fin de 1995, sera disponible. A terme, en l'an 2000, c'est une « bibliothèque immatérielle » de 300 000 volumes environ qui sera proposée aux lecteurs de la Bibliothèque Nationale de France. Ce sont de nouveaux modes de « lecture assistée par ordinateur » qui vont s'instaurer. C'est peut-être aussi un autre type de relation à la lecture et à la littérature qui va apparaître.

Il n'est pas jusqu'aux relations intellectuelles internationales qui risquent d'en être transformées. Déjà, l'emploi de la télémessagerie a permis, dès 1986, en France, à Villeneuve-lès-Avignon avec *Marco Polo*

⁴ *Surfaces* : Internet acces via FTP anonymous : harfang.cc.umontreal.ca

ou le *Nouveau Livre des merveilles*⁵ ou à Montréal, au Canada avec *La Plissure du Texte*⁶, de faire l'expérience de tentatives d'écriture collective, francophones, qui ont associé des écrivains de nationalités différentes et qui ont connu depuis des émules aux Etats-Unis et au Canada, et jusqu'en Australie, en langue anglaise. Depuis, l'imbrication des réseaux de télécommunications « Internet »⁷, « Renater »⁸, « Refer »⁹, a permis aux universités, françaises ou étrangères, de se connecter et de multiplier les échanges, depuis l'organisation de colloques jusqu'aux relations personnelles entre chercheurs, par exemple par l'intermédiaire du réseau *Humanist*¹⁰ (sur les humanités en général) ou du réseau *Balzac*¹¹ (sur la littérature française du XIX^e siècle), ou l'organisation de séminaires virtuels, par exemple sur la *Divine Comédie* de Dante. Les frontières entre les littératures et entre les individus tendent ainsi à s'estomper. Le phénomène ne fait que commencer. Les réseaux de télécommunication deviendront-ils de nouvelles « arènes électroniques » créatrices ?

Tels étaient les principaux domaines d'application concrète, immédiate, de l'informatique dont on pouvait faire état en 1995. Qu'ils aient existé alors ne signifiait pas pour autant qu'ils eussent été à cette date d'une pratique courante dans les universités françaises. Mais ils n'étaient plus à proprement parler des domaines d'exploration.

II. LES DOMAINES DE DECOUVERTE

De fait et de pair avec les transformations incessantes de l'informatique et des nouvelles technologies de l'information, les domaines de découverte ne cessent de se multiplier aussi. Certes, les recherches qui les concernent et dont on peut faire l'état sont encore modestes. De nouvelles perspectives se dessinent toutefois, de nouveaux territoires de

⁵ Adiaffi (J.-M.), Caron (L.), Delay (F.), Kacem (A.), Labou Tansi (S.), Lacarrière (J.), Savoie (J.), Visage (B.) avec la participation d'I. Calvino et d'U. Eco : *Marco Polo. Le Nouveau Livre des Merveilles*, La Chartreuse-lès-Avignon - Malakoff ; CIRCA- Solin, 1985.

⁶ Ascott Roy : *La Plissure du texte*, Paris, Musée d'Art Moderne, Exposition « Electra », 1985.

⁷ Internet : International network (pour « réseau international de télécommunications »).

⁸ Renater : Réseau National pour l'Enseignement et la Recherche (entre les universités françaises).

⁹ Refer : Réseau Francophone pour l'Enseignement et la Recherche (entre les universités francophones). Consulter en particulier le site : <http://www.refer.qc.ca/AUPELF.html>

¹⁰ *Humanist* : <http://www.princeton.edu/mccarty/humanist/>

¹¹ *Balzac* : balzac-l@cc.montreal.ca

recherche apparaissent, à mesure que s'affirment de nouvelles formes d'édition électronique, d'analyse des textes et de réflexion critique.

Embryonnaire jusqu'au début des années 1990 avec la constitution ici ou là de fonds documentaires textuels numérisés, la banque *Frantext*¹² en France sur la littérature française du XVI^e au XX^e siècles, l'*Oxford Text Archives* en Angleterre, à l'Université d'Oxford, sur la littérature anglaise, pour ne citer que ces deux exemples, l'édition électronique des textes littéraires ne cesse de se développer depuis 1992. Il n'existait guère en 1992, sur disques compacts, qu'une dizaine de titres « littéraires », toutes catégories confondues et en forçant sur le sens du terme « littéraire ». En 1995, le principal répertoire international qui existait, *The CD-Rom Directory* de TFPL Publishing aux Etats-Unis identifiait plusieurs centaines de titres dans ses rubriques « Literature », « Language », « Arts ». Il recensait des bibliographies mais aussi des anthologies, des éditions de textes, sélectives ou systématiques, multimédias ou non. Il en prévoyait le triplement ou le quadruplement à partir de 1996. Les perspectives annoncées sont vertigineuses. Les réalisations effectives le sont aussi. Les éditions Champion auront publié ainsi, en français, depuis la fin de 1995, les *Essais* de Montaigne¹³, sur disquettes et les *Oeuvres complètes* de Rabelais¹⁴, au début de 1996, sur un disque compact depuis les éditions princeps du XVI^e siècle jusqu'aux éditions plus récentes du début du XX^e siècle. Ce sont de nouveaux matériaux de travail qui seront proposés aux chercheurs. Ce sont peut-être d'autres conceptions de l'édition savante de textes électroniques littéraires qu'il serait urgent de concevoir.

Ce sont aussi d'autres formes d'analyse des textes, appropriées à la nature mouvante et éminemment complexe de ce qu'on appelle un « texte littéraire », qu'il conviendrait aussi de définir et d'explorer. Une fois transféré sur un support informatique, un « texte » se révèle constitué par un tissu de mots éminemment mobiles, malléables, décomposables ou transformables à l'infini. Le « statut »¹⁵ du texte change. L'approche des textes ne se limite plus aux seules analyses morphologiques, syntaxiques ou sémantiques auxquels trop d'informaticiens et de linguistes les réduisent. Ce sont d'autres "modèles d'analyse (...) propres au texte électro-

¹² *Frantext* réunissait en 1995 environ 3 000 textes d'oeuvres complètes d'écrivains français, du XVI^e au XX^e siècle, reproduits en texte intégral et consultables en ligne.

¹³ Montaigne (Michel de) : *Essais*, Paris, Champion « Electronique », 1995.

¹⁴ Rabelais (François) : *Oeuvres complètes*, Paris, Champion « Electronique », 1996.

¹⁵ Lusignan (Serge) : "Quelques réflexions sur le statut épistémologique du texte électronique", dans *Computers and the Humanities*, Osprey (FL), 1985, n°19, p. 210.

que" ¹⁶ littéraire, à sa forme, à son genre, à ses conventions rhétoriques, à ses licences, à ses procédés d'expression et jusqu'à sa mélodie, sa musique, qu'il faudrait imaginer et intégrer. Encouragées pendant un temps, entre 1960 et 1985, et transposés à des objets littéraires, les méthodes statistiques, les concordances, les index, utilisés d'une manière isolée, n'ont guère contribué à transformer l'approche des oeuvres, même si Ch. Muller puis E. Brunet ont publié en France, en ce domaine, des études déterminantes sur P. Corneille ou sur J. Giraudoux, V. Hugo, E. Zola, M. Proust ou, plus récemment, sur J. Gracq ou sur Rabelais. Là encore, par un paradoxe qui s'explique en partie par les cloisonnements qui existent entre les disciplines, la stylistique assistée par ordinateur apparaît comme un domaine de découverte qui est très exploré par les mathématiciens et par les statisticiens, et dont les résultats ne sont guère exploités par les littéraires.

La réflexion critique, épistémologique et méthodologique, gagnerait sans doute à se renouveler. A chaque fois en effet qu'un texte quelconque est enregistré sur un support technologique nouveau, informatique, optique ou électronique, c'est un nouvel objet « littéraire » qui est fabriqué, en fonctions de choix, de considérations techniques et de modes de découpage qui n'ont souvent aucune relation avec les préoccupations des littéraires, voire avec les exigences les plus élémentaires de la rigueur intellectuelle. Que penser par exemple de ces éditions ou de ses traductions de textes littéraires qui circulent sur « Internet » ou sur disquettes, sans indication sur l'édition manuscrite ou imprimée dont elles seraient la transcription, sans assurance que l'intégralité du texte ait été respectée, et dont la présentation visuelle et typographique n'a plus aucun rapport avec les caractéristiques des éditions d'origine ? Ce sont des formes d'édition « sauvages » qui vont proliférer, en toutes les langues, avec l'extension des « autoroutes de l'information ». Ce sont aussi les fondements mêmes de la critique philologique et littéraire qui se trouvent brutalement remis en question.

Les domaines de découverte que l'informatique entrouvre à la recherche littéraire ne manquent pas, des plus exaltants aux plus périlleux. Ils ne manqueront pas non plus de se diversifier à partir de 1997, quand l'explosion attendue de l'édition électronique se produira.

¹⁶ Ibid, p. 211.

III. LES DOMAINES DE PROSPECTIVE

Aussi, pour autant que l'on puisse prendre appui sur une connaissance des prévisions industrielles et technologiques qui prédéterminent l'avenir de l'informatique, faut-il tenter de prévoir ce qu'en seront les futurs territoires dans les études littéraires. Ces domaines de prospective seront nombreux. On n'essaiera pas de les dénombrer tous. L'on se contentera d'en privilégier trois secteurs, en fonction des évolutions qui se dessinaient dès 1995 dans l'édition savante, dans la traduction littéraire et enfin dans la littérature générée par ordinateur.

L'édition savante, à des fins de connaissance érudite, pourrait se trouver renouvelée, à proportion même de l'ampleur que l'édition « sauvage » des textes serait susceptible d'atteindre. L'édition critique se sert déjà depuis longtemps de l'informatique dans l'établissement des textes de manuscrits anciens ou nouveaux. Dès 1949, Don Roberto Busa s'était servi de calculateurs pour établir les premiers index informatisés de la *Somme théologique* de Saint Thomas d'Aquin. Un disque compact, *Perseus*¹⁷, propose ainsi, sur un seul disque, la quasi totalité des textes de la littérature grecque ancienne, du VII^e siècle avant J.-C. au V^e siècle après J.-C. D'autres projets d'édition analogues, entrepris dès 1995 ou même auparavant, portaient de la même manière sur l'édition savante de tous les manuscrits lyriques en langue occitane du XII^e siècle, sur la littérature du XVI^e siècle, ou sur des essais d'édition des oeuvres complètes des plus grands auteurs connus, de Dante à Goethe ou Shakespeare. Le disque compact de M.-L. Demonet et d'E. Brunet sur les *Electro-Chroniques sur François Rabelais*¹⁸, diffusé en 1995 par Les Temps qui courent, montre ce qu'il pourrait en être. Ces éditions devraient devenir multilingues et multimédias. Les disques produits en France par les éditions Arborescence sur *Matisse, Aragon, Prokofiev*, puis sur *Gauguin, Baudelaire, Tchaikovski* et enfin sur *Monet, Verlaine, Debussy*¹⁹ en sont une préfiguration. Une autre voie, celle des éditions multilingues, était déjà explorée en Italie, dès 1995, par les éditions E.M.M.E. avec un premier exemple d'édition synoptique des *Evangelies*²⁰ en italien, en français, en anglais, en allemand, en portugais, en espagnol et en latin. De ce point de

¹⁷ *Perseus*, Cambridge (MA), Harvard University, 1992, 1 cd-Rom.

¹⁸ Demontet (Marie-Luce), Brunet (Etienne) : *Electro-chroniques sur François Rabelais*, Paris, Les Temps qui Courent, 1995, 1 cd-Rom.

¹⁹ Casaril (Guy) : *Gauguin, Baudelaire, Tchaikovski - Monet, Verlaine, Debussy - Matisse, Aragon, Prokofiev*, Paris, Arborescence, 1995, 3 cd-Rom.

²⁰ Carraro (G & R) : *Les Evangelies*, Rome, Editel, 1994, 1 cd-Rom.

vue, les perspectives de l'édition électronique littéraire paraissent illimitées.

La traduction littéraire et les ordinateurs pourraient devenir un autre lieu de réflexion prospective et technologique. Le processus était déjà amorcé avant 1995. Il existe des pratiques inavouées qui consistent à nier ou à passer sous silence l'aide, quelquefois considérable, que le recours aux correcteurs morphologiques, orthographiques, syntaxiques, sémantiques et surtout stylistiques apportent déjà à l'activité de traduction la plus traditionnelle, aussitôt que l'on se sert de certains systèmes de traitement de textes. Le procédé le plus commun consiste à analyser d'une manière empirique les caractéristiques du style d'un texte dans sa langue d'origine, à l'aide d'un correcteur stylistique, pour en définir un « spectre » stylistique que le texte traduit devra s'efforcer de transposer ou de restituer. Les recherches menées en informatique sur les accès multilingues aux banques de données, les réflexions également des linguistes, des traductologues et des informaticiens sur la conception d'aides terminologiques appropriées à la traduction des oeuvres d'imagination, qu'elles soient populaires ou élaborées, pourraient susciter des recherches nouvelles, dont la traduction en général pourrait tirer un parti.

Un autre domaine d'avenir, exploré pourtant en Europe depuis 1959, c'est celui de la littérature générée, directement ou indirectement, par des ordinateurs ²¹. P. Valéry en avait pressenti l'avènement dès 1926. R. Queneau puis J. Roubaud en France, I. Calvino et U. Eco en Italie, H. Mathews en Grande-Bretagne l'ont exploré depuis les années 1959-1960. Les universitaires en ont découvert l'existence entre 1990 et 1995. Or, dès avant 1995, la poésie et la création poétique avec les revues *Kaos* et *alire* en France étaient déjà devenues l'un des domaines les plus riches en tentatives, en réflexions et en expérimentations de l'Allemagne au Brésil en passant par la France, l'Italie, le Portugal, la Russie, la Hongrie et bien d'autres pays. Le roman, depuis les années 1970, a subi une influence indirecte de l'informatique, qui n'est pas étrangère par exemple à l'inspiration d'auteurs aussi différents que G. Perec, Ph. Sollers ou I. Calvino. Le théâtre, avec la pièce en anglais d'A. Ayckbourn, *Intimate Exchanges* ²², adaptée en français au cinéma en 1994 par A. Resnais dans son film intitulé *Smoking/No Smoking*, est un autre exemple de la façon

²¹ Voir Vuillemin (Alain), Lenoble (Michel) et alii : *Littérature et informatique : la littérature générée par ordinateur*, Arras, Artois Presses Université, 1995 et Paris, Aupelf-Uref, collection « Presses Universitaires Electroniques Francophones », 1995 (sur Internet et Refer : <http://www.refer.qc.ca/AUPELF.html>).

²² Ayckbourn (Alain) : *Intimate Exchanges, A Play*, London, French, 1985, 2 vol.

dont des notions de littérature « permutationnelle » ou « combinatoire », empruntées à l'informatique par l'Oulipo, ont pu influencer sur les conceptions que des écrivains reconnus ont pu se faire de l'acte même d'écrire. Ce sont de nouvelles formes d'expression littéraires inédites ²³, encore confinées en des cercles restreints, qui surgissent, et qui pourraient contribuer à transformer jusqu'à l'idée que l'on se faisait jusqu'à cette date, en cette fin du XX^e siècle, de la « littérature »...

Peut-on vraiment conclure sur ce que l'utilisation de l'informatique a déjà commencé à apporter aux études littéraires ? Son emploi certes demeurerait encore restreint en 1995, en France du moins, même si de premiers groupes de chercheurs et si de premières universités françaises avaient commencé à s'y impliquer. Il en existait déjà, avant 1995, des domaines d'application reconnus, la recherche bibliographique, les publications savantes, les échanges internationaux, au moins entre des spécialistes. Il en était aussi, en 1995, des domaines qui relevaient encore de la découverte, l'édition électronique, l'analyse des textes, la réflexion critique. Il en était enfin, toujours en 1995, des domaines de pure prospective ou presque, de l'édition savante ou à la traduction littéraire et jusqu'aux perspectives infinies des nouvelles formes d'expression de la littérature générée par ordinateur. Ce que l'informatique apportera aux études de lettres sera multiforme. Ce qu'en sera la « qualité » littéraire, intellectuelle et scientifique, est une autre question. L'impréparation des milieux littéraires y était manifeste. De ce point de vue, l'intégration de l'informatique et des nouvelles technologies de l'information dans les études littéraires ne se fera pas sans mal ni crise grave.

Alain VUILLEMIN

Professeur à l'Université d'Artois

Une première version de cet article est parue sous le titre « Informatique et Littérature » dans la lettre de l'AILF (Association des Informaticiens de Langue Française) n° 17 de janvier 1996.

²³ Pour en savoir plus et pour identifier les sites littéraires francophones nouveaux qui se créent sur « Internet », consulter : <http://www.swarthmore.edu/humanities/clicnet/litterature.0.html>