

Numérisation et nouvelles pratiques de lecture

par Yannick Maignien et Jean-Didier Wagneur

Responsable et adjoint de la mission scientifique de numérisation à la Bibliothèque nationale de France.

Lecteur, il est temps que cette navigation agitée trouve enfin un point où aborder. Est-il un port mieux fait pour t'accueillir qu'une grande bibliothèque ?

(*Si par une nuit d'hiver un voyageur.* / Italo CALVINO – Seuil, 1981, p. 271.)

Les fonds électroniques, les postes de lecture assistée par ordinateur, les bibliothèques digitales, la délocalisation des sources sur les autoroutes de l'information soulèvent des questions dont certaines gravitent très souvent autour de la mort annoncée du livre et des bouleversements qu'elle présuppose dans nos pratiques traditionnelles de lecture. Qu'en est-il réellement ? Que peuvent attendre les théories de la lecture des nouveaux supports qui traitent non seulement le texte mais aussi le son et les images à travers les hypermédias ? Quelles sont ses perspectives d'avenir ? Telles sont les questions ouvertes par ce que l'on nomme « la révolution numérique ».

L'idée n'est pas neuve...

Avant d'aborder les pratiques liées aux nouveaux supports, il est nécessaire de revenir sur leur histoire. L'idée de fonds électronique n'est pas neuve. Dès les années 1950-1960, l'Institut national de la langue française commence à saisir sur cartes perforées une base de données qui deviendra *Frantext*. Environ 3 600 œuvres destinées à constituer les entrées et les exemples de son dictionnaire : le *Trésor de la langue française*. A la suite de cette expérience, d'autres laboratoires ont imité cette initiative en constituant des banques de données textuelles et en dé-

veloppant des logiciels pour répondre à des besoins de type statistique et lexicographique. A cette époque, on ne parlait pas encore de numérique, ni même de réseaux et ces projets en laissaient plus d'un sceptique.

C'est Théodore Nelson qui va apporter l'élément moteur en introduisant un nouveau concept : l'hypertexte. A la suite des travaux de Vannevar Bush, Théodore Nelson imagine en 1965 une configuration globale de lecture assistée où un lecteur accède par réseau à la totalité des informations disponibles, les archive dans sa bibliothèque personnelle, les annote, les relie les unes aux autres. Ce dispositif qu'il baptise Xanadu, est proprement un coup de Trafalgar dans l'univers de l'informatique, et l'hypertexte se constitue vite en paradigme absolu, quasiment une utopie aux dimensions mondiales dont Théodore Nelson poursuit actuellement la formalisation et la réalisation sur le réseau Internet.

Nos dispositifs modernes de lecture/écriture assistée par ordinateur sont les héritiers de cette histoire. Ils sont nés de la conjonction entre fonds numériques et un bureau électronique chargé de faciliter les échanges et les opérations. Favorisés par l'apparition des micro-ordinateurs, la vitesse des processeurs et les interfaces graphiques, les dispositifs de LAO permettent de passer du simple texte numérisé sur lequel les opérations restaient jusqu'alors assez limitées, à la constitution de

véritables bibliothèques textuelles dotées d'outils de structuration, de surlignage, d'annotation et de liens hypertextuels. C'est le passage du local au global, vers le réseau planétaire et le travail collaboratif avec, notamment, les expériences initiées par Douglas Engelbart, précurseur de ce que l'on nomme le *groupware*.

Lecture sur écran

C'est dans ce contexte qu'ont été développés des dispositifs de lecture assistée comme celui de la Bibliothèque nationale de France et des hypertextes comme Hypercard, Hyperties, Intermedia, Arachnid, Storyspace, DynaText... Mais toutes ces configurations – dont certaines sont étonnantes au plan technologique – reposent implicitement sur une philosophie de la lecture et sur des usages historiquement éprouvés. Ainsi, lorsque Vannevar Bush donne pour titre à l'article qui décrit son projet Memex, "*As you may think*", il l'inscrit dans une approche cognitive, celle de nos associations mentales, opération triviale dans tout processus de lecture. « Ne vous est-il jamais arrivé », écrit Roland Barthes, « lisant un livre, de vous arrêter sans cesse dans votre lecture, non par désintérêt, mais au contraire par afflux d'idées, d'excitations, d'associations¹ ? » Lorsque les hypertextes automatisent des liens permettant une navigation dans des hyperdocuments, ils reprennent des procédés mis au point dès le Moyen-Âge au travers des tables de concordances et des index. Les fonctions d'annotation, enfin, participent éminemment des pratiques que tout lecteur utilise quotidiennement dans la lecture de l'imprimé. La dimension réticulaire qui gouverne l'ensemble simule les réseaux traditionnels des communautés savantes.

Est-ce à dire qu'il n'y a rien de nouveau et que la lecture d'un hypermédia ou d'un hypertexte est de même nature que celle d'un document imprimé ? Bien évidemment non. Tout « lecteur » de CD-Rom ou de fonds numérique en fait vite l'expérience. Et, pour reprendre une expression aujourd'hui courante : on ne « lit » plus exactement ces médias-là, on les « consulte ».

Lire un texte sur écran est plus fastidieux que le faire livre en main. Notre perception de l'espace du volume est alors totalement modifiée et notre expérience émotionnelle du texte qualitativement bien différente. S'il nous faudra encore

du temps pour « dévorer » un tel fichier, en revanche l'expression « se perdre dans un livre » reçoit ici une intonation nouvelle. Car, lorsqu'il est numérisé en mode image (bit-map, fac-similé), le livre est réduit à une pure représentation en deux dimensions. Le consulter oblige à allumer un micro-ordinateur, à ouvrir un fichier, à retrouver l'endroit où l'on s'était arrêté. Le feuilleter nécessite la médiation d'un clavier ou d'une souris. Le livre a perdu tout poids, toute densité, toute réalité palpable. Lorsqu'il est numérisé en mode texte, la perte est alors encore plus importante. Privé des attributs typographiques de l'original il n'existe plus que sous la forme d'un pur texte. C'est une idée platonicienne de livre, un ensemble d'octets stockés dans une mémoire de masse.

Dans *Les Usages de l'imprimé*² Roger Chartier recherchait « l'ensemble des gestes neufs secrétés par une production de l'écrit et de l'image en une forme nouvelle ». Ce programme pourrait être applicable aux nouveaux supports du texte. Mais quels seraient alors les gestes « neufs » ? Il est nécessaire ici de différencier méthodologiquement deux types de lecture. Tout d'abord, la lecture ordinaire, privée, pratiquée au quotidien pour la quête d'information et/ou le « plaisir du texte » et celle définie essentiellement comme une opération critique de lecture/écriture.

En ce qui concerne la lecture littéraire on comprendra vite que le livre, malgré tous les Cassandre, garde un avenir certain et que toute lecture passera d'abord par le livre tant qu'il reste le support premier du texte lu. Il en est de même pour la lecture dite linéaire ou séquentielle (bien que ce terme doive être amendé par les résultats des études sur la psychologie de la lecture). Car l'univers du numérique privilégie ce que l'on nomme le non-linéaire, où la lecture progresse par bonds entre des nœuds d'informations. L'exemple en est, sur le plan littéraire, les œuvres de création multimédias que des réseaux comme Internet diffusent dès aujourd'hui. Héritières pour la plupart d'une théorie de la littérature combinatoire (Raymond Queneau, Georges Perec), elles ont intégré dans leur écriture les ressources hypertextuelles et hypermédias, et annoncent dès à présent l'émergence de nouvelles formes d'expression que l'on pourrait qualifier globalement encore de *res literaria* jusqu'à ce qu'elles se sédimentent en genres. Comme dans la rhétorique classique,

ces œuvres ménagent une place au lecteur implicite. Il est promu pour l'occasion lecteur interactif, navigateur voire parfois scripteur. L'intégration de l'image et du son dans de telles productions implique en même temps qu'une nouvelle syntaxe entre ces composants (au sein desquels le primat culturel du texte semble soumis à une révision) de nouveaux parcours tandis que le concept de « montage » qui les organise offre des solutions inédites aux problèmes esthétiques de la représentation. Sur de telles œuvres, la lecture accomplit son essence proprement ludique, telle que l'a définie Michel Picard dans son excellent essai. « La lecture comme jeu » est portée même à la limite avec la dimension ludique induite par tout support virtuel.

L'ordinateur personnel, épicentre de la famille moderne

Les CD-Rom disponibles aujourd'hui s'inscrivent pour la plupart dans une ligne générale qui est celle de la bibliothèque privée des ménages. Les dictionnaires, les encyclopédies, les corpus à usage pédagogique, les multimédias culturels visent des pratiques de lecture bien définies, essentiellement la recherche et la gestion d'informations. Dans ces entreprises éditoriales, la vocation pédagogique est constamment affirmée. Les CD-Rom, en grande partie, visent l'enrichissement culturel car le problème n'est plus tant, comme hier, l'alphabétisation des individus mais l'accès à la connaissance. De la même manière que l'édition du XIX^e siècle l'a fait avec des publications populaires, les nouveaux supports visent un encyclopédisme à l'usage des familles en utilisant au mieux les ressources didactiques et interactives du média. C'est un peu un « magasin d'éducation et de récréation » électronique tel que l'auraient rêvé Hetzel et Jules Verne.

Des publicités, comme celle d'*Apple* sur les machines multimédias, font même de l'ordinateur personnel l'épicentre de la famille moderne. Le micro est capable de gérer le budget de la maison, l'accès aux réseaux, les CD-Rom, les CD-Vidéo, les activités ludiques, et la domotique est déjà prévue pour demain. Ce scénario préfigure le rôle joué par l'ordinateur à l'avenir. Vendu aujourd'hui avec son lecteur de CD-Rom – certains fabricants proposant même une bibliothèque électronique où figurent, notamment, les œuvres complètes

1. « Écrire la lecture » in *Le Bruissement de la langue*, Roland Barthes – p. 33, Seuil, 1984.

2. *Les Usages de l'imprimé*, Roger Chartier – Fayard, 1986.

de Shakespeare – l'ordinateur personnel devient la machine à lire et à mémoriser, le central de communication sur lequel chacun naviguera et stockera ses bases de connaissances. Une telle représentation préfigure quelque peu le portrait du lecteur du futur pris dans un réseau d'informations provenant de plusieurs sources, livres, CD, réseaux... Cela nécessitera de sa part l'alternance de plusieurs activités, lire, regarder, écouter et beaucoup de méthode pour naviguer, trouver et trier. En un mot il reste à construire un nouvel art de lire.

Si les CR-Rom et la littérature numérique tracent des chemins de lecture à travers des corpus et donc d'une certaine façon conditionnent l'appropriation du texte selon une cartographie définie, les machines à lire expérimentales offrent en revanche à leur utilisateur la possibilité de faire, dès à présent, de la navigation hauturière dans l'univers de l'information.

La lecture n'est pas seulement le parcours, l'appropriation d'unités documentaires constituées (le livre) mais l'entrée vers des ensembles métadocumentaires plus conséquents, de collections, de bibliothèques. « La littérature n'est pas seulement faite d'œuvres singulières, mais de bibliothèques, de systèmes dans lesquels les diverses époques et traditions organisent les textes *canoniques* et *apocryphes*. A l'intérieur de ces systèmes, chaque œuvre est différente de ce qu'elle serait si elle était isolée ou insérée dans une autre bibliothèque », nous dit Italo Calvino dans *La Machine littérature*³. Le projet de numérisation de la Bibliothèque nationale de France a fait ressortir avec évidence cette dimension. La lecture assistée par ordinateur, c'est l'accès, *via* un écran et une image en deux dimensions, à des masses sans égales d'informations. Mais c'est l'ensemble des fonctions bibliothéconomiques qui s'en trouvent affectées.

Le catalogue, d'une part nécessaire, s'avère cependant insuffisant pour constituer et présélectionner des extraits, des articles, des anthologies, des références formant la « bibliothèque personnelle » au sein de laquelle le lecteur va se mouvoir. Cette signalétique de départ est indispensable, mais insuffisante. Au-delà des notices, toute indication fondée sur des extraits significatifs, tables des matières, mais aussi *index rerum* ou *nominarum*, bibliographies, citations, quatrièmes de couverture, éventuellement résumés ou

préfaces, permettra une « lisibilité » du document et instruira une commande de sélection. L'utilisation du paratexte à des fins d'indications croisées doit permettre des recoupements qui « localisent » les sources autoriales.

Entre la notice informatisée et le document numérisé, notamment si celui-ci n'est qu'en mode image, ces « couches documentaires » permettront d'orienter la recherche de façon fructueuse, bien que non déterministe. Les « voisinages », les proximités intellectuelles seront ainsi repérés de proche en proche. On pourrait rajouter avec Calvino : « La bibliothèque idéale, vers laquelle je tends, pour ma part, est celle qui gravite vers le *dehors*, vers les livres *apocryphes*, au sens étymologique du mot, c'est-à-dire les livres *cachés*. La littérature est la recherche du livre caché au loin, qui modifiera la valeur des livres connus ; elle est la tension vers le nouveau texte apocryphe à découvrir, ou à inventer⁴. ».

Depuis Leibniz jusqu'à Borges, le livre est compris comme cet atome constitu-

tif d'une infinité de possibles que sont les bibliothèques. C'est cet infiniment grand qui est à lire, qui seul donne valeur aux « bons ouvrages rares » sur fond d'innombrables écrits sans valeur qu'il faut pourtant conserver.

La bibliothèque, ensemble de connaissances à réorganiser

La bibliothèque est d'abord cet ensemble encyclopédique de connaissances toujours à repenser, à réorganiser, à reclasser en fonction de l'émergence ou de la dérive de continents disciplinaires. La nouveauté est que l'informatique permet d'hypertextualiser ces différentes couches documentaires hiérarchisées que sont les rubriques formatées des notices, les indexations et les autorités, puis les extraits significatifs, puis les documents eux-mêmes.

On sait que c'est l'enjeu technique actuel que d'instaurer des continuités de traitement dans la lecture exploratrice de ces couches hiérarchisées d'information,

4. *Idem*. p. 53.

3. *La Machine littérature*, Italo Calvino – Seuil, 1984.

depuis le signalement jusqu'à la matière même du document. Enjeu complexe, puisque les formes ou structures éditoriales à prendre en compte sont rien moins que standardisées ; enjeu nécessaire cependant si l'on veut qu'un encyclopédisme renouvelé puisse s'appuyer sur la richesse du patrimoine écrit.

En effet les rubriques informatisées (par exemple en format MARC) peuvent être reprises en outil de saisie structurée SGML, ce langage de description devenant lui-même outil de gestion de bases de données, permettant de rechercher contextuellement d'après la structure logique de l'information. SGML permet enfin d'établir les liens entre documents, entre couches documentaires (par exemple noms propres cités et noms propres en bibliographie). Un tel métalangage de description des objets textuels permet aussi l'établissement de familles de classes logiques ouvrant sur la possibilité de liens et de « navigations ».

En parallèle, l'univers des services en ligne s'est de son côté rapproché de la notion de bibliothèque. Certes des serveurs comme W3 sont encore en attente de structurations normalisées, mais d'ores et déjà la « navigation », mot-clé remplaçant celui de lecture comme parcours, est la forme canonique de la lecture butinage, aléatoire et fructueuse, imprécise mais prometteuse.

Ce qui est en jeu dans cette mutation, c'est la déconnexion entre les structures d'accès que permet la numérisation et l'hypertexte d'une part, des structures éditoriales initiales que revêt encore l'écrit imprimé d'autre part. Les structures autoriales et éditoriales initiales risquent dans cette aventure de disparaître, ou en tout cas d'être réutilisées au profit de recompositions parfois productives, mais souvent hasardeuses, que permettent les reconstructions sur le terminal de travail ou de lecture qu'est l'ordinateur.

Le lecteur renvoyé à ses responsabilités

Mais comme l'écrit Barthes de façon prémonitoire : « La lecture : ce serait là où la structure s'affole⁵. »

Le lecteur est ainsi renvoyé dans des responsabilités que, certes, déjà la lecture de l'écrit relié lui avait consenties, mais qui là se révèlent constamment indispensables. « Le texte postule la coopération du lecteur comme condition d'actualisation »,

disait déjà Umberto Eco⁶, p. 65) pour désigner cette place *stratégique* que l'auteur fait consciemment ou inconsciemment jouer à son lecteur dans la littérature. Un merveilleux exemple est là encore illustré par Italo Calvino, dans *Si par une nuit d'hiver un voyageur...*, livre composé de débuts ou morceaux d'œuvres disparates, au gré de la quête initiatique du « lecteur », lui-même à la fois image spéculaire et connivence avec le vrai lecteur (vous ou moi) et véritable héros du livre. Les titres de chacune de ces œuvres fragmentaires ne forment-ils pas eux-mêmes un texte... ? Ces titres sont tous au catalogue de la bibliothèque, mais aucun, pour des raisons diverses n'est disponible... Ce roman de Calvino est un merveilleux traité sur le plaisir de lire sous toutes ses formes (y compris celles de la Lectrice⁷). C'est aussi le paradigme du livre composé de tous les livres ou fragments de livres, livre-bibliothèque, livre infini, tangentiel. (La lecture est comme l'analyse, interminable !) Le fonds numérisé et la lecture assistée par ordinateur reprennent de même ce double ou cette duplicité : lire, c'est cette façon plus autonome de parcourir les univers déconstruits et complexes de multiples livres fragmentaires et de textes pourtant homogènes⁸. Devant son ordinateur, le lecteur est implacablement contraint de construire son univers : sa « navigation » est hauturière ; il doit prendre les droites de hauteur qui le situent dans son activité lectrice pour savoir où il se dirige, et même d'où il vient tant la masse des espaces documentaires parcourus peut accabler sa mémoire. Cette lecture assistée, au sens où l'on parle d'une conduite assistée dans les univers de simulation ou de parcours dans des espaces virtuels sensibles, doit disposer des bons outils dynamiques de localisation, de cartographie documentaire ou encyclopédique pour cultiver le plaisir phénoménologique qu'est la lecture de l'écran.

Ces outils doivent pouvoir capter et capitaliser le profil du lecteur au fur et à mesure de son enquête assistée par ordinateur. En retour, ce « profil » doit piloter des recherches de voisinages sémantiques propres à alimenter la curiosité du

lecteur. Là encore, nous retrouvons des thématiques communes aux débats ouverts par les bibliothèques virtuelles et électroniques que sont les ressources distribuées sur Internet⁹.

L'intentionnalité que met le lecteur dans sa lecture, doit encore se doter des bons outils sur lesquels doit s'appuyer le désir personnel du parcours et de l'exploration, allant du retour motivé, à la balade rapide, du survol, à l'attention scrupuleuse et annotatrice. Les auteurs et les éditeurs ont pendant des siècles pris en charge ce guidage du lecteur dans le bon mode d'emploi du texte : œuvres picaresques, baroques, techniques, encyclopédiques (les recommandations de lecture de d'Alembert en préface de l'*Encyclopédie*), philosophiques (les instructions de Spinoza pour lire l'*Éthique* entre lemmes, scolies et théorèmes). Imperceptiblement, ces stratégies d'appropriation du texte, quelquefois déliées de leur stricte linéarité qu'impose l'imprimé, trouvent ici une assistance qui passe à discrétion du lecteur. Mais tout livre, dès lors qu'il est publié, ne suppose-t-il pas ce lecteur qu'il veut séduire, convaincre, instruire, conduire, ou s'attacher à une cause transcendante ? Mais lecteur, tu t'en laisses de moins en moins compter (conter ?). Tu revendiques ton émancipation. Toi aussi tu veux être maître du texte. N'est-ce d'ailleurs pas toi, Lecteur-Modèle, dont on scrute à l'envie la « demande » ? N'est-ce pas toi, lecteur supposé, qui tires le marché de l'édition et es le véritable auteur de toute cette production littéraire ? N'est-ce pas toi, nouveau copiste électronique, qui produis du texte avec du texte ?...

Risquons qu'un des plaisirs renouvelé attaché à cette assistance automatisée, en quoi il faut bien reconnaître une forme de lecture, consistera dans l'établissement nouveau d'ordres de lecture « inédits », se rajoutant aux ordres auctoriaux et éditoriaux initialement et matériellement préétablis.

Toujours supposé par l'Auteur, le Lecteur s'arroge pourtant sur celui-ci des pouvoirs qui excèdent les seules vertus de l'écriture. L'informatique vient-elle donner corps à ce nouvel imaginaire ? « La conclusion à laquelle je suis arrivé, c'est que la lecture constitue une opération sans objet ; ou qui n'a pas d'autre véritable objet qu'elle-même. Le livre est un support accessoire, ou même un prétexte¹⁰. »

6. Umberto Eco, *Lector in fabula* – Grasset et Fasquelle 1985, p. 65.

7. « LECTRICE, voici que tu es lue ». *Si par une nuit d'hiver un voyageur*, Italo Calvino – Seuil, 1981, p. 166.

8. Dans le roman de Calvino, l'homogène est du côté de la mythologie orale (le « Père des récits ») ; il peut être aussi du côté du vil traducteur fabricant d'apocryphes, Hermès Marana, représentant par ailleurs de l'OEPHLW (Organisation pour la production électronique d'œuvres littéraires homogénéisées)...

9. *Bulletin des bibliothèques de France*, n° 2, tome 40, 1995, « Bibliothèques virtuelles ».

10. Italo Calvino, *op. cit.* p. 273.

NDLR – Les intertitres sont de la rédaction.

5. Roland Barthes, *ibid.* p. 65.