Texte et ordinateur

Jacques Anis

▶ **To cite this version:**

Jacques Anis. Texte et ordinateur. Bulletin de l'EPI (Enseignement Public et Informatique), 1988, 49, pp.150-158. edutice-00000999

HAL Id: edutice-00000999
https://edutice.hal.science/edutice-00000999

Submitted on 21 Oct 2005

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

TEXTE ET ORDINATEUR

Jacques ANIS

Dans les applications de l'informatique que j'évoquerai aujourd'hui (1), le traitement de texte et le vidéotex, les textes sont traités dans leur matérialité, en tant que chaînes de caractères affectés de certains attributs graphiques, sans prendre en considération ni leur sens ni, en règle générale, leur structuration globale ; ils sont codés pour être manipulés sur le mode du calcul. Cependant, ces deux technologies modifient en profondeur la production et la réception de l'écrit (2).

CODAGES INFORMATIQUES ET SYSTEMES GRAPHIQUES

On pourrait définir le texte comme une suite de caractères alphabétiques ou alphagrammes, jalonnée de marqueurs syntagmatiques et énonciatifs ou topogrammes (signes de ponctuation et modificateurs graphiques) ; des caractères non alphabétiques ou logogrammes (signes conventionnels divers) peuvent facultativement se substituer à certaines séquences alphabétiques. Notons tout de suite l'hétérogénéité d'un système qui oscille entre un système graphique minimal, représenté par l'écriture manuelle courante, comprenant alphagrammes, signes de ponctuation, et un seul autre topogramme, le soulignement, les logogrammes étant peu employés et un système graphique maximal, représenté par l'imprimé, qui comporte un jeu complet de variantes comme l'italique, le gras, le souligné, les petites capitales, etc. Ces variantes s'intègrent à un

(1) Ceci est le texte d'un exposé présente le 28 novembre au siège de l'EPI dans le cadre du groupe "INFORMATIQUE ET LITTÉRATURE" dirigé par M. Barbot et B. Gicquel. Je les remercie, ainsi qu'A. Vuillemin, de m'avoir donné l'opportunité de cette communication.

(2) Je me permets de renvoyer le lecteur intéressé par ce bref exposé à *Linx* 17, coord. ANIS, Jacques et LEBRAVE, Jean-Louis, "LE TEXTE ET L'ORDINATEUR", C.R.L. de Paris X Nanterre, novembre 1987, notamment aux volets 1 et 2 consacrés respectivement au traitement de texte et au vidéotex (ce numéro de revue, compte-rendu d'une table ronde tenue en novembre 1986, peut être commandé directement, au C.R.L., Université de Paris X, bureau F 514, 200 avenue de la République, 92001 Nanterre Cédex). Sur la PAO, on peut signaler un ouvrage de référence : LA MICRO-ÉDITION SELON SEYBOLD, de J. Seybold et F. Dressler, Dunod (Bordas), 1987.

fonctionnement global complexe et riche, qui fait intervenir tous les niveaux, du caractère à l'ouvrage complet, en passant par la ligne et la page ; les paramètres sont liés aux types de texte et même aux textes particuliers. Dans le traitement informatique du texte, on pourrait considérer que système minimal et compléments sont dissociés, l'un relevant d'un code de base universaliste, les autres de codages très divers, selon les programmes et les matériels.

CODE ASCII

Code crée aux USA pour les Télécommunications (*American Standard Code for Information Interchange*), il a été adopté par tous les fabriquants de micro-ordinateurs ; dans la version originelle, un caractère était codé sur 7 bits et il y avait donc 2^7, soit 128 caractères disponibles, ce qui peut paraître beaucoup : c'est peu, si l'on considère que les 32 premiers codes sont réservés aux caractères de contrôle et ne sont donc pas des caractères au sens courant du terme et qu'il faut ajouter aux lettres les signes de ponctuation, les chiffres, les opérateurs arithmétiques, les logogrammes les plus courants ; ce code peut être adapté aux normes nationales et ainsi par exemple accepter les lettres portant l'accent aigu et l'accent grave et le ç : mais c'est au prix de la perte de signes utiles comme les crochets ou les accolades ; quant aux lettres portant l'accent circonflexe, elles ne peuvent être traitées qu'avec trois caractères (la lettre simple, l'accent et un code de recul). On peut s'étonner que la Direction Générale des Télécommunications, définissant la norme du vidéotex, se soit basée sur un système restreint, ce qui impose pour écrire un é, d'utiliser trois codes ! La micro-informatique a adopté un code élargi (8 bits) comportant 255 caractères ; avec l'inconvénient de contribuer au recul de la standardisation.

MISE EN FORME MATÉRIELLE

Jacques Virbel désigne sous ce terme l'ensemble des paramètres grapho-textuels qui affectent les chaînes de caractères : styles, polices de caractère, disposition, etc. La MFM est plus ou moins complexe et plus ou moins riche. Elle oscille entre les systèmes graphiques riches et pauvres. Le traitement de texte bas de gamme n'intègre qu'une partie des valeurs graphiques, il est comparable à la dactylographie. De même le vidéotex ; mais cette pauvreté est partiellement compensée par des traits visuels propres à l'écran : inversion vidéo, clignotement, couleurs ou niveau de

gris. Le traitement de texte haut de gamme, notamment l'édition de bureau (*desktop-publishing*) ou PAO (publication assistée par ordinateur), tend à se rapprocher de la complexité et des finesses de l'imprimé classique. Une des problématiques les plus cruciales en informatique est celle du mode de gestion de la MFM ; deux orientations ont possibles, l'insertion de codes d'impression dans le fichier texte, la visualisation sur l'écran des caractéristiques du futur texte imprimé, qui a reçu le nom de *wysiwyg* - "what you see is what you get" -, franciser en affichage conforme à l'impression ; si ce procédé à permis la PAO en donnant aux non professionnels la possibilité de gérer une mise en page, par exemple, il rencontre certaines limites, sur lesquelles nous reviendrons.

LE TRAITEMENT DE TEXTE ET LA PAO

La délinéarisation de la production du texte

Dans le mode traditionnel de production et de publication du texte, les étapes sont assez bien définies : plan et/ou notes, brouillon, mise au net (manuscrite ou dactylographiée), épreuves, imprimé. Si le scripteur ne parcourt pas obligatoirement ces étapes dans un ordre absolument linéaire, il existe des clivages marqués entre ces différents types de documents, qui sont produits sur des supports et par des instruments d'écriture spécifiques. D'autre part, un document écrit sur papier est figé, en raison des difficultés d'effacement et de la place limitée. Le texte informatisé, quant à lui, peut évoluer quasiment sans rupture de l'état de plan (sous certaines réserves, en tout de l'étape du brouillon) à l'état de texte mis au net ou même de texte imprimé. Le document, toujours propre, à l'écran comme sur les sorties papier, reçoit ajout (mode standard dit insertion), suppressions, substitutions. Le scripteur n'est pas astreint à une rédaction séquentielle ; si l'on prend l'exemple d'un article scientifique, il est possible de circuler entre le début et la fin, entre le texte principal et les annexes telles que les notes et la bibliographie, et même, si l'on dispose d'un micro-ordinateur avec une mémoire vive de l'ordre d'un Mo et d'un disque dure, entre le texte lui-même et des données graphiques ou numériques relevant d'autres logiciels.Certaines fonctions évoluées du traitement de texte peuvent aider le scripteur à porter un regard critique sur son texte : recherche ou, mieux encore, indexation signalent les répétitions ; la présentation sous forme de plan et la table des matières fournissent un point de vue

synthétique ; enfin les correcteurs orthographiques, encore limités quantitativement et qualitativement, car ce ne sont que des dictionnaires de formes, peuvent aider à éliminer une partie des fautes d'orthographe et beaucoup de fautes de frappe.

Le travail sur l'objet grapho-textuel

La philosophie qui inspire *wysiwyg* et PAO et celle du *Do it yourself* ; l'auteur, autrefois coupé de la fabrication matérielle, peut y participer, dans des proportions qui peuvent varier considérablement, il est vrai. Participation minimale : remise à l'imprimeur d'un document papier propre, où peuvent figurer des attributs graphiques déjà visualisés, tels que l'italique, le gras, les titres et sous-titres ; participation maximale : remise à l'imprimeur du document définitif, qui sera reproduit en offset ; participation moyenne : remise d'un texte sur disquette, qui pourra être retravaillé soit en PAO, soit en photocomposition classique. En tout état de cause, que la mise en forme élaborée par l'auteur soit remaniée plus ou moins profondément, il me semble que la disponibilité demarques graphiques raffinées, singulièrement si celles-ci sont visualisées à l'écran, peut modifier l'écriture elle-même ; le texte devient un objet plus concret, la production se rapproche de la réception. L'avant-texte, déjà matériellement lisible, donne déjà une anticipation du texte ; nous rejoignons ici la délinéarisation évoquée tout-à-l'heure. Cependant, le *wysiwyg* a ses limites, elles sont liées notamment à la taille et à la définition des écrans : certes les écrans format A4 ou plus sont promis à une large diffusion, certes les écrans même cathodiques et les cartes graphiques progresseront en définition ; il n'est reste pas moins vrai qu'écran et papier sont des espaces graphiques spécifiques. D'autre part, il n'est pas sûr que la confusion des rôles soit souhaitable ; l'imprimerie est un métier et si certaines règles traditionnelles peuvent sembler désuètes, la plus grande partie d'entre elles correspond à des impératifs de lisibilité, à des équilibres esthétiques, à des acquis culturels ; un logiciel de mise en page et une imprimante à laser ne donneront pas au non spécialiste la compétence typographique.

Le Vidéotext

Ce terme désigne la vidéographie grand public supportée par le réseau téléphonique, que l'on qualifie d'interactive, car l'usager peut non seulement recevoir mais aussi émettre ; cette caractéristique est fondamentale, elle faut du Minitel un média pluri-fonctionnel et multi-

directionnel. Il faut insister en second lieu sur le support-écran, pauvre par ses dimensions et sa définition médiocre, mais riche par l'espace-temps qu'il offre.

Types de communication

Le nombre de services sur Télétel augmente à un rythme effrayant, même si on se limite aux services grand public. J'esquisserai une typologie fondée sur deux critères, le sens de circulation des messages et la finalité de la communication.

On peut donc dans un premier temps distinguer :

- orientation descendante : dirigée de l'organe de presse, de l'institution ou de la société industrielle et commerciale vers le public ; un journal télématique, l'auto-présentation d'une université.
- orientation ascendante : dirigée du public vers l'organe de diffusion ; sondage par Minitel, questions qui seront posées à un homme politique dan une émission radiophonique ou télévisuelle.
- orientation horizontale : le service sert de véhicule aux échanges entre utilisateurs ; petites annonces, et surtout messageries (en direct ou en différé) ; les échanges peuvent être individuels, collectifs ou mixtes (ainsi une petite annonce comporte un message adressé à tous pour obtenir des réponses individuelles).

Sur un autre plan, on pourrait distinguer les fonctions sociales :

- information : générale ou spécialisée, étatique ou privée, etc ;
- action : sous ce terme, on pourrait désigner des utilisations pratiques, relevant notamment du commerce, telles que la commande à distance ou la réservation de places de train ou d'avion.
- contact : messageries, petites annonces relationnelles ;
- jeu : des jeux variés sont proposés, de chiffres, de lettres, de connaissances culturelles, d'aventures, etc.

L'analyse concrète d'un service fera apparaître évidemment l'imbrication des fonctions : l'information débouche sur l'action, quand elle se fait publicitaire ou promotionnelle ; la messagerie associe contact et jeu, sur les fameuses messageries roses.

Mise en page/Mise en scène

Le principe général étant le même pour toutes les applications – un écran de dimensions limitées va accueillir quelques instants des données textuelles et graphiques –, l'exploration de l'espace-temps peut varier considérablement.

Parlons d'abord du journal télématique. Jean-Yves Michalik a proposé la notion de *mise en scène*, estimant que le concept de *page-écran* n'était qu'une transposition du texte-papier, lié aux origines télévisuelles non interactives du système. De fait, les premiers observateurs insistèrent sur l'atomisation du texte, ainsi suacussonné en petites unités irrémédiablement isolées par le mode de consultation ; de fait, assez vite, on mit en place des éléments stables afin d'assurer la continuité de la lecture ; l'aboutissement est une composition par zones, chacune évoluant à son propre rythme. Le flash AFP, mis en place d'abord sur *Libération* télématique puis sur d'autres organes, offre ainsi des écrans construits sur un sous-sommaire et une zone où s'affichent les dépêches sélectionnées.

Le modèle informatique de la banque de données est également utilisé pour certains services d'informations : la plus importante de ces banques de données est l'annuaire ; d'autres proposent listes de restaurants ou programmes de cinéma. Le principe général est la sélection d'un ou plusieurs éléments correspondant aux critères indiquées par l'usager. Une banque de données est comparable à un annuaire ou répertoire, mais le consultant n'a plus accès à l'ensemble d'un corpus qu'il dépouillerait lui-même ; l'efficacité du support informatique face au support papier dépend de l'étendue du corpus, de la quantité et de la qualité des critères proposés, du traitement des indications du consultant (tolérances orthographiques, élimination des éléments inutiles, repérage des mots clés), de l'habileté de celui-ci. Ainsi l'annuaire par noms permet à l'utilisateur d'accéder à la liste des abonnés de tout le territoire et en règle générale de trouvez assez vite ce qu'il cherche (donc ne pas dépasser les trois minutes gratuites). En revanche, la consultation sur Minitel de l'*Officiel des Spectacles* ou de *Pariscope* coûtera beaucoup plus de temps et d'argent et sera moins efficace que celle des magazines imprimés.

Les messageries gèrent de manière assez complexe l'espace et le temps, car l'écran est tour à tour espace de lecture et espace d'écriture. Une messagerie individuelle en direct comme *Aline* du *Nouvel Observateur* offre une page d'accueil où s'affiche la liste des personnes

connectées, alternant avec une page de communication partagée entre le rappel du dernier message, le message du partenaire, et une zone disponible pour écrire la réponse. Un salon comme celui d'*Antigel* partage l'écran en bandeaux de trois lignes correspondant à chaque participant ; il est significatif que l'écran soit assimilé fictivement au lieu salon et que les participants puissent se désigner par l'emplacement de leur zone d'écriture ("le 5" ou "toi là-haut"). L'écran est en effet non seulement un espace de lecture-écriture mais aussi un espace projectif et fantasmatique coupé du réel, où s'agitent des êtres purement verbaux et qui peut sombrer à tout moment dans le néant, pour peu qu'un flux d'électrons soit coupé.

Texte et ordinateur

Revenant à une perspective plus générale, je voudrais conclure en évoquant deux thématiques qui me paraissent se dégager des divers modes de traitement informatique du texte (j'envisage ici aussi le vidéodisque ou le CD-ROM, banque de données sur disque optique).

Le texte et l'image

Quand l'imprimerie a succédé au manuscrit médiéval, image et texte ont été dissociés ; l'image faisait appel à différents procédés, tels que la lithographie ou la photographie. La numérisation homogénéise à nouveau le traitement des mots et des images ; le traitement graphique du texte permet d'améliorer la qualité de la page imprimée ; les pictogrammes reviennent en force dans le monde de l'écrit ; les images mosaïques, construites comme des combinaisons de caractères, balisent le monde du vidéotex.

Le texte multiple

On connaît les textes modulaires, construits sur l'association d'un texte de base et d'éléments variants, utilisés pour la promotion et la publicité, mais qui pourraient aussi être exploités esthétiquement ; on connaît aussi les générateurs de textes littéraires mis au point par l'ALAMO et J-P. Balpe, dont la productivité exponentielle donne le vertige. Moins spéctaculairement sans doute, la délinéarisation que nous avons évoquée tout-à-l'heure débouche sur un processus ininterrompu d'écriture : on peut multiplier les versions, on peut construire un patchwork textuel alimenté à diverses sources. La page-écran crée une combinaison unique et éphémère de textes que rassemble l'intention

momentanée d'un individu ou de plusieurs. Au texte d'encre et de papier, fixe et délimité de l'imprimerie se substitue le texte mobile et non borné d'électrons et d'écran.

J. ANIS
(Université Paris X Nanterre)