



HAL
open science

Lecture, compréhension et recherche d'informations dans les hypertextes.

Jean-François Rouet

► **To cite this version:**

Jean-François Rouet. Lecture, compréhension et recherche d'informations dans les hypertextes.. 1998.
edutice-00000074

HAL Id: edutice-00000074

<https://edutice.hal.science/edutice-00000074v1>

Submitted on 20 Oct 2003

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.

LECTURE, COMPREHENSION ET RECHERCHE D'INFORMATIONS DANS LES HYPERTEXTES



*Jean-François Rouet
Chargé de recherche au CNRS
Université de Poitiers
Laboratoire Langage et Cognition*

Cet article reprend l'essentiel de la conférence donnée au Centre Régional de Documentation Pédagogique Poitou-Charentes en janvier 1998. Il est directement inspiré d'un article précédent, également diffusé par le CRDP

Titre courant: Lecture des hypertextes

INTRODUCTION

Ces dernières années ont vu l'essor rapide de logiciels informatiques permettant le stockage, la gestion et la consultation de grandes quantités d'informations documentaires: hypertextes, hypermédias, multimédias "fermés" (sur CD-ROM), "ouverts" (sur Internet) ou mixtes. Ces systèmes, que l'on peut regrouper sous l'expression de "technologies de l'information", permettent à l'utilisateur de sélectionner et consulter les informations utiles, et de "naviguer" dans des réseaux complexes par l'intermédiaire de "liens" entre différentes unités d'information.

L'utilité de ces systèmes dans des contextes de recherche d'information personnelle ou professionnelle est évident. La question est de savoir si ces systèmes ont également un potentiel en tant qu'outils pédagogiques, permettant de favoriser la construction par les enfants de solides compétences en matière de production, compréhension et recherche d'informations documentaires. Les acteurs du système éducatif ont dans leur immense majorité déjà répondu par l'affirmative, si l'on en juge par l'engouement que suscitent les PC multimédia et l'Internet dans les établissements scolaires; engouement accompagné et parfois précédé par de vastes (et coûteux) plans d'équipement.

On ne peut cependant émettre aucune hypothèse sérieuse sur le potentiel éducatif des technologies de l'information sans examiner la nature des activités éducatives induites par ces technologies. Ce qui compte en effet, c'est moins les propriétés techniques des outils proposés (par exemple, moteurs de recherche sophistiqués) que les processus cognitifs que leur utilisation va évoquer voire provoquer (ou pas) chez l'élève. Programme ambitieux, car la diversité des activités en jeu semble constituer un obstacle à toute tentative d'arriver à des conclusions globales en la matière.

Cependant, si l'on dépasse les différences superficielles entre différentes situations d'usage des TI à l'école, on constate rapidement que ces situations passent toutes par un même sous-ensemble d'activités cognitives fondamentales: lire, comprendre, sélectionner l'information, produire (des notes, des comptes-rendus...). Mesurer l'impact des TI sur les apprentissages, c'est chercher à comprendre comment ces activités sont impliquées et transformées par leur usage. Il faut notamment prendre en compte la nature précise des processus cognitifs en jeu, le rôle des connaissances initiales dont l'élève dispose, la façon dont ces connaissances sont investies dans l'activité de lecture, et les transformations (p.ex. acquisition de connaissances nouvelles) qui peuvent résulter de cette activité.

Dans cet article j'examinerai plus en détail l'un des processus impliqués dans l'usage des TI: la lecture-compréhension dans un contexte de recherche d'informations. J'essaierai de dégager les conséquences d'une transposition de l'activité de lecture-compréhension dans les systèmes hypertextes. Le savoir-lire élémentaire suffit-il pour maîtriser ces systèmes? Sinon, quelles sont les nouvelles compétences/connaissances requises? Par ailleurs, quels bénéfices pour l'élève peut-on attendre de l'utilisation de ces systèmes? Un accès plus facile et plus rapide aux informations utiles? Une meilleure compréhension? Une évolution positive des stratégies de lecture?

Dans une première partie, je vais présenter quelques notions relatives à la psychologie de la compréhension, afin de souligner la diversité et l'importance des connaissances initiales impliquées dans ce processus. Puis j'aborderai le problème de la lecture et de la recherche d'information dans les systèmes " non-linéaires ", en essayant de montrer l'importance d'une approche basée sur les capacités cognitives des utilisateurs.

1. APPROCHE PSYCHOLOGIQUE DE LA COMPREHENSION DE TEXTE

1.1. Quelques éléments de théorie de la compréhension

D'un point de vue très général, on peut considérer tout texte comme une hiérarchie d'unités sémantiques en relation: les mots véhiculent des significations qui peuvent être composées pour former des énoncés. Ceux-ci se composent à leur

tour pour former des paragraphes, et de cette composition naissent de nouvelles significations. Le même raisonnement s'applique aux entités plus complexes (textes et groupes de textes) sur lesquels reposent la plupart des activités de lecture "fonctionnelle".

Selon une théorie psychologique maintenant largement répandue, pour comprendre un texte, le lecteur doit en construire une représentation "cognitive", c'est à dire extraire les significations véhiculées à chaque niveau et les intégrer aux connaissances dont il dispose initialement. Les mécanismes de cette construction ont été décrits dans un article devenu classique de Walter Kintsch et Teun A. van Dijk (1978; traduction in Denhière, 1984). Kintsch & van Dijk distinguaient plus exactement deux niveaux de processus permettant la construction de représentations de textes:

- **les micro-processus**: il s'agit de l'ensemble des processus psycholinguistiques permettant l'identification des mots, la construction des propositions sémantiques (combinaisons de prédicats et arguments), et de leur organisation "locale" au sein des phrases. Le résultat de l'application de ces processus est une représentation "littérale" du contenu du texte, appelée "base de texte" dans la théorie de Kintsch et van Dijk.

- **les macro-processus**: il s'agit des processus permettant de sélectionner et d'organiser les informations importantes de la base de texte pour former une représentation structurée et stable à long terme. En effet, le lecteur conserve rarement en mémoire une trace "mot à mot" d'un texte. Au contraire, il tend à mémoriser les informations les plus centrales, les plus importantes, et/ou celles qui permettront le mieux d'atteindre l'objectif fixé initialement. Pour cela, il lui faut notamment supprimer les informations inutiles ou redondantes, résumer les groupes de propositions relatives au même thème, et créer des propositions permettant de mettre en relation celles du texte, quand cette mise en relation n'est pas explicite. Le résultat des macroprocessus est une représentation propositionnelle condensée et hiérarchisée des informations initiales du texte.

Cette théorie a fait l'objet de nombreuses validations expérimentales. Elle rend compte de façon satisfaisante de la façon dont le lecteur reconstruit les significations véhiculées explicitement par le texte. Mais rend-elle compte pour autant de la "compréhension"? Il faut pour cela expliquer comment se réalise l'intégration des informations du texte aux connaissances du lecteur: comment relier "ce qui est dit" dans un texte à ce que l'on connaît du monde en général?

Dans leur ouvrage de 1983, van Dijk et Kintsch ont tenté d'apporter une réponse à cette question en faisant appel à la notion de "modèle de situation". Selon cette nouvelle hypothèse, comprendre c'est construire non seulement une représentation "propositionnelle" du texte, mais encore une représentation plus globale de la "situation" décrite par le texte. Dans cette représentation "situationnelle" figurent non seulement les informations en provenance du texte, mais aussi les connaissances antérieures que le lecteur a activées et mises en relation avec le texte.

Une propriété importante du modèle de situation est que, contrairement aux représentations propositionnelles, il n'est pas nécessairement conforme à la structure énonciative du texte. Les informations peuvent être réorganisées en fonction de ce que le sujet sait de la situation de référence. Ainsi, on a pu montrer par des expériences que deux objets décrits comme "proches" spatialement au début d'un texte resteront étroitement associés dans la mémoire du lecteur même si l'un des deux seulement est évoqué dans la suite du texte.

Une entité particulière du système cognitif joue un très grand rôle dans le déroulement des processus de compréhension: il s'agit de la "mémoire de travail". Dans le contexte de la lecture-compréhension, la mémoire de travail peut-être conçue comme l'espace actif de la mémoire, où sont réalisés l'ensemble des traitements qui permettent la compréhension: activation du sens des mots, construction des propositions, macroprocessus, élaboration du modèle de situation. Or, on sait que la mémoire de travail est d'une capacité relativement limitée (Baddeley, 1993). En d'autres termes, le lecteur ne peut avoir activement à l'esprit à un instant donné qu'une petite partie des informations du texte. Pour lire et comprendre efficacement, il lui faut donc gérer correctement le " flot " d'informations: identifier les informations du texte, activer les connaissances correspondantes, sélectionner, hiérarchiser et mettre en relation les propositions, et construire peu à peu une image de la situation évoquée par le texte. Etre un lecteur performant, c'est aussi savoir conduire cette " gestion " temporelle des multiples traitements requis par la compréhension.

En résumé, la compréhension d'un texte résulte d'une interaction entre des informations externes, apportées notamment par le texte, et les connaissances initiales du sujet, par l'intermédiaire du "moteur" actif mais d'une capacité réduite qu'est la mémoire de travail. Je vais maintenant revenir plus en détail sur la nature des connaissances initiales impliquées dans la compréhension, car il s'agit d'un point central pour aborder le problème des hypertextes.

1.2. Le rôle de différents types de connaissances dans la compréhension.

La rapide synthèse présentée ci-dessus fait apparaître l'importance des connaissances initiales dont dispose le lecteur. L'expression de "connaissances initiales" évoque bien entendu tout d'abord les connaissances du référent, c'est à dire le thème ou la situation évoquée par le texte. Ainsi, pour comprendre un article de journal sur la situation en Tchétchénie début 1995, il faut avoir un minimum de connaissances initiales sur cette partie du monde, son

histoire, etc. Mais le lecteur utilise également d'autres formes de connaissances lorsqu'il aborde un texte: d'une part, des connaissances rhétoriques, c'est à dire relatives à l'organisation des textes et des discours; d'autre part, des connaissances stratégiques, relatives aux tâches ou objectifs pour lesquels on doit lire et comprendre. Ces connaissances indiquent au lecteur la façon optimale de traiter les informations en fonction des buts ou objectifs poursuivis.

Le fonctionnement des connaissances initiales chez les lecteurs en âge scolaire a fait l'objet de multiples études théoriques et expérimentales ces dernières années. Je n'en résumerai ci-dessous que quelques-unes, à titre d'illustration. Pour une introduction plus complète aux problèmes de compréhension vus sous l'angle de la psychologie cognitive, le lecteur pourra se reporter aux ouvrages de Denhière (1984); Fayol (1985); Fayol et coll. (1992).

• 1.2.1. Le rôle des connaissances référentielles

Chacun sait qu'il est plus facile de comprendre un texte lorsque l'on est déjà familiarisé avec son "réfèrent", c'est à dire le domaine de connaissances qu'il évoque. Cependant, au delà du simple constat, il est difficile de caractériser la nature et les modalités d'activation des connaissances initiales. Quelles sont les "unités" de connaissance? Quel est leur mode d'organisation en mémoire? Par quels procédés sont-elles activées et intégrées avec les informations d'un texte? Plusieurs études expérimentales tentent de répondre à ces questions. Par exemple, Freebody & Anderson (1983) ont demandé à des enfants de 11-12 ans de lire la description d'un jeu présenté en deux versions:

- ○ - Version **familière**: Dans la version familière, le texte décrit un jeu bien connu des enfants américains ("horseshoe").

- Version **peu familière**: La version peu familière décrit le même jeu mais cette fois sous un autre nom, et présenté comme étant d'origine indienne.

Après la lecture on demande aux élèves de rappeler les informations du texte, d'en faire une présentation résumée, et de vérifier une liste d'affirmations dérivées du texte. Les performances sont largement meilleures dans le groupe ayant lu la version familière. Cette influence de la familiarité thématique s'avère aussi grande que celle du niveau général en lecture des élèves et de la difficulté lexicale (que l'on a également fait varier de façon indépendante).

Cette expérience illustre bien que la compréhension s'opère en grande partie grâce à l'activation de "schémas" en partie pré-construits, et qui peuvent varier selon la population d'origine des lecteurs (niveau éducatif, culture, connaissances d'un domaine, etc.). Des expériences similaires ont montré que le fait de disposer de connaissances initiales élevées permet aux lecteurs de rappeler ou de résumer le texte d'une façon mieux structurée, et plus "conforme" au texte initial.

Enfin, d'autres études ont permis de montrer que la mise en oeuvre des connaissances n'est pas toujours automatique, mais peut être renforcée par des exercices appropriés. Ainsi Stahl et coll. (1989) ont démontré que si l'on communique aux élèves des informations d'arrière-plan avant la lecture d'un texte, le niveau de compréhension s'améliore. Cet effet est d'autant plus important que les informations communiquées sont centrales par rapport à la thématique du texte.

• 1.2.2. Le rôle des connaissances rhétoriques

Si le rôle des connaissances référentielles a été abondamment démontré et illustré dans les recherches psychologiques, le rôle d'autres formes de connaissances commence tout juste à être étudié. C'est le cas par exemple des connaissances "rhétoriques", que nous pouvons définir comme les connaissances relatives non pas au contenu mais à la forme des textes.

On sait que la plupart des textes possèdent des caractéristiques structurales bien définies. Tout d'abord, les idées ne s'expriment pas de façon désordonnée mais selon des schémas rhétoriques précis: On ne produit pas le même type de texte lorsque l'on *raconte* une histoire que lorsque l'on *défend* une théorie ou que l'on *décrit* un paysage. De plus, au sein d'un même "genre" rhétorique les textes se composent de figures typiques. Le texte "explicatif", par exemple pourra comporter des figures de type "liste", "comparaison", "problème-solution", etc.

Tout texte se présente donc comme une entité composée de parties qui remplissent des fonctions informatives distinctes. Ces parties sont le plus souvent signalées par des marques de structures, telles la division en paragraphes, sections, chapitres, qui va de pair avec la présence de titres, sous-titres, énoncés de transition, etc.

Le lecteur doit être capable de repérer ces structures et de les utiliser, en particulier lorsqu'il se trouve confronté à des textes longs, complexes, ou relatifs à un domaine peu familier. A l'inverse, l'absence d'une structure clairement identifiable peut être un handicap, comme l'illustre une expérience de Kintsch & Yarbrough (1982). Au départ, les auteurs veulent montrer qu'il existe bien plusieurs "niveaux" de compréhension d'un texte (voir ci-dessus, point 1.1.).

Ensuite ils veulent démontrer que le lecteur utilise des stratégies de compréhension basées sur ses connaissances de structures textuelles typiques.

Les auteurs préparent quatre textes représentatifs de schémas d'explication classiques (classification, illustration, comparaison, procédure). Par exemple, l'un des textes traite des *"formes de locomotion chez les primates"*.

Pour chacun de ces textes les auteurs préparent deux versions:

- **Versión conforme au schéma**: les informations sont présentées dans l'ordre le plus conforme au schéma rhétorique et avec les marqueurs linguistiques appropriés. Par exemple, le texte sur les formes de locomotion comporte un paragraphe introductif puis évoque les formes de locomotion selon des degrés d'évolution croissants.

- **Versión moins conforme**: cette version présente une structure et des marqueurs différents de la version conforme.

Les deux versions comportent par ailleurs les mêmes phrases.

Après avoir lu le texte les sujets (étudiants) réalisent deux épreuves:

- ○ - Test de clôture: on efface un mot sur cinq du texte et on demande au sujet de le rappeler.
- Questions de compréhension générale.

La version moins conforme entraîne des difficultés au niveau des questions de compréhension mais pas au niveau du test de clôture. Selon Kintsch et Yarbrough, cette "interaction" entre type de schéma rhétorique et type d'épreuve indique que les connaissances rhétoriques seraient utilisées lors de la construction du modèle de situation, mais pas au niveau des "microprocessus" (voir ci-dessus, section 1.1.).

• 1.2.3. Connaissances métacognitives ou "stratégiques"

Les situations dans lesquelles l'on peut être amené à lire ou consulter un texte sont extrêmement variées: compréhension des idées générales, recherche d'un détail particulier, comparaison de différentes notions, etc. L'activité de lecture est donc extrêmement flexible, et le "modèle général" évoqué plus haut ne rend compte au mieux que d'un "commun dénominateur" de toutes ces activités. En d'autres termes, même si l'on admet que la compréhension repose sur un ensemble fini de "processus cognitifs", il faut toutefois examiner comment ces processus sont adaptés aux exigences de telle ou telle situation de lecture particulière. Pour cela, le lecteur doit disposer de connaissances métacognitives ou "stratégiques", qui concernent l'activité de lecture en elle-même, ainsi que ses relations avec les contraintes du contexte.

Supposons par exemple un élève en train de lire un texte dans lequel il rencontre le mot "cimaise", dont le sens lui est inconnu. Si ce mot est indispensable à la compréhension du passage de texte, le lecteur va devoir mettre en oeuvre une recherche d'informations qui consistera à interrompre la lecture, prendre un dictionnaire, rechercher le mot "cimaise", comprendre la définition, reprendre la lecture et utiliser le sens du mot pour comprendre le passage.

Toutes ces opérations d'apparence anodine requièrent en fait une capacité à "réguler" sa propre activité qui est loin d'être acquise d'emblée au lecteur. Il faut notamment repérer la difficulté de compréhension (donc savoir si oui ou non on connaît le sens d'un mot), choisir une stratégie pour résoudre cette difficulté (ici, prendre un dictionnaire. NB: une autre stratégie pourrait être d'essayer de "deviner" le sens du mot à l'aide du contexte). Il faut également savoir "mettre en attente" le processus de compréhension du texte qui a été interrompu, puis reprendre ce processus une fois la définition obtenue. Sans compter les difficultés supplémentaires qui peuvent se poser, par exemple lorsque plusieurs définitions sont proposées, ou lorsque le mot ne peut être trouvé.

Comme suggéré plus haut, il serait erroné de penser que ces connaissances stratégiques se développent "automatiquement" avec l'acquisition du savoir-lire élémentaire. Une expérience de Williams, Taylor & Ganger (1981) est tout à fait illustratrice du décalage qui peut exister. Ces chercheurs ont demandé à des enfants de 9 et 11 ans ainsi qu'à des adultes de lire des passages descriptifs contenant des phrases "anormales" (sans relation au thème ou incohérentes). Les résultats indiquent que les enfants comme les adultes mettent plus de temps à lire les passages comportant de telles anomalies. Cependant, les adultes signalent une plus forte proportion d'anomalies que les enfants. Cette expérience suggère que si les jeunes lecteurs éprouvent une "gêne" à la lecture de phrases incohérentes, ils ne sont pas toujours en mesure d'identifier précisément la cause de leur problème.

Paris & coll. (1982) ont développé la notion de conscience en lecture (reading awareness) pour caractériser la connaissance dont l'élève dispose à propos de sa propre activité. Pour valider cette notion, ils ont étudié une population d'élèves dont les capacités de lecture ont tout d'abord été évaluées, ce qui a permis de les classer en "bon" et "moins bons" compreneurs. Après avoir demandé aux élèves de lire un texte comportant des mots difficiles, ils ont observé que les élèves de 9 ans "bons compreneurs" ont de meilleures stratégies de lecture (p.ex., utiliser un

dictionnaire, poser des questions sur le sens des mots). Les moins bons compreneurs posent également des questions, mais celles-ci concernent principalement la prononciation des mots. La moins bonne compréhension paraît donc liée à une mauvaise maîtrise des techniques permettant de comprendre.

Ces capacités stratégiques, encore appelées "métacognitives" (parce que relatives à l'activité propre de chacun) ont fait l'objet d'un intérêt croissant en psychologie, ces vingt dernières années. John Flavell (1981) a consacré toute une série de recherches à leur acquisition chez l'enfant. La question du développement métalinguistique est également traitée dans un ouvrage de Gombert (1990).

1.3. Résumé et conclusions

Pour évaluer le potentiel des systèmes hypertextes en éducation, il est nécessaire de partir d'une analyse détaillée des processus cognitifs qui sous-tendent l'activité de lecture. Dans cette partie nous avons montré que cette activité implique une hiérarchie complexe de processus, ainsi que plusieurs types de connaissances: référentielles, textuelles et stratégiques.

En quoi ces données peuvent-elles alimenter une réflexion sur les systèmes hypertextes? On a vu que les connaissances rhétoriques et stratégiques correspondent à des structures linguistiques conventionnelles, telles qu'elle se présentent dans une culture donnée, et telles que l'élève les appréhende dès son plus jeune âge. En tant que lecteurs "matures", nous exploitons donc de façon intensive -- quoique généralement automatique, sans nous en rendre compte -- une somme considérable de connaissances sur les formes textuelles typiques que nous avons acquises peu à peu, au fil de nos " expériences " personnelles de lecture et d'utilisation de textes.

Le texte n'est donc pas une entité abstraite, virtuelle, qui pourrait prendre toutes sortes de formes ou d'apparences au gré des éditeurs (ou des informaticiens). Au contraire, il s'agit d'un objet précisément défini, en référence à des normes culturelles qui ont été établies, renforcées et adaptées lentement, au fil de l'évolution des technologies, des usages et des besoins.

Or, le propre des nouvelles technologies de l'information, et en particulier des hypertextes, c'est de proposer de nouvelles formes, donc de nouvelles pratiques de lecture, dont le but serait d'enrichir les possibilités d'interaction entre le lecteur et le texte.

Mais, si l'objectif visé ne peut être que loué, il faut absolument s'interroger sur les "moyens" cognitifs nécessaires à une telle évolution. Tout comme le manque de connaissances appropriées entrave l'élève dans sa lecture du texte conventionnel, il risque de compromettre l'apport potentiel des hypertextes. Il faut donc étudier en premier lieu l'impact des hypertextes sur le processus de lecture, puis dégager les moyens de doter les lecteurs -- y compris les adultes -- d'une bonne maîtrise de ces formes de lecture. Alors seulement on peut espérer mettre au point, en relation avec des équipes de développeurs, des systèmes réellement efficaces pour l'éducation et la formation. Cette question sera reprise dans la deuxième partie.

2. LECTURE ET UTILISATION DES HYPERTEXTES

Dans cette seconde partie je voudrais aborder l'utilisation des hypertextes du point de vue des processus de compréhension mis en oeuvre par l'élève. Je me centrerai sur les problèmes d'orientation et les stratégies de recherche d'information et je présenterai quelques études expérimentales sur ces questions.

Comme nous l'avons vu dans la première partie, la question centrale qui se pose est la suivante: Quels savoirs, quels savoir-faire nouveaux sont et seront requis pour parvenir à une maîtrise correcte des nouveaux systèmes d'informations? Quelles sont les différences avec le processus de lecture conventionnelle, sur papier? Quelle formation doit-on envisager pour préparer les futurs utilisateurs à l'emploi de ces technologies?

Après avoir rapidement retracé l'évolution du concept d'hypertexte, je préciserai quels types de problèmes ces systèmes peuvent poser au plan du fonctionnement cognitif de l'utilisateur. Puis je présenterai quelques expériences visant à identifier précisément ces problèmes et à en tirer quelques implications pour la conception de systèmes mieux adaptés.

2.1. Aperçu de l'histoire des hypertextes

L'histoire du "concept" d'hypertexte et celle des premiers systèmes ont été retracées dans plusieurs ouvrages, et je ne vais donc pas en faire une nouvelle présentation détaillée ici (pour des définitions générales et des exemples, voir Balpe, 1990).

Il faut toutefois rappeler que l'idée de mettre au point une technologie permettant d'optimiser le classement et la récupération d'informations documentaires n'est pas nouvelle. On s'accorde pour la faire remonter à l'article fondateur de V. Bush (1945), intitulé "*As we may think*". Bush critiquait les systèmes d'information classiques basés sur l'indexation alphabétique. Partant du principe que la mémoire humaine fonctionne par associations, il appelait de ses

voeux un système de gestion de l'information où les informations "proches" par leur contenu, leur utilisation, etc. seraient reliées entre elles. Bush avait conçu un projet de système basé sur des microfiches, le "MemEx" (ou "MEMory EXpander"), qui ne vit cependant jamais le jour, à défaut d'une technologie adaptée.

Il revient à T. Nelson d'avoir proposé le premier le terme d'HYPERTEXTE pour caractériser la mise en relation associative d'éléments d'information au sein des systèmes documentaires.

- ○ *"Je propose le mot 'hypertexte' pour représenter un ensemble de documents écrits ou figuratifs, interconnectés de façon si complexe qu'il ne pourrait être représenté de façon satisfaisante sur du papier (...)" "Un tel système, correctement conçu et géré, aurait un potentiel important pour l'éducation, car il augmenterait le répertoire de choix de l'élève, son sentiment de liberté, sa motivation, ses capacités intellectuelles."*

On trouve dans cette citation les principes essentiels des hypertextes :

- d'une part, la notion de réseaux d'informations interconnectées de façon complexe, de telle sorte qu'il soit difficile de représenter les mêmes informations de façon linéaire, c'est-à-dire selon un ordre particulier.

- d'autre part, l'espoir que les hypertextes vont augmenter le potentiel d'apprentissage de l'élève. La présentation " non-linéaire " des informations est censée encourager une participation active de l'élève. De plus, si les informations sont organisées en fonction de leurs relations sémantiques, alors la " navigation " dans un tel environnement permettrait à l'élève d'accroître ses connaissances et sa capacité à se procurer les informations.

Il faudra toutefois attendre encore environ 15 ans pour qu'apparaissent les premiers dispositifs informatiques correspondant à la définition de Nelson. Depuis le milieu des années 1980, une véritable explosion s'est produite, qui a vu les hypertextes devenir le centre d'un nouveau courant de recherche et développement. Une particularité en est la pluridisciplinarité, phénomène tout à fait remarquable qui veut que soient associés à ces recherches non seulement des informaticiens, mais aussi des psychologues, linguistes, didacticiens, enseignants, etc.

En 1987 s'est tenue une première conférence internationale de recherche sur les hypertextes à Chapel Hill, aux Etats-Unis. Depuis, cette conférence se tient tous les ans, soit en Europe soit aux Etats-Unis. Bien qu'elle soit largement dominée par la présentation de techniques et de nouveaux produits (ce qui lui donne parfois l'air d'un "salon" plus que d'une conférence scientifique), elle reste toutefois un forum privilégié pour l'échange et la confrontation des idées et des expériences. Au plan français, il faut noter l'existence des journées "Hypermédias et Apprentissages" dont les deux premières éditions (en 1991 et 1993) ont donné lieu à des actes publiés (voir de la Passardière et Baron, 1991).

Mais peut-être n'est-il pas inutile de rappeler brièvement ce que recouvre actuellement le terme d'hypertexte, au sens le plus large.

Un hypertexte se compose tout d'abord d'une base de données, dont les unités d'informations sont des passages de texte. A cette base de données on associe un réseau sémantique, c'est à dire un système de liens qui mettent les unités textuelles en relation les unes avec les autres. Le système est enfin doté d'une interface permettant de créer et de modifier les données, de consulter des réseaux déjà existants, d'afficher sur écran des unités textuelles, et de naviguer d'une unité à l'autre par l'intermédiaire des liens.

Jusqu'à présent, les recherches dans le domaine des hypertextes ont été fortement teintées d'optimisme, voire d'une certaine exaltation à laquelle la nouveauté et la puissance (réelle ou fantasmée) des technologies utilisées n'est sans doute pas étrangère. On sait à l'heure actuelle faire beaucoup de choses avec les hypertextes, et l'on assiste à une prolifération de systèmes commerciaux de toutes sortes qui s'inspirent du principe d'hypertexte. Outre les systèmes hypertextes proprement dits, on voit également toutes sortes d'applications (bases de données ou logiciels de bureautique, par exemple) intégrer progressivement des "composants" hypertextes.

2.2. Les hypertextes: Une solution ou un problème pour le lecteur?

Dans ce foisonnement créatif qui rappelle d'autres engouements " technophiles ", il n'est guère que certains psychologues pour prétendre que le coeur du problème ne se situe pas au niveau du fonctionnement d'un système en lui-même (si nouveau et séduisant soit-il), mais bien au niveau du couple système-utilisateur, c'est à dire de la communication qui s'établit (ou pas) entre le lecteur et le système d'informations. Problème sous-estimé voire négligé par beaucoup d'équipes de conception mais qui revêt tout de même une certaine importance, si l'on considère que les systèmes en question sont promis -- selon leurs auteurs -- à un avenir brillant dans les milieux de l'éducation ou de la formation.

Or jusqu'à présent les concepteurs se sont rarement aventurés à mettre rigoureusement leurs prototypes à l'épreuve d'utilisateurs réels, ni même à s'interroger sur les théories cognitives, ergonomiques, didactiques - si tant est qu'elles existent - qui pourraient justifier l'aura de révolution médiatique qui entoure les hypertextes. Les quelques expériences

d'utilisation ont donné des résultats pour le moins mitigés. Bien souvent, les utilisateurs montrent une certaine propension à perdre leur chemin dans les réseaux d'informations; ils se disent désorientés, hésitent quant aux décisions à prendre, et tous comptent faits en viennent à regretter la structure et la sécurité inhérentes aux bons vieux textes imprimés (voir à ce sujet Charney, 1994; Rouet, 1992).

On se trouve donc devant un paradoxe qui veut qu'un mode de présentation en théorie plus "riche" semble parfois poser plus de problèmes qu'il n'en résout. Avant d'examiner ces problèmes de plus près, examinons ce que pourraient être des arguments "pro-" ou "anti-" hypertextes.

Du côté "pro-", l'hypertexte permet une présentation de l'information analogue au domaine de connaissances concerné. Imaginons par exemple la description topographique d'un musée avec ses entrées, ses étages, ses escaliers, ses salles, etc.. Une présentation sur papier impliquera un certain choix dans l'ordre de présentation, ce qui entraîne une nécessaire "déformation" liée à la linéarisation d'un référent au départ tridimensionnel. Au contraire, la présentation "hypertexte" permet de conserver les relations telles qu'elles se présentent naturellement. Ainsi, si la salle A ouvre sur deux salles adjacentes B et C ainsi que sur le couloir D, la description de A dans l'hypertexte contiendra trois "boutons" qui mèneront respectivement aux descriptions de B, C et D. C'est le lecteur, qui, parcourant l'hypertexte, construira son propre itinéraire selon ses besoins, ses préférences, le temps dont il dispose, etc., exactement comme s'il visitait "en direct" le musée.

Bien entendu, cette métaphore spatiale s'applique également à des hypertextes portant sur des concepts bien plus abstraits, dans lesquels les notions apparentées peuvent être reliées entre elles de façon non-linéaire. Le lecteur de l'hypertexte navigue donc directement dans un "réseau de connaissances", plutôt que d'avoir à le reconstruire d'après une structure textuelle linéaire, donc nécessairement déformée et contraignante.

Voilà, présenté de façon sommaire, le principal argument en faveur d'une présentation non-linéaire de l'information. Mais, comme l'a fort bien montré D. Charney (sous presse), cet argument se heurte à quelques obstacles: par exemple, comment fait-on pour s'orienter dans un système non-linéaire? Dans l'exemple ci-dessus, comment s'assurer que la *présentation* en réseau va permettre au lecteur de construire une *re-présentation* cohérente? Le livre imprimé, du fait même de son organisation linéaire, permet au lecteur de se "situer" en permanence. En revanche, dans l'hypertexte se pose en permanence le problème de savoir "où" l'on est dans le réseau.

Des avantages d'une représentation analogique ou des difficultés à gérer l'orientation/navigation dans un réseau, il reste à vérifier ce qui l'emporte. Ceci ne peut se faire que par un patient travail d'analyse théorique des mécanismes de sélection, de traitement et d'évaluation impliqués dans cette activité. Et encore des principes théoriques ne peuvent-ils jamais être acceptés tels quels, sous forme de postulats. Il faut au contraire vérifier par l'expérience les prédictions ou les conséquences de telle ou telle théorie.

2.3. Apprendre à naviguer dans les hypertextes

Dans le cadre d'une thèse de doctorat réalisée au Laboratoire "Langage et Communication" de l'Université de Poitiers (unité associée au CNRS), sous la direction de Stéphane Ehrlich, j'ai réalisé plusieurs expériences dont le principe consistait à demander à des élèves de niveaux scolaires variables mais sans expérience initiale d'utiliser de petits systèmes hypertextes expérimentaux dans le but d'accomplir diverses tâches de lecture-compréhension (Rouet, 1991).

Les principales questions de recherche étaient les suivantes:

- Les lecteurs en âge scolaire sont-ils capables de s'orienter dans un hypertexte? Quels sont les processus cognitifs nécessaires à cette orientation?
- La recherche d'informations requiert-elle des stratégies de lecture particulières? Quand et comment ces stratégies sont-elles acquises par les élèves?
- L'utilisation de documents multiples via un système hypertexte permet-elle l'acquisition de connaissances?

Je vais maintenant résumer deux des études réalisées dans le but d'apporter des débuts de réponse à ces questions. Un compte-rendu plus détaillé de ces expériences se trouve dans Rouet (1994).

2.3.1. un système hypertexte expérimental

Pour ces expériences un système expérimental a été construit et adapté pour le "Nanoréseau". Le système, très simple dans sa conception, permettait de présenter de courtes "pages" de texte reliées par des menus. La Figure 1 présente un exemple d'hypertexte simple sur la société néandertalienne, destiné à des élèves de 11 à 14 ans.

Insérer la Figure 1 ici

- **2.3.2. Prise en compte des relations et navigation dans l'hypertexte.**

L'hypothèse générale qui a guidé la première expérience était que l'orientation dans un hypertexte nécessite l'identification du thème de chaque passage, ainsi que des relations entre thèmes, telles qu'elles sont représentées par les liens de l'hypertexte. Ces opérations nécessitent la mise en oeuvre de connaissances rhétoriques dont on ignore si et quand elles sont acquises par les élèves. Cependant, compte tenu des connaissances actuelles sur le niveau moyen en lecture des élèves de 11 à 14 ans (voir la première partie), il est probable que les connaissances rhétoriques ne soient qu'incomplètement mises en place au niveau de la sixième.

Pour vérifier cette hypothèse, deux versions de chaque hypertexte ont été préparées. Dans la version avec "relations implicites", chaque passage se terminait par une question dont la réponse pouvait se trouver dans un passage relié. Dans la version explicite, l'identité de ce passage relié était présentée explicitement (Cf. Figure 1).

148 élèves ont participé à l'expérience. La moitié environ étaient des élèves de sixième (11-12 ans) et l'autre moitié des élèves de quatrième (13-14 ans). Pour tous les élèves, la consigne consistait à parcourir l'hypertexte en essayant de prendre en compte les relations entre thèmes. Une explication détaillée ainsi qu'un exemple précédaient l'expérience proprement dite.

L'expérience comportait deux séances espacées de quelques jours, afin d'étudier si la familiarisation avec le dispositif entraînerait une évolution des stratégies des élèves. A chaque séance, chaque élève étudiait un hypertexte dans l'une des deux versions proposées, puis devait en rédiger un court résumé et répondre à une série de questions.

On s'est penché sur l'influence des relations explicites sur le type de parcours réalisé par les élèves. L'hypothèse était que la présence de relations explicites entraînerait des parcours plus conformes à l'organisation de ces relations. En fait, il s'avère que lors de la première séance, les relations sont prises en compte dans seulement 50% des cas en moyenne. Elles n'influencent significativement que les parcours des élèves les plus âgés. Il faut attendre la seconde séance pour que la prise en compte des relations soit effective à environ 80%, mais toujours légèrement plus chez les élèves les plus âgés. Par ailleurs, la prise en compte des relations s'améliore également dans le cas où ces relations sont implicites, ceci aux deux niveaux scolaires.

Quelles conclusions peut-on tirer de ces observations? D'abord, il y a bien une évolution de la 6ème à la 4ème dans la capacité des élèves à prendre en compte des énoncés organisateurs insérés dans un texte explicatif. Ensuite, cette capacité peut évoluer rapidement, puisqu'une première familiarisation suffit à faire évoluer significativement la prise en compte des énoncés relationnels. Ainsi, la lecture de l'hypertexte pourrait peut-être faciliter la mise en place de stratégies de lecture basées sur la thématisation et la mise en relation. Mais ce n'est là qu'une hypothèse qui reste à vérifier par des expériences à plus long terme.

Quoi qu'il en soit, les réserves émises plus haut quant à la capacité d'utilisateurs sans expérience préalable à s'orienter dans un réseau d'information trouvent ici une première confirmation.

- **2.3.3. Recherche d'informations dans un hypertexte simple.**

Si la première expérience a bien fait apparaître la relation entre compréhension des énoncés et orientation dans l'hypertexte, elle n'est représentative que d'une situation de lecture particulière: la consultation exhaustive d'un réseau d'informations. Or la présentation en hypertexte peut se prêter à d'autres formes d'activités, notamment la recherche d'une information particulière dans un ensemble.

On peut une nouvelle fois se demander si les élèves possèdent les connaissances nécessaires à la conduite d'une telle activité. La recherche d'informations requiert en effet au moins trois étapes distinctes: L'élève doit pouvoir évaluer son besoin d'information, par rapport à la question ou problème initial. Il doit ensuite sélectionner une "catégorie" d'information parmi celles qui sont proposées dans l'hypertexte. Il doit enfin comprendre l'information figurant dans cette catégorie (c'est à dire le passage de texte), et procéder à une nouvelle évaluation: l'information acquise est elle suffisante ou doit-on en chercher d'autres? On commence à mieux connaître les propriétés de ce cycle évaluation/ sélection/ compréhension au plan des processus cognitifs (voir à ce sujet Rouet et Tricot, 1995). Cependant, sa mise en oeuvre par l'élève fait encore l'objet de nombreuses questions.

Dans une deuxième expérience, on a utilisé un système hypertexte similaire à celui présenté ci-dessus, mais en demandant cette fois aux élèves de rechercher les informations nécessaires pour répondre à des questions. Le but était de déterminer dans quelle mesure les élèves sont capables de prendre en compte la nature de l'objectif dans leurs sélections d'informations, et si la familiarisation avec le dispositif et la tâche s'accompagnerait d'une évolution des stratégies de recherche.

Les participants étaient 60 élèves, dont 25 en sixième et 35 en quatrième. Chaque élève participait à quatre séances d'expérimentation, au cours desquelles il devait utiliser un hypertexte pour répondre à une série de quatre questions sur des thèmes variés (Homme de Neandertal, atmosphère, etc.). La moitié des questions étaient "complexes", en ce qu'elles renvoyaient à plus d'une unité de l'hypertexte (par exemple, question portant sur une relation entre deux unités). L'autre moitié des questions ne portait que sur le contenu d'une seule unité. De plus certaines questions étaient "explicites" car leur formulation reprenait les titres des unités de l'hypertexte. L'autre moitié était "implicite", car il était nécessaire de produire une inférence pour identifier la ou les unités pertinentes.

Le principal résultat observé est que le temps de recherche des informations-cibles décroît fortement avec la familiarisation. Ceci ne semble pas dû à un désintérêt progressif des élèves pour la tâche, car le rapport qualité des réponses/temps de recherche tend à augmenter. Il y a donc acquisition de stratégies de sélection plus efficaces, et peut-être même un début d'automatisation de certaines composantes du processus de recherche. Mais l'évolution est également *qualitative*: alors qu'en début d'expérience les élèves utilisent des parcours similaires pour différents types de questions, on voit peu à peu une certaine différenciation se mettre en place. Les élèves passent relativement plus de temps à examiner les menus lorsque la question est implicite, ce qui peut indiquer qu'ils produisent des inférences plus élaborées (un compte-rendu plus complet des résultats se trouve dans Rouet, 1992).

Au total, cette deuxième expérience montre qu'une stratégie de recherche efficace n'est pas acquise d'emblée par les élèves, mais qu'elle peut se mettre progressivement en place avec un entraînement. D'autres études similaires, visant à analyser de façon plus détaillées les mécanismes cognitifs de la recherche d'informations sont actuellement en préparation.

CONCLUSIONS

La lecture et l'utilisation de l'information textuelle reposent sur des processus complexes et des connaissances de différents types. Certaines de ces connaissances, en particulier celles qui sont relatives aux types de textes et aux stratégies permettant d'en faire le meilleur usage ne sont acquises que tardivement, et par une partie seulement des lecteurs adultes.

L'application des nouvelles technologies de l'information à des activités éducatives pose le problème des savoir-faire nécessaires à leur utilisation optimale. Il est vrai qu'on peut attendre des systèmes hypertextes des avantages importants en termes de rapidité et de souplesse dans l'accès à l'information. Mais il faut pour cela que ces systèmes soient bien adaptés aux caractéristiques de leurs utilisateurs, en particulier pour ce qui concerne leur interface, c'est à dire les outils permettant de naviguer et de s'orienter dans l'hypertexte. Les expériences résumées dans la deuxième partie ont montré que la maîtrise de ces outils n'est acquise que progressivement. Il ne faut donc en aucun cas sous-estimer l'effort de formation qu'exige l'utilisation ces systèmes.

Il faut ensuite se demander pour quelles tâches les hypertextes peuvent être réellement bénéfiques par rapport aux systèmes traditionnels. C'est l'objectif d'une nouvelle série de recherches dans lesquelles nous tentons d'évaluer le potentiel des hypertextes pour la présentation de documents pédagogiques dans le domaine de l'histoire (Britt, Rouet et Perfetti, 1996). Ici encore, nous travaillons à partir d'un système très simple, permettant la consultation d'un ensemble de documents par un système à menus. On a demandé à des étudiants d'université d'utiliser le système pour étudier des documents relatifs à un problème de type "controverse historique". Il s'agit d'une situation pour laquelle l'utilisation de documents sur papier nécessite des stratégies élaborées et s'avère complexe pour la plupart des individus. Il y a donc un réel besoin de faciliter l'accès et le traitement de cette information. L'hypertexte permet de prendre connaissance de différentes explications et de comparer des interprétations parfois contradictoires par l'intermédiaire d'une interface d'emploi aisée. L'hypothèse étudiée ici est qu'une telle présentation pourrait faciliter l'intégration de l'information à partir de sources multiples.

Les premiers résultats ont montré que la présentation en hypertexte permet un accès facile à l'information dans le cas où la structure d'ensemble des documents est clairement représentée dans le système. Toutefois, la supériorité d'une présentation en réseau par rapport à une présentation "linéaire" n'a pas été démontrée jusqu'ici. Ces recherches se poursuivent actuellement pour déterminer quels paramètres peuvent augmenter l'efficacité de la présentation informatisée de systèmes documentaires.

Pour conclure, je voudrais souligner que les travaux présentés dans cet article ne représentent qu'une première exploration d'un ensemble d'activités dont l'étude scientifique apparaît désormais nécessaire. En effet, l'adaptation des technologies de l'information aux besoins et aux capacités cognitives des utilisateurs ne pourra faire l'économie d'un modèle cognitif général des processus de lecture et de compréhension, et de leur transposition aux systèmes électroniques. Il faudra pour cela poursuivre et développer les recherches du type de celles présentées dans cet article.

Les progrès de ces recherches impliqueront une collaboration étroite entre les spécialistes des sciences cognitives, des sciences de l'information et bien entendu des éducateurs. Il faut donc souhaiter que ces collaborations, encore très ponctuelles à l'heure actuelle, puissent se développer dans les prochaines années.

Bibliographie

Balpe, J.P. (1990). *Les hyperdocuments - Hypertextes - Hypermedia*. Paris: Eyrolles.

- ○ Britt, M.A., Rouet, J.-F. & Perfetti, C.A. (1996). Using hypertext to study and reason about historical evidence. In J.-F. Rouet, J.J., Levonen, A.P. Dillon, and R.J. Spiro (Eds.). *Hypertext and Cognition* (pp. 43-72). Mahwah, NJ: Lawrence Erlbaum Associates.

Charney, D. (1994). The impact of hypertext on processes of reading and writing. In S. J. Hilligoss & C. L. Selfe (Eds.), *Literacy and computers: The complications of teaching and learning with technology* (pp. 238-263). New York: Modern Language Association.

- de la Passardière, B. & Baron, G.-L. (Eds., 1991). *Hypermédiat et Apprentissages*. INRP, MASI

Denhière, G. & Baudet, S. (1992). *Lecture, Compréhension de Texte et Science Cognitive*. Paris: P.U.F.

Fayol, M. (1985). *Le récit et sa construction*. Paris: Delachaux et Niestlé.

- ○ Fayol, M., Gombert, J.-E., Lecocq, P., Sprenger-Charolles, L. et Zagar, D. (1992). *Psychologie cognitive de la lecture*. Paris: Presses Universitaires de France.

Flavell, J.H. (1981). Cognitive Monitoring. In P. Dickson (Ed.) *Children's Oral Communication Skills*. New York: Academic Press.

Freebody P. & Anderson R.C. (1983). - Effects of vocabulary difficulty, text cohesion and schema availability on text comprehension. *Reading Research Quarterly*, **18**, 277-294.

Gombert, J.E. (1990). - *Le développement métalinguistique*. Paris: Presses Universitaires de France.

Kintsch, W. & van Dijk, T.A. (1978). Toward a model of text comprehension and production. *Psychological Review*, **85**, 363 - 394.

Nelson, T.H. (1965). - A file structure for the complex, the changing, the indeterminate. *ACM 20th National Conference Proceedings* (84-100). Cleveland (Ohio).

Paris S.G., Lipson J.E. & Wixson K.K. (1982). - Becoming a strategic reader. *Contemporary Education Psychology*, **8** (3), 293-316.

Rouet, J.-F. (1991). *Compréhension de textes didactiques par des lecteurs inexpérimentés dans des situations d'interaction sujet-ordinateur*. Thèse de Doctorat, Université de Poitiers, Laboratoire de Psychologie du Langage.

Rouet, J.-F. (1992). Apprendre à lire un hypertexte: une étude expérimentale. *Cahiers de Linguistique Sociale*, **21**, 81-92.

Rouet, J.-F. (1994). Naviguer sans se perdre: lecture et acquisition de connaissances à l'aide des hypertextes. *Enseignement Public et Informatique*, **73**, 97-107.

Rouet, J.-F. & Tricot, A. (1995). Recherche d'informations dans les systèmes hypertextes: des représentations de la tâche à un modèle de l'activité cognitive. *Sciences et Techniques Educatives*, **2**(3), 307-332.

Stahl, S.A., Jacobson, M.G., Davis, C.E. & Davis, R.L. (1989). - Prior knowledge and difficult vocabulary in the comprehension of unfamiliar text. *Reading Research Quarterly*, **24** (1), 27-43.

van Dijk, T.A. & Kintsch, W. (1983). *Strategies of Discourse Comprehension*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Williams, J.P., Taylor, B.M. & Ganger, S. (1981). Text variations at the level of individual sentence and the comprehension of simple expository paragraphs. *Journal of Educational Psychology*, **73**, 851-865.

La société des hommes de Neandertal était déjà bien organisée. Ils vivaient dans des tribus où chacun avait un rôle à jouer. Les activités importantes étaient organisées en commun sous l'autorité d'un ou

plusieurs chefs. Les hommes de Neandertal furent les premiers à prendre soin des vieillards et des infirmes.

<i>relation implicite (*)</i>	<i>relation explicite (*)</i>
Quelles activités se sont développées grâce à cette vie communautaire?	Cette vie communautaire a favorisé le développement d'activités traditionnelles et artistiques.

Choisis un titre:

1. Les conditions d'existence
2. L'utilisation du feu
3. L'organisation de la société
4. Les activités traditionnelles
5. Les outils
6. L'aspect physique

Figure 1: un système hypertexte expérimental (*explication dans le texte).

