

PUDELKO, B., CRINON, J., ET LEGROS, D. (2002). LECTURE ET COMPRÉHENSION DES HYPERTEXTES. DANS D. LEGROS, ET J. CRINON (DIR.) *PSYCHOLOGIE DES APPRENTISSAGES ET MULTIMÉDIA*, (PP. 84-106). PARIS : ARMAND COLIN.

CHAPITRE IV

Lecture et compréhension de textes

Jacques Crinon, Denis Legros et Béatrice Pudelko

Introduction

Ce chapitre présente une synthèse de travaux conduits sur les outils et systèmes multimédias dédiés à la lecture et à la compréhension de textes. L'examen de la littérature consacrée à cette thématique nous a conduits à diviser ce vaste domaine en deux parties : la première partie est consacrée à la synthèse des recherches menées sur les systèmes et outils conçus pour favoriser l'apprentissage de la lecture. La deuxième partie, consacrée à la synthèse des travaux centrés sur la problématique de la lecture et de la compréhension des textes électroniques (construction de la signification du texte lu), est dominée par la présentation des travaux sur les systèmes hypertextes.

En effet, les documents électroniques sont de plus en plus souvent fondés sur des techniques hypertextes et les problèmes de lecture et de compréhension qu'ils engendrent ont donné lieu à une littérature abondante (voir Goldman, 1996 ; Reinking, McKenna, Labbo & Kieffer, 1998). Dans ce chapitre nous analysons les recherches sur les hypertextes, uniquement sous l'angle de leur lecture et de leur compréhension, soit encore comme aides à la lecture et à la recherche d'informations. C'est pourquoi nous considérerons l'hypertexte dans son sens restreint de "hyper-texte", c'est-à-dire un système dans lequel les unités de la couche d'information sont de nature textuelle. Nous adoptons ce point de vue, bien que le concept d'*hypertexte* soit employé en général de façon interchangeable avec celui d'*hypermédia*, sa caractéristique principale étant en effet sa facilité d'établir des liens entre les différents formats des informations (texte, graphique, image, son) et un accès par navigation à toutes les informations stockées (Conklin, 1987 ; Marchionini, 1988 ; Spiro & Jehng, 1990 ; Balpe, 1997).

On constate que l'évolution des documents sur écran vers un format multimodal constitue une tendance forte : de plus en plus souvent, l'écrit y voisine avec l'image et le son. La facilité d'intégrer les différents formats de représentation constitue un atout majeur des systèmes d'aide à l'apprentissage de la lecture (Mayer & Moreno, 1998 ; Vandendorpe, 1999), et leurs effets seront analysés dans la première partie de ce chapitre. Cependant, étant donné l'ampleur des travaux sur la problématique de l'hypertexte, nous avons choisi, par souci de cohérence, d'analyser les aspects multimodaux des hypertextes dans le chapitre sur la construction des connaissances.

1. L'apprentissage de la lecture

Des logiciels d'aide à la maîtrise du code

La synthèse vocale

Les systèmes de synthèse vocale permettent à une machine de prononcer des mots ou des séquences de mots à partir des caractères d'un texte écrit. Ces systèmes, utilisés surtout dans les pays de langue anglaise, ont montré leur efficacité avec des lecteurs en difficultés de décodage et de reconnaissance de mots. Ainsi Olson et Wise (1992) ont comparé les effets de ce système sur les performances de sujets présentant une déficience dans le décodage phonologique et la reconnaissance de mots. Un premier groupe de sujets a utilisé un tel système pendant un semestre à raison de seize périodes d'une demi-heure prise sur leur horaire d'enseignement de la langue maternelle ; le groupe témoin a suivi l'enseignement normal. Les progrès dans le décodage phonologique ont été quatre fois plus importants dans le premier groupe que dans le groupe témoin, et les progrès en reconnaissance de mots deux fois plus importants. Cette étude met en évidence également des différences de performances en fonction du niveau initial des élèves et du soutien individuel apporté par l'adulte.

Lovett, Barron, Forbes, Cuksts et Steinbach (1994), ainsi que Wise (1992) ont comparé les performances dans trois conditions de restitution orale des mots : son par son, distinction de l'attaque et de la rime, et par mots entiers. Les résultats sont différents dans les deux études. Les élèves qui ont fait l'objet d'un diagnostic d'atteinte neurologique (étude de Lovett & *al.*) ont tiré un meilleur bénéfice d'un entraînement incluant la segmentation (attaque/rime et phonèmes). La population d'élèves en difficulté de lecture ne présentant pas de troubles neurologiques a progressé de façon identique, quel que soit le type d'entraînement.

Dans une étude de Wise et Olson (1992), le système propose des feed-back vocaux de la lecture orale de mots. Deux versions du système ont été utilisées. La première version fournit, à la demande de l'élève, la prononciation du mot entier. La seconde propose la prononciation correcte de ce qu'il a tenté de prononcer, que le mot soit complet ou incomplet. Cette procédure permet à l'élève de comparer son essai au résultat souhaité. Cette seconde version du système a été favorable aux progrès du décodage phonologique avec les enfants de dix à douze ans. Les auteurs n'ont en revanche pas constaté de différences avec les enfants de sept à neuf ans et de treize à quatorze ans¹.

Scrase (1998) a évalué les performances obtenues par des élèves dyslexiques à une batterie d'exercices de remédiation présentée sur ordinateur. Les résultats de cent cinquante élèves ont été enregistrés pendant quatre ans. Le rythme des progrès obtenus dans les tâches de décodage a été mesuré, ainsi que les effets du programme sur ces progrès. Les données permettent de conclure à l'efficacité du système que les auteurs attribuent principalement à deux fonctionnalités du logiciel. D'une part la possibilité de choisir la couleur des lettres et du fond, qui permet de réduire le " *stress visuel* " des élèves atteint du syndrome de Meares-Irlen ; d'autre part, la synthèse vocale, qui favorise le développement des capacités de segmentation phonémique.

On notera donc, à la fois, l'intérêt de ces approches pour la remédiation et la rééducation de certaines difficultés de lecture. Leur efficacité avec de jeunes enfants n'est cependant pas clairement affirmée.

L'entraînement auditif et visuel

Les logiciels visant à entraîner les élèves à la correspondance entre les phonèmes et les graphèmes ou bien à développer les aspects idéo-visuels de la lecture sont nombreux sur le marché. Les premiers visent le public des jeunes enfants en début d'apprentissage de la lecture, les seconds s'adressent à des élèves plus âgés (fin de l'école primaire ou collège).

Dans la première catégorie, on trouvera par exemple, en France, des logiciels parascolaires (*Adibou, Atout Clic CP*, Hachette...) ou ludo-éducatifs (*Moi je sais lire* du Club Pom...).

Atout Clic CP propose ainsi différentes activités d'aide aux premiers apprentissages de la lecture : extraire une graphie ou un son parmi plusieurs, trier des lettres, des mots, des syllabes, compléter des mots, écrire sous la dictée, remettre les lettres d'un mot en ordre, repérer un mot lu dans une liste de mots, s'entraîner à l'identification rapide de mots. Il s'agit d'exercices dont la présentation se veut ludique et humoristique. Un personnage crayon anime les séquences, présente les activités, fait

¹ Olofsson (1992) conclut aussi à l'absence de résultats avec les plus jeunes enfants.

des commentaires sur le temps mis par l'enfant pour réaliser un travail ou intervient pour rappeler l'utilisateur à l'ordre s'il n'agit pas sur la machine...

Moi je sais lire a pour objectif d'entraîner l'enfant à la discrimination auditive et visuelle, à l'acquisition d'un capital de mots et de l'orthographe de mots simples, à la construction des mots à partir de lettres et de syllabes, à la construction de phrases, et enfin à la construction de la signification des phrases.

Dans la seconde catégorie, les logiciels d'entraînement visuel fondés sur les facilités d'affichage de l'ordinateur (*ELMO* ou *ELSA* de l'Association française pour la Lecture, *Lirebel* des Éditions Chrysis), permettent de développer un ensemble d'habiletés (*skills*) du lecteur expert : mémorisation des mots et de leur signification, élargissement de l'empan de lecture, accélération de la vitesse de lecture, anticipation (exercices de *closure*), recherche non linéaire d'une information dans une page...

L'intérêt des logiciels destinés aux enfants en début d'apprentissage de la lecture semble confirmé par des études comme celle de Boone, Higgins, Notari et Stump (1996), qui porte sur 143 élèves de jardin d'enfants. Un logiciel hypermédia propose des jeux permettant d'associer les lettres aux sons correspondants et à des images d'objets dont le nom contient ce son. L'étude, fondée sur la comparaison entre les données d'un groupe expérimental et celles d'un groupe témoin, conclut à l'efficacité de cet entraînement sur la reconnaissance de lettres. Cette efficacité est plus marquée pour les élèves bons ou moyens que pour les élèves en difficulté.

L'efficacité sur l'apprentissage de la lecture de ce type de logiciel d'entraînement reste cependant discutée. Les évaluations disponibles sont certes positives, dans le cadre même de ce type d'entraînement², mais les données expérimentales fiables et généralisables sont rares. Augmenter la vitesse de lecture est certes un objectif à poursuivre pour améliorer les compétences des mauvais lecteurs. Les limites de la mémoire de travail rendent en effet peu efficace une lecture trop lente. Des études expérimentales ont mis en évidence, non seulement le rapport de causalité entre la vitesse de lecture et la compréhension, mais aussi l'amélioration de la compréhension lorsqu'on accélère la vitesse de lecture (voir Breznitz, 1997). S'agit-il d'accélérer l'activité de décodage phonologique ou de favoriser l'accès direct au lexique mental ? Breznitz (1997) fournit des arguments expérimentaux en faveur des deux stratégies avec des lecteurs débutants. Il montre, de plus, l'intérêt de la seconde stratégie pour améliorer les performances des enfants présentant un dysfonctionnement dans l'activité de reconnaissance des mots liée à un déficit phonologique. On peut s'interroger sur la pertinence de la focalisation des entraînements sur ces processus dans le développement de la capacité de lecture. Mais cette question interroge non pas la pertinence des outils techniques, mais la pertinence des théories de l'apprentissage de la lecture qui fondent la conception de ces outils.

Si les évaluations des gains immédiats dus à des exercices informatisés sont positives, reste posée la question de la généralisation de ces gains hors du contexte des exercices sur ordinateur. Ainsi, Olson et ses collègues (1997) ont suivi pendant deux ans une population d'enfants présentant des difficultés de lecture. Ces enfants ont suivi un programme de remédiation sur ordinateur, incluant pour les deux groupes expérimentaux une aide à la lecture d'histoires sous forme de synthèse vocale des mots difficiles, et, pour seulement l'un des deux groupes, un entraînement explicite à l'analyse phonologique. Les effets immédiats du programme, pour l'un comme pour l'autre groupe, ont été importants. Mais l'hypothèse selon laquelle le groupe "analyse phonologique" obtiendrait, avec le temps, de meilleurs scores de reconnaissance de mots que l'autre groupe ne s'est pas vérifiée. Après deux ans, on n'observait pas de différences entre les performances des deux groupes, même là où le premier groupe réussissait mieux à l'issue de l'entraînement à la conscience phonologique et au décodage. Les auteurs attribuent ce résultat à un temps d'entraînement insuffisant et à un manque d'explication de la manière d'utiliser sans l'aide de l'ordinateur les compétences acquises.

La généralisation et le transfert des compétences à des situations réelles de lecture nécessitent une conception de l'apprentissage et donc de l'enseignement qui facilite l'explicitation des acquis dans le cadre du fonctionnement cognitif de l'élève.

² Voir par exemple, pour *Lirebel*, www.education.gouv.fr/sec/servlet/lireb3.htm.

Les livres animés

Les livres animés qui se sont multipliés dans l'édition grand public³ offrent des fonctionnalités séduisantes pour faciliter les premiers apprentissages de la lecture :

- les aides orthographiques, orales, imagées visent à faciliter la compréhension des mots ;
- la même information est présentée de manière multiple. Les apprenants peuvent utiliser la modalité d'aide qui leur convient le mieux ;
- les mots nouveaux peuvent être prononcés ;
- la possibilité de lire et relire indéfiniment améliore la fluidité de la lecture.

La compréhension des textes semble meilleure sur les livres CD-ROM que sur les livres papiers (Matthew, 1996). Les enfants passent plus de temps devant un livre CD-ROM que devant un album lu par un adulte. Les images et les animations maintiennent l'intérêt de l'élève pour l'activité (Goldstein, Olivares & Valmont, 1996). En outre, des résultats expérimentaux déjà anciens portant sur l'effet de la lecture et de l'écoute simultanées sur la compréhension d'un texte peuvent être transposés à ces situations. Neville (1975, cité par Topping, 1999) a montré, avec un échantillon de 180 enfants, que les performances de compréhension de textes sont meilleures lorsque les textes enregistrés sont écoutés et, en même temps, suivis visuellement sur le papier.

Il reste à s'interroger sur l'origine de ces effets : faut-il les attribuer aux fonctionnalités du produit ou à un effet de motivation lié à la nouveauté ? Le caractère ludique des animations, qui peut détourner l'enfant du texte et modifier ses activités de traitement, ne semble pas constituer un facteur d'aide à la compréhension. Dans une recherche expérimentale conduite avec des enfants de quatre à sept ans, Oyen et Bebko (1996) ont comparé les effets de deux situations d'apprentissage : le jeu sur ordinateur et la leçon. Le nombre d'informations rappelées par les élèves dans la condition " leçon " était significativement plus important que dans la condition " jeu ".

Il ne semble cependant pas que la présence de l'animation fasse obstacle à l'activité de compréhension. Underwood et Underwood (1996) ont observé 62 garçons et filles de huit ans travaillant en tandem sur ces livres animés. Les animations sont souvent utilisées mais, comme l'indiquent les résultats d'une épreuve de rappel libre, ce sont les éléments importants de l'histoire qui sont retenus.

Conclusion

Les critiques souvent portées contre les tentatives d'aide informatisée à la lecture sont en réalité le plus souvent des critiques portées contre les théories implicites de l'apprentissage qui les inspirent. Ainsi les logiciels d'aide à la lecture peuvent favoriser des traitements locaux et porter remède à des difficultés ponctuelles. Ils sont cependant souvent inefficaces pour aider les élèves en échec et incapables non seulement de comprendre les textes, mais encore d'assigner au langage écrit ses fonctions de communications et d'outils cognitifs pour penser et pour apprendre (McKenna, Reinking & Labbo, 1998). À cet égard, certaines pratiques pédagogiques de lecture sont plus favorables que d'autres pour activer et élargir le champ de connaissances et le champ d'expériences des élèves, en particulier les pratiques qui favorisent l'échange, le partage de la lecture, le travail coopératif. Un des apports essentiels des nouvelles technologies est de permettre le développement du travail collaboratif à distance et des communautés virtuelles⁴. On connaît par ailleurs l'importance du rôle du milieu familial dans la compréhension par l'enfant du sens de la lecture et certains spécialistes fondent de grands espoirs sur le rôle que pourrait jouer l'ordinateur à la maison dans l'implication des parents dans cet apprentissage (voir Topping, 1999).

De nombreuses applications informatiques inspirées de la " *whole language approach* " – approche globale visant à inclure l'apprentissage des codes dans des activités langagières authentiques – ont vu le jour aux États-Unis (pour une revue, voir McCullough, 1995). Des logiciels incluant éventuellement des images et du son qui permettent ainsi de lier lecture et production de textes dans des

³ On peut citer quelques exemples heureux et bien diffusés en France : *Just Grand'Ma and me* (collection " *Living Books* " de Broderbund, proposant des versions en différentes langues du même livre) ou les adaptations proposées par Bayard Presse de la collection " J'aime lire " (*Un prince à l'école, La princesse Dragonne*).

⁴ Voir la présentation du CSCL (*Computer supported collaborative learning*) dans le chapitre v : " Apprendre à écrire avec l'ordinateur " ?

projets d'apprentissage interdisciplinaire existent aussi en France (*L'auteur en herbe*, etc.). Mis à part les effets motivationnels souvent relevés, il est cependant difficile d'évaluer l'apport spécifique de l'ordinateur dans l'apprentissage de la lecture qui se fonde sur une conception aussi globale des apprentissages, et des situations aussi complexes.

Des logiciels ont aussi été conçus à l'intention des enfants d'école maternelle en vue de les sensibiliser au monde de l'écrit. Une étude ethnographique de Labbo (1996) montre comment, grâce à des activités adaptées à un éditeur multimodal (*Kid Pix 2*), l'écran se révèle être un espace d'activités symboliques. Il permet en effet aux enfants d'utiliser des images et des signes typographiques divers pour exprimer leurs idées et les communiquer. Il leur permet d'explorer des formes non conventionnelles d'"écriture", et il favorise ainsi la construction de symboles grâce à une exploration active et coopérative de l'environnement écrit. Shilling (1997) a observé pendant cinq mois le coin ordinateur d'une classe de jardin d'enfants en vue de tester la compréhension des conventions du langage écrit. Les enfants avaient accès à un traitement de texte muni d'un système de synthèse vocale et ils pouvaient jouer et créer des textes et des dessins. Cette étude montre que l'ordinateur peut favoriser chez les jeunes enfants la compréhension de "l'univers de l'écrit". Les échanges dont il est le support contribuent en outre, dès deux ou trois ans, au développement de l'expérience sociale du jeune enfant et de son langage.

2. La lecture et la compréhension des textes électroniques

Comprendre un texte consiste à construire une représentation compatible à la fois avec les données de la situation symbolique (évoquée dans le texte), avec la tâche à réaliser et avec les connaissances qui sont en mémoire (cf. Kintsch, 1998 ; Kintsch & van Dijk, 1978 ; van Dijk & Kintsch, 1983, pour une revue voir Graesser, Gernsbacher, & Goldman, 1997). Cette construction passe par la récupération des éléments linguistiques et leur analyse en propositions de forme sujet-prédicat, l'interprétation sémantique de ces éléments de la "micro-structure", leur condensation en macro-éléments de signification, leur intégration aux connaissances antérieures du sujet. Elle aboutit à une représentation cohérente de ce que dit le texte (voir Denhière & Baudet, 1989) : le "modèle de situation" (Perrig & Kintsch, 1985 ; Mannes & Kintsch, 1987) ou le "modèle mental" du texte et de son contenu (Johnson-Laird, 1980 ; 1983).

Paramètres physiques des textes électroniques

La présentation du texte sur écran est traditionnellement signalée comme un handicap pour la lecture. Il est en fait difficile de comparer de manière générale la lecture sur papier et la lecture sur écran. Ainsi, lorsqu'on utilise des écrans à haute résolution, les vitesses de lecture sont équivalentes. Les recherches sur l'ergonomie des écrans montrent que les procédés de présentation jouent en outre un rôle non négligeable sur les performances des lecteurs, qu'il s'agisse de typographie ou autres marques paralinguistiques⁵ (Caro & Bétrancourt, 1998), la présence d'"escamots" (Bétrancourt & Caro, 1998) ou les caractéristiques des fenêtres d'affichage.

Ce dernier aspect a été étudié par Piolat et Roussey (1999). Deux études expérimentales ont été conduites dans le but d'étudier les effets de types de présentation du texte : page par page et dans des fenêtres défilantes (*scrolling*). Dans la première, les sujets devaient lire un texte, replacer des phrases dans le texte et résumer celui-ci ; la seconde proposait une tâche de révision du texte lu. Dans un cas comme dans l'autre, les performances du groupe "présentation par pages" ont été supérieures à celles du groupe "*scrolling*" : meilleure localisation des informations importantes, meilleur rappel des idées principales, plus grand nombre de corrections portant sur le niveau global du texte. La présentation par page semble permettre une meilleure représentation mentale de la hiérarchie des informations. Des expériences déjà anciennes avaient montré que la lecture sur écran est à la fois plus lente que la lecture sur papier (Gould & Grischkowski, 1984 ; Gould & al., 1987) et plus fatigante (Wilkinson & Robinshaw, 1987).

Dillon, dans une revue de question publiée en 1992, conclut que la vitesse de lecture est probablement plus lente sur écran, mais que la compréhension n'en est pas affectée.

⁵ Comme la couleur, les caractères spéciaux, la disposition du texte, la place des figures...

Les livres électroniques

Les livres électroniques, grâce à leurs fonctionnalités, semblent ouvrir aux étudiants et aux chercheurs de nouvelles perspectives de travail. Ces fonctionnalités constituent en effet une aide à l'exploration tabulaire des textes. Parmi ces livres électroniques, on citera le pionnier en la matière, le monumental *Discotext 1* (CNRS-InaLF/Hachette), mais aussi la collection des "Hyperlivres" (Éditions Ilias) ou encore les productions des éditions Acamédia (*Alexandre Dumas, Balzac, Chateaubriand*). Ces éditions numérisées des classiques de la littérature sont sans doute moins commodes pour une lecture cursive que les éditions sur papier, mais elles permettent de retrouver rapidement un passage et de le copier/coller dans un traitement de texte, d'établir rapidement des listes d'occurrences de mots ou d'expression, de travailler sur les contextes grâce aux méthodes de l'analyse lexicométrique, de faire appel à des dictionnaires intégrés...

Le développement rapide sur Internet de bibliothèques de textes numérisés, en particulier en France, sur le serveur Gallica de la bibliothèque nationale de France ou encore sur le site de l'association des bibliophiles universels⁶), ainsi que l'apparition récente du nouveau support que constitue le livre électronique rechargeable semble promettre un bel avenir à ce type d'outils.

Les recherches sur le traitement cognitif des textes électroniques sont encore très peu développées (voir Gillingham, 1996 ; Goldman, 1996 ; Dillon, 1996). Les applications pédagogiques n'en sont sans doute qu'à leurs balbutiements. Les travaux de Pierre Muller, menés à l'INRP dès le début des années quatre-vingt, ont permis d'expérimenter et d'éditer des outils lexicométriques pour les études littéraires au lycée (collection "Pistes" au CNDP). On peut évoquer aussi l'exemple du système LYRE, décrit par Bruillard (1997). Un affichage en surbrillance de certains mots produit des effets sur l'activité de lecture, en activant successivement tel ou tel réseau sémantique. À cet égard, un simple traitement de texte peut constituer un instrument de lecture. La transformation de la présentation ou de la mise en page d'un texte permet de faciliter l'activation d'autres significations ou d'autres superstructures (van Dijk, 1977) restées inactivées avec la présentation linéaire du texte.

Les aides " en ligne "

Parmi les aides à la lecture que proposent les livres électroniques, les fonctions hypertextuelles ont été les plus étudiées : qu'il s'agisse de notes de vocabulaire ou de questions visant à favoriser une lecture active (par exemple, aux États-Unis, le projet *Electrotext*, Horney et Anderson-Inman, 1994, 1999 ; en France, Collinot et Saustier, 1991), voire de conseils sur les stratégies de lecture à employer (Salomon, Globerson & Guterman, 1989). Quelle est l'efficacité de ce type d'aides, apportent-elles autre chose que leur équivalent dans les manuels traditionnels ?

La présence d'aides en ligne facilite la compréhension d'un texte : plusieurs recherches l'indiquent nettement. MacArthur et Haynes (1995) montrent les effets sur dix élèves de quinze à dix-sept ans présentant des difficultés de lecture, d'aides hypermédias destinées à compenser ces difficultés. Deux versions d'extraits d'un manuel de sciences sont présentées sur ordinateur : l'une est une simple reproduction du texte imprimé ; dans l'autre, on a ajouté une synthèse vocale, un glossaire, des liens entre les questions et le texte, un soulignement des idées principales, des explications supplémentaires. De meilleures performances de compréhension sont obtenus par les élèves utilisant la version hypermédia.

Rouet et Levonen (1996) résument ainsi une expérience de Reinking et Rickman (1990) :

"Ceux-ci ont demandé à des élèves de cinquième et sixième année (dix à onze ans) d'étudier un court texte scientifique dans une des quatre conditions de présentation suivantes : texte imprimé avec un dictionnaire, texte imprimé avec un glossaire de termes choisis, texte sur ordinateur avec les définitions soit optionnelles, soit obligatoires. Les sujets dans la condition "ordinateur et définition optionnelle" ont été beaucoup plus disposés à lire les définitions des mots que dans la condition "texte imprimé et glossaire". De plus, les deux conditions avec ordinateur ont abouti à un meilleur apprentissage du vocabulaire que sur papier."

Cet effet positif des aides en ligne dépend cependant de la conception de l'aide. Lachman (1989) a obtenu, avec des étudiants du début de l'enseignement supérieur, des résultats différents aux tests de compréhension portant sur un même texte, selon que les définitions proposées comme aide concernaient

⁶ <http://www.bnf.fr> et <http://cedric.cnam.fr/ABU>.

des mots et expressions plus ou moins importants (au sens du modèle de la compréhension des textes proposé par van Dijk et Kintsch, 1983).

Rouet et Levonen concluent à un effet positif sur la compréhension des textes des aides en ligne – effet qui ne semble pas dû au format différent des textes sur papier et sur écran. Les lecteurs tirent davantage profit des éléments non linéaires d'aide sur écran que sur papier. La présentation sur écran diminue en effet le coût cognitif de l'accès aux informations extérieures au texte et favorise chez les étudiants la compréhension des textes. Avec de jeunes enfants, les aides en ligne semblent moins efficaces. Encore faut-il s'entendre sur ce que l'on entend par aides en ligne. Celles-ci peuvent varier sur le contenu, mais aussi sur la modalité de présentation. Les lecteurs de documents numériques peuvent être perturbés par le passage d'une modalité sensorielle à une autre. Ainsi, dans l'expérience de Black, Wright, Black et Norman (1992), des lecteurs pouvaient obtenir des définitions des mots du texte. Dans l'un des groupes, cette définition était écrite et affichée dans la marge. Dans l'autre groupe, la même définition était donnée oralement. Contrairement aux hypothèses, les sujets n'ont pas plus utilisé l'aide orale que l'aide écrite. En outre, ils ont demandé à réutiliser les aides avant de répondre aux questions plus souvent dans la condition "aide orale".

Les liens lecture/écriture

Certains lecteurs éprouvent le besoin de gloser et de commenter les textes. Souligner, annoter un passage dans la marge, le recopier dans un cahier sont des pratiques courantes. Pratiques parfaitement rationnelles d'ailleurs : prendre des notes en lisant améliore la compréhension du texte lu (voir van Oostendorp, 1996). Cependant des expérimentations conduites par van Oostendorp ne mettent pas en évidence une quelconque supériorité des effets de la prise de note sur écran par rapport à la prise de note sur papier. L'intérêt des espaces électroniques réservés à la prise de note est donc plutôt de fournir des environnements intégrés de lecture, d'annotation et éventuellement de travail collaboratif à distance⁷.

Les hypertextes

Hypertexte :

"[C']est une démarche cognitive qui constitue un paradigme spécifique dans les méthodologies de construction des connaissances. [...] Cette démarche cognitive postule une construction non linéaire, relativement personnalisée, adaptative de la connaissance : elle fait l'hypothèse de la nécessaire variété des parcours d'apprentissage. Elle implique une observation externe des données d'information, quelque chose comme une métaconstruction constituant une représentation des ces données." (Balpe, 1997, p. 12.)

Les travaux théoriques

Dans la perspective de la psychologie cognitive, l'activité de la compréhension de texte a trois composantes : la situation, les connaissances et la tâche. Cette modélisation, qui rapproche les situations de compréhension des textes de celles de résolution de problèmes, permet de mettre en évidence les caractéristiques importantes du traitement : le rôle de la finalité perçue de l'activité de lecture et le rôle des connaissances et stratégies métacognitives portant sur l'analyse de la situation, la sélection de procédures pertinentes, le contrôle des différentes activités et l'ajustement des stratégies en fonction de la tâche (Haberlandt, Graesser & Schneider, 1989 ; Haberlandt, Graesser, Schneider & Kiely, 1986). De nombreuses recherches sur les processus de compréhension des textes "classiques" (c'est-à-dire linéaires) ont montré le rôle des connaissances sur le domaine évoqué par le texte dans la construction de la cohérence de la représentation, surtout au niveau global, et l'élaboration des inférences : les lecteurs qui ont des structures conceptuelles correctes du domaine intègrent mieux une nouvelle information et présentent un meilleur rappel du texte (Deschênes, 1988).

Par ailleurs, les lecteurs confirmés disposent d'une variété de stratégies pour progresser à travers un texte, en fonction des buts de lecture, des connaissances sur les structures textuelles et des connaissances causales (Fayol, 1991; Kintsch & Yarborough, 1982 ; Trabasso & van der Broeck, 1985) La plus typique (et la plus étudiée) des structures textuelles est la structure narrative du récit, caractérisée par l'ordre de présentation de l'information qui correspond à la logique sous-jacente du récit, habituellement temporelle, et qui implique souvent la connexion causale entre les deux

⁷ Sur ce dernier point voir le chapitre de ce rapport consacré à l'écriture. Pour un exemple de stations de lecture destinées à des chercheurs et intégrant des outils d'annotation, voir Stiegler, 1995.

événements. Comme le souligne Goldman (1996), étant donné que la structure narrative constitue la structure privilégiée (et “ par défaut ”) du texte (voir Coirier, Gaonac’h & Passerault, 1996 ; Denhière, 1984), il n’est pas surprenant de constater que les recherches sur la compréhension ont montré que la lecture s’effectue de gauche vers la droite, commence au début du texte et progresse vers le bas... En effet, dans ce type de texte, les lecteurs n’ont aucun intérêt à “ sauter ” d’un endroit à un autre du texte et n’en tireraient d’ailleurs qu’un bénéfice minime, puisque l’auteur lui-même a écrit le texte en présupposant que les lecteurs le liront de façon linéaire et non en “ sautant ”.

La différence principale entre les textes et les hypertextes consisterait dans la non-linéarité des derniers. Les documents hypertextes sont appelés “ non-linéaires ” dans la mesure où leur lecture est supposée s’effectuer de façon différente de celle des textes imprimés et “ linéaires ” : c’est le lecteur qui détermine l’ordre de lecture des différents nœuds de l’hypertexte. Cependant, pour Espéret (1996) cette conception de la linéarité résulte de la confusion entre trois différentes caractéristiques du traitement de l’information verbale :

- l’organisation des unités de langage du niveau inférieur ;
- la façon dont l’information est stockée dans un médium (livre, bande, hypertexte) ;
- la façon dont le lecteur contrôle les processus d’accès à l’information.

L’examen des deux premières caractéristiques, qui concernent la structure du texte, montre que le livre et l’hypertexte impliquent tous deux l’ordre linéaire des unités de langage (phonèmes, mots, phrases) et que, de plus, les hypertextes, à la manière des textes, possèdent souvent des marques linguistiques pour coder la cohérence au niveau des paragraphes. Par ailleurs, la présentation physique des unités de base (paragraphes ou nœuds) contraint les processus d’accès à ces unités aussi bien dans les textes imprimés que dans les textes électroniques. Cependant, à la différence de l’imprimé, la façon dont le texte est stocké dans un hypertexte n’est pas directement accessible au lecteur : c’est l’ordinateur qui se charge de retrouver l’information en offrant généralement une gamme d’outils qui permettent l’accès aux unités. Certains de ces outils sont similaires à ceux des textes imprimés (renvois, tables de matières), d’autres (cartes ou *fisheyes*) ont été développés pour faciliter la consultation des hypertextes. C’est pourquoi, Espéret (1996) propose de remplacer l’opposition entre un format linéaire et non-linéaire du texte par une notion de continuum de la flexibilité dans l’accès à l’information.

D’autre part, l’opinion largement répandue selon laquelle la non-linéarité de l’hypertexte est synonyme de liberté pour le lecteur est contestable. En effet, si, lors de la lecture d’un texte imprimé, le lecteur peut aller à son gré d’une page à l’autre, dans un hypertexte, la progression d’une unité textuelle à une autre est contrainte par les liens imposés par l’auteur. La prétendue liberté du lecteur de l’hypertexte n’est donc pas le résultat de la structure hypertextuelle, mais des décisions du lecteur lui permettant de construire un parcours personnel dans l’hypertexte (Rouet & Levonen, 1996 ; Tricot, 1995). C’est pourquoi le problème majeur de la lecture des hypertextes réside dans la troisième caractéristique : la façon dont le lecteur contrôle l’accès à l’information.

Si la lecture des hypertextes présente des difficultés, c’est parce qu’elle oblige le lecteur à contrôler de façon continue son activité de lecture en fonction de ses besoins et de ses objectifs. Pour construire une représentation cognitive satisfaisante du texte, il doit choisir quelle information doit être lue, d’abord, quelle autre traiter ensuite, etc. L’obtention du but désiré implique un ou plusieurs cycles de la sélection, de l’évaluation et du traitement d’information. C’est pourquoi, on a avancé que la difficulté majeure rencontrée par un lecteur d’hypertexte consiste à ne pas activer les connaissances qu’il possède déjà sur la structure typique de textes. À leur place, les lecteurs doivent apprendre les avantages des différents formats de présentation (par exemple, l’accès hiérarchique ou direct), donc acquérir de nouvelles habiletés stratégiques de traitement qui peuvent éventuellement se combiner avec les capacités existantes (Dillon, 1996 ; Espéret, 1992 ; Rouet & Levonen, 1996).

Cet accent mis sur l’importance des stratégies des lecteurs et des connaissances sur la structure de l’hypertexte, est d’autant plus justifié qu’une grande partie des recherches sur l’utilisation des hypertextes concerne les activités de la recherche d’informations, dans lesquelles l’explicitation des buts de la lecture et la maîtrise des stratégies de navigation conditionne la réussite de la construction d’une représentation cohérente du texte lu.

La complexité de la problématique de la lecture des hypertextes commence à susciter un renouvellement de la réflexion théorique sur les processus cognitifs mis en jeu dans la compréhension des textes, et une réinterrogation des modèles et des méthodologies existants (Goldman, 1996 ; van Oostendorp, 1996 ; Rouet & Passerault, 1999). Depuis les premiers travaux sur les hypertextes dont

l'objectif était l'évaluation comparative de l'efficacité des textes linéaires et non linéaires, le contexte de la recherche a évolué : les travaux actuels définissent plus clairement les enjeux des études sur les processus cognitifs mis en œuvre dans les activités basées sur la lecture des hypertextes du point de vue aussi bien des apprentissages que de la conception des systèmes. On peut avancer que les efforts de l'évaluation de l'efficacité des apprentissages dans le contexte des hypermédias convergent avec les efforts dans les études des processus discursifs pour examiner les processus de construction des modèles mentaux des textes (Goldman, Pellegrino & Bransford, 1994). Goldman (1996) a souligné que la focalisation des recherches sur la structure de texte lu, plutôt que sur les stratégies que les lecteurs emploient pour construire la signification du texte, est l'expression de l'influence insidieuse des conceptions d'apprentissage basées sur la transmission du savoir vers un apprenant passif (Cognition and Technology Group at Vanderbilt, 1993).

Les recherches empiriques

De nombreuses recherches sur la lecture des hypertextes et leur compréhension ont été effectuées avec des résultats souvent contradictoires (Jakobson, 1994 ; Jacobson, Maouri, Mishra & Kolar, 1995 ; Nielsen, 1989 ; Rouet & Levonen, 1996). Chen et Rada (1996), qui ont procédé à une méta-analyse de trente et une recherches menées entre 1989 et 1993, ont constaté une grande hétérogénéité des résultats expérimentaux rapportés, ce qui les a conduits à mettre en doute leur validité même. Ces auteurs ont souligné que la diversité des situations expérimentales rend problématique toute tentative d'effectuer des comparaisons valides entre les différentes recherches et concluent à la nécessité de s'accorder sur une classification des tâches dans les environnements hypertextes.

Pour Foltz (1996), une grande part de cette variabilité est due aux différences non contrôlées dans les caractéristiques des tâches, des types de textes employés et des systèmes utilisés. En effet, étant donné que l'hypertexte incorpore plusieurs caractéristiques entièrement nouvelles par rapport au texte linéaire (liens, outils de recherche, l'aide en ligne), il est nécessaire tout d'abord d'examiner le rôle et les apports de ces différents outils afin de déterminer leurs effets sur les processus de la lecture en ligne et la compréhension des hypertextes.

De plus, la recherche dans le domaine des hypertextes est jeune et s'est focalisée tout d'abord sur le développement des nouveaux systèmes dans une perspective athéorique " endémique " (Spiro & Jehng, 1990) dont les risques de dérive technocentrée ont été soulignés par Rouet et Levonen (1996).

Nous avons analysé les données empiriques sur les processus de la compréhension et la construction de la représentation des hypertextes en procédant à un classement des différentes variables étudiées dans le but de clarifier la problématique complexe des recherches sur les hypertextes. Pour ce faire, nous avons suivi la distinction souvent opérée dans la littérature (Gall & Hannafin, 1994 ; McKnight, 1996 ; Rouet, 1992) entre les caractéristiques des usagers, des tâches et les caractéristiques des systèmes incluant les fonctionnalités des outils. Ce découpage ne peut éviter certaines redondances dans la présentation, en raison des facteurs qui ne peuvent pas toujours être isolés et de leurs effets qui sont souvent confondus.

Les caractéristiques des systèmes

La non-linéarité de la présentation des informations dans les hypertextes a souvent été présentée comme plus avantageuse, car permettant une interaction " naturelle " avec une diversité de sources d'information (Collier, 1987; Jonassen, 1989). Une méta-analyse de 92 études (Nielsen, 1989) a montré que dans la plupart des recherches, les performances de l'hypertexte sont moins bonnes ou, dans le meilleur des cas, équivalentes à celles du texte papier. D'autres études ont montré que la lecture des textes électroniques non linéaires est plus difficile que celle des textes électroniques linéaires (Gordon, Gustavel, Moore & Hankey, 1988 ; voir Anderson-Inman & Horney, 1998 ; McKenna, 1998). Une recherche bien contrôlée de Foltz (1993, cité par Foltz, 1996) a comparé les deux formats de présentation (sur l'écran) d'un chapitre d'un livre sur l'économie, en procédant à un recueil des différentes mesures de la compréhension et les stratégies des lecteurs (temps de lecture, nombre de macropropositions des résumés, questionnaires à choix multiples). Par ailleurs, il a manipulé les objectifs des lecteurs, dont la moitié avait pour but le recueil des connaissances générales, tandis que l'autre moitié avait pour but de trouver des informations spécifiques. Ses résultats ne montrent aucune différence significative entre les performances des lecteurs de la version linéaire et celles des lecteurs de la version hypertexte, quel que soit l'objectif de lecture. La seule différence concerne les meilleures performances des lecteurs de la version hypertexte dans les questions portant sur la structure des documents.

Afin d'expliquer les résultats mitigés de ces études comparatives, deux sortes d'arguments ont été avancés. D'une part, dans la plupart des expériences de ce type, la conversion d'un texte linéaire en hypertexte consiste habituellement à restructurer l'information tout en maintenant les marques de progression (titres de chapitres, énumération, etc.), ce qui permet de lire l'hypertexte de la même façon linéaire que le texte papier, tout en y incluant des outils supplémentaires de navigation. C'est la stratégie adoptée par les concepteurs de *SuperBook* (Egan, Remde, Landauer, Lochbaum & Gomez, 1989 ; Landauer, Egan, Remde, Lesk, Lochbaum & Ketchum, 1993). D'autre part, faute de connaissances ou d'entraînement, les lecteurs n'utilisent pas ou peu les différents outils de navigation. Ils lisent les hypertextes de la même façon que les textes imprimés (Anderson-Inman, Horney, Chen & Lewin, 1994 ; Messing, 1996) ou encore abandonnent leurs tentatives de navigation en raison des difficultés éprouvées dans l'utilisation des outils de navigation mal conçus des hypertextes expérimentaux (Goldman, 1996).

La navigation. L'accès aux unités d'information formant la structure non linéaire de l'hypertexte s'effectue par " navigation ", ce qui permet au lecteur de passer d'un nœud hypertextuel à un autre, et de choisir ainsi son propre chemin dans " l'espace sémantique ". La navigation s'accompagne du phénomène de désorientation ou de " perte dans l'hyperespace " qui a très tôt été diagnostiqué comme la difficulté principale de la lecture des hypertextes (Barab, Bowdish & Lawless, 1997 ; Beasley & Waugh, 1996 ; Conklin, 1987 ; Gall & Hannafin 1994 ; Halasz, 1988 ; Nielsen, 1990 ; Rouet et Tricot, 1995). En revanche, les défenseurs des hypertextes affirment que la navigation hypertextuelle n'est qu'un faux problème (Landow, 1990). Dans les petites applications, elle ne constitue qu'un problème mineur (voir Dillon, McKnight, & Richardson, 1993). La récurrence dans la lecture (le fait de passer plusieurs fois par les mêmes nœuds hypertextes), peut même constituer un avantage pour les lecteurs dans les hypertextes complexes, puisque " la répétition fournit une force structurante puissante, et un moyen qui aide les lecteurs à synthétiser leur expérience de la lecture... " (Bernstein, Joyce & Levine, 1992).

Foss (1988, cité par Tricot, 1995) a identifié deux types de problèmes à la base de la désorientation. D'une part, le problème de la localisation : le lecteur ne sait plus où il se trouve, il ne sait pas comment accéder à quelque chose qu'il croit exister, il arrive à un endroit et ne sait plus pourquoi il est là, il se perd en digressions, il ne sait pas s'il reste des documents pertinents dans le système...). D'autre part, le problème de la charge de traitement : le lecteur a le sentiment de voir trop d'informations, de ne pas posséder d'outil pour les traiter, de ne rien retenir, il oublie les sélections faites précédemment, n'est pas capable de se représenter une vue d'ensemble ou de construire un résumé cohérent.

Des études ont confirmé que la désorientation est un phénomène important, existant même dans des hypertextes simples (Rouet, 1990). Etant donné l'accroissement de la charge cognitive lors de la lecture, il n'est pas surprenant de constater que les comparaisons des performances obtenues lors de la lecture du texte imprimé (situation familière) et la lecture de l'hypertexte, sont rarement en faveur de ce dernier (pour une revue voir Rouet & Levonen, 1996). Cependant, ce constat doit être modulé étant donné que la lecture dépend fortement des types d'interface (Schroeder, 1992), des différents modes d'accès à l'information (Leidig, 1992), de la structure même de l'hypertexte (Rouet, Britt, Mason & Perfetti, 1992), ce qui amène certains chercheurs à poser la question de la définition même des hypertextes, à savoir si un réseau d'informations à la stricte organisation hiérarchique constitue encore un hypertexte (Lai & Waugh, 1995).

Cette problématique concerne également la distinction, souvent effectuée, entre les hypertextes " ouverts " et " fermés ". En effet, si l'on peut envisager de structurer un hypertexte " fermé " (par exemple un CD-ROM) afin d'augmenter son potentiel pédagogique (Rhéaume, 1991), cela devient impossible avec les environnements ouverts, comme le *World Wide Web*. Comme le souligne Bélisle (1999), avec l'interconnexion des espaces due aux progrès technologiques (par exemple la possibilité d'appeler un site Web depuis un CD-ROM), la frontière entre les systèmes ouverts et les systèmes fermés tend à disparaître, ce qui rend la navigation encore plus difficile à maîtriser.

Les aides à la navigation. L'existence des différents outils de navigation que l'on peut catégoriser comme ponctuels (boutons d'aide), structurels (cartes, *fisheyes*, filtres, index) et historiques (historiques du parcours, marqueurs des liens visités) (Bruillard & La Passardièrre, 1994 ; La Passardièrre & Dufresne, 1992) constitue probablement la différence principale entre le texte imprimé et hypertexte (Dee-Lucas, 1996). En effet, les outils de navigation définissent les possibilités de la

localisation et de l'accès à l'information et, par là, influencent la façon dont les lecteurs vont interagir avec l'hypertexte.

Dans le cas des hypertextes pédagogiques, la conception de ces outils est particulièrement critique puisqu'elle détermine la façon dont les lecteurs pourront atteindre les buts de la tâche. Cependant, l'utilisation adéquate des outils de navigation nécessite un apprentissage spécifique qui accroît la charge cognitive liée à la gestion de la tâche de lecture (McKnight, Dillon & Richardson, 1990). Par ailleurs, les études sur les effets des différents types d'outils navigationnels sur la performance de la recherche d'information et la compréhension du contenu de l'hypertexte n'ont pas réussi à démontrer la supériorité d'un type d'outil sur un autre. Ainsi, on n'a trouvé aucune différence entre une recherche par index et une recherche par menu dans une encyclopédie hypertextuelle ou entre une recherche par index et une recherche par carte, que ce soit dans un hypertexte à structure linéaire ou à structure hiérarchique (Wright & Lickorish, 1990). Allinson et Hammond (1989) ont comparé cinq conditions de recherche dans un hypertexte avec un but exploratoire *vs* spécifique. Dans une première condition, les lecteurs effectuaient leur recherche à l'aide des seuls liens hypertextuels, dans chacune des trois autres conditions, un outil différent a été rajouté : la carte, l'index ou le parcours guidé. Dans la cinquième condition, les lecteurs disposaient de tous ces outils simultanément. Les résultats de tests de connaissances n'ont révélé aucune différence significative entre les cinq conditions, que ce soit pour le groupe avec des buts de recherche généraux ou le groupe avec des buts spécifiques. Leader et Klein (1996) ont examiné l'effet du type d'outil (carte, index, *browser*) sur les performances des lecteurs adultes dans une tâche de lecture d'une base de données hypertextuelles en fonction des styles cognitifs (dépendance à l'égard du champ). L'analyse de la variance des mesures recueillies (nombre d'écrans visités, temps passé sur chaque écran, test des connaissances et questionnaire sur la satisfaction) a montré un effet significatif de l'interaction type de recherche/style cognitif au test des connaissances : les sujets catégorisés comme "indépendants" réussissent mieux en utilisant l'index et la carte. En revanche, il n'y a pas de différence entre les "dépendants" et les "non dépendants" lors de l'utilisation de *browsers*. Cependant, les résultats doivent être interprétés avec prudence, d'autant plus que d'autres expériences similaires n'aboutissent pas aux mêmes résultats (Leidig (1992).

Ces résultats conduisent les auteurs à souligner la nécessité de développer des systèmes hypertextes offrant une grande variété d'outils de navigation, afin de faciliter leur usage par les différentes catégories d'apprenants.

Les structures. Du fait qu'elle constitue le fondement de la conception des hypertextes, la problématique des structures de l'hypertexte a été étudiée de façon approfondie. Selon l'approche hypertextuelle classique (Balpe, 1997), la conception des hypertextes consiste à faire dépendre les informations portées par les médias – quelle qu'en soit la nature – à deux couches de structuration différentes :

- une couche de premier niveau matérialisée par une structure topologiquement constituée des médias portant l'information terminale (la couche des informations),
- une couche de second niveau dont la matérialisation peut emprunter des modalités diverses – liens, ancres, modalités de calcul – portant un graphe de représentation plus ou moins conceptuelle du champ d'informations du premier niveau et permettant une circulation transversale⁸.

La séparation entre la couche des informations (contenu) et celle des connaissances (structures) constitue à la fois la richesse et la faiblesse des hypertextes (Nanard, 1995). En effet, contrairement au livre imprimé, où la connaissance apportée par l'auteur est indissociable de l'information, dans un hypertexte, elle peut rester non accessible au lecteur, qui tentera alors de deviner pourquoi l'auteur a proposé un tel lien plutôt qu'un autre (Charney, 1994).

Les recherches indiquent que cette séparation de la structure et du contenu influence grandement la lecture et l'acquisition des connaissances à partir des hypertextes. Ainsi, de nombreuses études empiriques ont montré que le manque de perception de la structure de l'hypertexte diminue de façon importante la lisibilité des hypertextes (Levonen, 1996), alors que les hypertextes structurés accroissent les performances des sujets. Les résultats de l'expérience de Dee-Lucas et Larkin (1995) montrent que

⁸ La presque totalité des CD-ROM actuels sont conçus selon une approche hypertextuelle classique : un ensemble fini et fixe de la couche des informations sur lequel est défini un ensemble fini de parcours prédéterminés, c'est-à-dire des réorganisations de la topologie de surface des informations (Balpe, 1997).

la structure hiérarchique de l'hypertexte, rendue accessible à travers un sommaire hiérarchique, favorise la recherche d'information et le rappel de l'organisation des titres, comparativement à la version non structurée de ce même hypertexte, où l'accès aux informations s'effectuait à l'aide d'un index alphabétique. De même, Dee-Lucas (1996) a montré qu'un hypertexte hiérarchique facilite la sélection de l'information pertinente pour les objectifs de la tâche, favorise le rappel de l'information principale et subordonnée et, par conséquent, la construction de la macrostructure mesurée à l'épreuve du résumé. D'autres études ont également montré que les hypertextes hiérarchiques conduisent à des performances meilleures que d'autres structures hypertextuelles (Girill & Luk, 1992 ; Mohageg, 1992 ; Simpson & McKnight, 1990 ; Xu, Dainoff & Mark, 1999 ; Lai & Waugh, 1995).

Cependant, l'effet des différentes structures dépend de l'objectif fixé à la tâche de lecture. Dee-Lucas et Larkin, (1995, expérience 2), ont constaté que la lecture de l'hypertexte orienté par un but spécifique (résumer le document) diminue les écarts dans les performances constatées antérieurement entre les différents formats de présentation. Ces résultats conduisent ces auteurs à proposer l'interprétation qu'une plus grande implication des sujets dans la tâche d'apprentissage peut permettre de surmonter l'influence des différents formats de présentation.

L'ensemble de ces résultats est compatible avec les données des recherches sur la compréhension de textes qui ont montré l'importance des marques de structure textuelles dans la construction d'une représentation mentale du contenu d'un texte. L'analyse des protocoles verbaux enregistrés lors de la lecture d'un hypertexte sur l'économie (Foltz, 1993, cité par Foltz, 1996) indique que les sujets utilisent des stratégies pour maintenir la cohérence de l'information traitée.

C'est pourquoi l'absence des marqueurs de la macro-structure du texte (titres, pages, table de matières) peut interférer avec les stratégies usuelles de la lecture et conduire à un manque de cohérence des textes non linéaires (Gillingham, 1993 ; Martindale, 1993 ; Rouet & Tricot, 1995). Deux expériences de Lacroix (1999) ont comparé l'effet du type des titres (rhétoriques ou sémantiques) et du format de présentation (linéaire vs non linéaire) sur la construction de la macro-structure de la représentation d'un texte informatif. L'analyse des performances a été faite à la fois en termes de produits (résumés et rapports) et de processus de lecture (progression de la lecture, nombre et endroit des pauses, temps de lecture). Cependant les résultats n'indiquent pas que le type de titre exerce une influence sur la construction de la macrostructure.

Les connaissances et l'expertise. Hillinger et Leu (1994) ont évalué les possibilités qu'offre un hypertexte à remédier aux manques de connaissances du domaine. Leurs résultats sur la lecture d'un hypertexte décrivant la maintenance d'une turbine d'avion indiquent que les apprenants novices obtiennent les mêmes résultats que les apprenants possédant déjà des connaissances sur le domaine. Ces résultats ont été retrouvés aussi bien en ce qui concerne les informations spécifiques (définies comme étant l'objectif de la lecture) que les informations générales.

Cette étude a également montré que les lecteurs novices obtiennent de meilleurs résultats sous le contrôle du système, alors que les apprenants possédant des connaissances sur le domaine ont de meilleures performances lorsqu'ils contrôlent eux-mêmes leur apprentissage. Ces résultats permettent à Hillinger et Leu (1994) d'avancer que l'interaction entre le niveau des connaissances initiales et le type de contrôle de l'environnement, devrait être prise en considération lors de la conception des systèmes hypertextuels sur les domaines complexes.

Un résultat similaire rapporté par Balcytiene (1999) a montré que les étudiants avec peu de connaissances initiales bénéficient davantage de l'étude d'un domaine hypertexte structuré que des étudiants possédant déjà des connaissances sur le domaine. Un hypertexte structuré permettrait ainsi à ces lecteurs de former d'abord une vue globale de la complexité du domaine sans les surcharger des détails.

L'expertise de la lecture constitue également un facteur important pour la réussite de la lecture d'un hypertexte (Leu, Gallo, & Hillinger, 1995, cités par Leu & Reinking, 1996). Les bons lecteurs obtiennent des meilleurs résultats lors de rappel. Les données et les analyses des protocoles de l'interaction effectuées par Leu, Gallo, et Hillinger (1995) indiquent que le niveau d'expertise de la lecture conduit les lecteurs à exploiter différemment les caractéristiques des hypertextes : les bons lecteurs utilisent des éléments graphiques inclus pour évaluer leur compréhension des textes, alors que les lecteurs faibles s'en servent pour essayer de construire la signification des passages textuels.

Les stratégies. Bien que l'importance des stratégies et des connaissances stratégiques lors de la lecture des hypertextes a été fréquemment soulignée, les études consacrées prioritairement à l'analyse des processus stratégiques de lecture, que ce soit au niveau micro- ou macrostructurel sont quasi inexistantes (Leu & Reinking, 1996).

Cependant, de nombreuses études ont montré de façon indirecte que les lecteurs des hypertextes sont fort concernés par le besoin d'établir une représentation de la macrostructure du texte. De plus, on a souligné que la lecture d'un texte non linéaire s'apparente davantage à l'activité de résolution de problèmes dans un espace-problème non familier que la lecture d'un texte linéaire (Rouet & Levonen, 1996). Cette dernière caractéristique de l'activité de lecture peut aussi être modélisée en termes de l'heuristique du "gravisement de la montagne" ("*hill climbing*", cf. le modèle CE+ de Lewis & Polson, 1990), selon laquelle afin de trouver une information spécifique les sujets doivent d'abord appairer leurs objectifs de lecture avec tous les nœuds d'information possibles et ce n'est qu'ensuite qu'ils procèdent de façon descendante en cherchant parmi ceux-là les nœuds pertinents. Lewis et Polson (1990, cité dans Rouet & Levonen, 1996) suggèrent qu'une des façons de rendre efficace la progression des lecteurs est de concevoir des hypertextes permettant aux lecteurs d'avoir des conjectures valides (successful guessing) sur la structure de l'espace-problème, ce qui peut être facilité par la présence des marques de la macrostructure du texte comme les titres. En effet, les sujets qui ont des objectifs de lecture généraux peuvent utiliser des titres comme aide à déterminer les relations entre nœuds et choisir les transitions les plus cohérentes.

Dans l'état actuel de connaissances, on peut dire que les stratégies des lecteurs des hypertextes évoluent avec l'entraînement (van Oostendorp, 1996 ; Rouet, 1990 ; 1991) et que leur ajustement en fonction de la spécificité de l'hypertexte correspond au besoin de la cohérence textuelle (Foltz, 1996) et au besoin d'accéder au niveau supérieur de l'hierarchie de l'information (Britt, Rouet & Perfetti, 1996b ; Dee-Lucas, 1996).

Cependant, le fait que les recherches sur la lecture des hypertextes sont effectuées le plus souvent avec les étudiants du niveau supérieur, présente le risque d'ignorer les aspects développementaux de l'activité de la lecture et de leur interaction avec les processus mis en œuvre lors de la construction des connaissances à l'aide des hypertextes (Alexander, Kulikowich & Jetton, 1994 ; Leu & Reinking, 1996). En effet, les études menées avec des lecteurs jeunes sont plus nombreuses sur les processus d'écriture hypertextuelle ou sur les stratégies coopératives (Oshima, Bereiter & Scardamalia, 1995 ; Signer, 1992 ; Turner & Dipinto, 1992), que sur la lecture *stricto sensu* dans les environnements hypertextes.

Les différences individuelles. Les stratégies et les connaissances stratégiques sont liées au problème de contrôle méta-cognitif des processus de lecture des hypertextes (cf. les sections de ce chapitre sur les problèmes de la navigation hypertextuelle). Dwyer (1994), à la suite de Gray (1990), a constaté que les étudiants utilisant les hypertextes éprouvent un sentiment de contrôler davantage leur activité, ce qui s'accompagne de niveaux d'implication personnelle plus élevés que dans les environnements traditionnels de lecture.

Les résultats de la méta-analyse effectuée par Chen et Rada (1996) indiquent qu'on ne peut pas conclure à l'existence des différences individuelles en termes de "styles cognitifs" ou de "*locus of control*", malgré l'importance accordée par certains chercheurs aux avantages de "l'indépendance à l'égard du champ" lors de la lecture des hypertextes (cf. l'étude de Leader & Klein, 1996). La seule variable individuelle qui semble jouer un rôle important lors de la lecture des hypertextes concerne les habiletés spatiales (Vicente & Williges, 1988) et celles-ci peuvent être notablement atténuées par l'implémentation des outils d'aide graphiques (cartes) qui favorisent la visualisation de l'organisation structurelle de "l'espace sémantique" (Chen & Rada, 1996). C'est pourquoi certains auteurs ont souligné l'importance des processus de la création d'une "carte cognitive" correspondant à "l'espace sémantique" d'un hypertexte (Gaines & Shaw, 1995 ; Jonassen, 1993 ; Lai & Waugh, 1995 ; Nigay, 1994). D'autres chercheurs ont relevé les limites conceptuelles des métaphores spatiales et ont proposé de développer plutôt une rhétorique propre aux hypertextes (Dillon, 1991 ; pour une revue voir Dillon, McKnight, & Richardson, 1993). En effet, les représentations topographiques construites par les lecteurs des hypertextes correspondraient aux structures rhétoriques typiques des textes et des documents, telles que la structure linéaire, hiérarchique, grille à plusieurs dimensions, tableau à double entrée (Gray, 1990, cité par Tricot, 1995).

Dans les environnements hypertextes, il devient particulièrement important de distinguer l'intérêt *situationnel*, habituellement transitoire de l'intérêt *individuel*, plus permanent et résultant des expériences individuelles antérieures à l'expérience (Alexander, Kulikowich & Jetton, 1994). L'intérêt situationnel peut avoir un impact négatif sur la performance des lecteurs, qui peuvent être distraits de leur tâche principale par des détails "séducteurs", des informations intéressantes, mais non importantes ou non pertinentes. Plusieurs recherches ont montré que ce que l'on considère généralement comme un avantage essentiel des hypertextes, à savoir la libre navigation dans "l'espace informationnel", n'est que rarement mise à profit par les apprenants, qui, bien souvent, s'intéressent davantage aux images et aux animations qu'au contenu de l'hypermedia (Anderson-Inman, Horney, Chen & Lewin, 1994 ; Barab, Bowdish & Lawless, 1997). C'est pourquoi, les chercheurs s'intéressant au rôle de ces facteurs doivent veiller à ne pas confondre les mesures de l'intérêt situationnel, effectuées après l'expérience et les mesures de l'intérêt individuel qui sont effectuées avant l'expérience.

Selon les données de Becker et Dwyer (1994), les lecteurs ont le sentiment de contrôler leur environnement et de s'impliquer davantage, lors de la lecture d'un hypertexte que lors de la lecture d'un même matériel présenté de façon traditionnelle. De façon similaire, la recherche de Hill et Hannafin (1997) sur la navigation dans le Web à l'aide de *Netscape* indique que le nombre de stratégies utilisées nécessaires à la recherche d'informations pertinentes est proportionnel au niveau de perception de son efficacité personnelle. Par ailleurs, les recherches expérimentales peuvent difficilement isoler l'effet du facteur motivation des autres facteurs (Alexander, Jetton & Kulikowich, 1995 ; Dillon, 1991).

Les caractéristiques des tâches

L'examen des recherches effectuées sur les processus de compréhension mis en jeu dans le traitement des textes dans les environnements traditionnels et dans les environnements électroniques montre que ces derniers sont considérés davantage comme des *outils* plutôt que comme des *supports* de connaissances (Platteaux & Rickenmann, 1998). Les résultats soulèvent de nouvelles questions sur les objectifs de l'activité de la lecture et leurs effets sur la compréhension. La construction d'une représentation cohérente de ce qui est dit dans un hypertexte ne constitue plus le but exclusif de la lecture, comme l'illustre l'utilisation de la métaphore de "broutage" (*browsing*) pour décrire l'activité typique d'un lecteur d'hypertexte qui ne possède pas d'objectif précis. Pour Tricot (1995), la possibilité pour le lecteur d'avoir des buts "flous" est une grande nouveauté des hypertextes comme outils cognitifs. Il considère que la navigation libre (*browsing*) peut constituer un avantage pour la consultation d'une base de données, puisqu'elle permet une exploration de l'espace sémantique et la construction progressive et cohérente des connaissances (Tricot, 1995). Cependant, selon Tonta (1991), les systèmes hypertextes ne sont pas conçus pour une recherche rapide et efficace des informations, mais plutôt pour une recherche informelle, lente et temporellement coûteuse. C'est pourquoi, la consultation libre d'un hypermédia n'est pas adaptée à l'apprentissage d'un domaine complexe.

Mais lire un hypertexte c'est également "harponner" l'information (Nanard, 1995), puisqu'un système hypertexte a pour caractéristique principale de faciliter la tâche de la recherche d'information. La recherche d'information est considérée soit comme une activité d'apprentissage fondamental (Jonassen & Grabinger, 1990), soit comme un cas spécial de résolution de problème (Marchionini, 1989) caractérisé par des stratégies itératives, non linéaires et opportunistes (Marchionini, 1992).

D'autres modélisations des activités cognitives de la recherche d'informations dans les systèmes hypertextes ont mis l'accent sur la planification des différentes étapes de la recherche. Ainsi, Guthrie, Britten et Barker (1991) ont proposé un modèle cognitif de la recherche d'information en cinq étapes : formation de but, sélection de la catégorie (la sélection de la branche), extraction de l'information, intégration et recyclage. Leurs résultats montrent que l'étape cruciale de la recherche, et donc la plus coûteuse en ressources cognitives, est la sélection de la catégorie (la branche) pertinente au but formulé. C'est pourquoi les hypertextes à structure hiérarchique facilitent cette étape en permettant de réduire le temps alloué à la recherche de la branche adéquate.

Les recherches empiriques consacrées à l'activité de recherche d'information ont étudié principalement les différences entre les buts de la tâche et ont montré que la représentation du but influence le traitement du contenu. Les résultats d'une expérience de Foltz (1996) qui a manipulé les buts de lecture (recherche d'informations spécifiques vs générales) indiquent que les sujets dont l'objectif était de rechercher des informations spécifiques ont été plus rapides et ont manifesté une moins bonne compréhension lors des questions sur les connaissances générales. Cependant, dans cette tâche, le format hypertexte ne présente pas toujours un avantage comparativement à un texte linéaire

(Leidig, 1992) et, de plus, l'effet du format varie en fonction des buts du lecteur (Dee-Lucas & Larkin, 1999). Rouet (1994) a montré que l'entraînement favorise l'amélioration des stratégies de la recherche des utilisateurs novices, et conduit à l'élaboration des stratégies différentes en fonction du type de question (simple vs complexe).

De plus, on constate que le développement des stratégies efficaces est meilleur lorsque les lecteurs ont des buts spécifiques lors de la consultation des hypertextes à structure hiérarchique (*Hypertext Research Group*, 1999). De façon similaire, lorsqu'on indique aux lecteurs qu'ils vont avoir à reconstruire la structure logique du domaine, ils sont moins désorientés, et les résultats indiquent une corrélation positive entre l'acquisition des connaissances structurelles et la rétention des faits (Beasley & Waugh, 1996).

La recherche d'information peut être considérée comme une tâche de lecture des textes multiples dans laquelle le but de lecture est de rechercher la solution à des problèmes et à des questions. Pour ce faire, le lecteur doit déterminer si ce qu'il lit est pertinent par rapport à son objectif de recherche, effectuer des comparaisons des contenus entre les textes, les synthétiser pour en faire un tout cohérent. De plus, une lecture efficace des sources multiples nécessite la maîtrise d'autres habiletés tels que prendre les notes, organiser l'information, détecter les inconsistences et les redondances. Les chercheurs qui se sont intéressés à la problématique de la lecture des textes multiples ont montré que les adolescents et les adultes ont recours à des stratégies très variées pour construire la représentation basée sur la lectures de multiples sources (Britt, Rouet & Perfetti, 1996 ; Perfetti, Britt & Georgi, 1995 ; Reinking & Rickman, 1990 ; Wright, 1991) et cette tâche est d'une grande difficulté pour les enfants de onze ou douze ans (Goldman, 1997). L'exploration de multiples points de vue dépend également de l'expertise du domaine. Les travaux de Spiro, Feltovich, Jakobson et Coulson (1992) ont montré que l'entrecroisement ("*criss crossing*") est bénéfique pour les lecteurs qui possèdent un haut niveau d'expertise du domaine. Les novices bénéficient davantage d'une présentation linéaire et singulière, au moins lorsqu'ils abordent un domaine pour la première fois.

Conclusion et perspectives

L'ensemble de ces recherches indique qu'il existe une interaction complexe entre les caractéristiques des systèmes, les caractéristiques des lecteurs et les caractéristiques des tâches. La complexité de la problématique de la lecture des hypertextes et la nécessité de poursuivre des recherches empiriques sur les processus à l'œuvre dans environnements interactifs de lecture ont été soulignées par Rouet et Levonen (1996). En effet, des recherches empiriques fiables sont nécessaires pour guider la conception des environnements hypertextuels efficaces.

Par ailleurs les chercheurs soulignent fréquemment la nécessité de contrôler la complexité des situations expérimentales utilisant les technologies des hypertextes et d'augmenter la fiabilité des méthodologies employées. De même, on constate qu'il est de plus en plus urgent de développer un cadre théorique de l'interaction apprenant-système hypermédia (Rouet & Passerault, 1999).

En effet, on peut avancer que :

“ La spécificité de la lecture des hypertextes ne se trouve ni dans la diversité médiatique qu'elle sollicite, ni dans son fonctionnement “rhizomatique” ou sa non-séquentialité. En fait, la navigation déplace le centre d'activité dans son rapport à l'information. [...] Les mouvements de lecture qu'impose l'hétérogénéité inhérente à un hypermédia demandent une prise en compte de la diversité des contextes, des visées ou des niveaux d'information ; ils obligent à un travail de prise de conscience conceptuelle, qui était jusque-là de l'ordre de la métacognition, et qu'on peut appeler ici une *métalecture*. ” (Bélisle, 1999, p. 66.)

Selon les membres de l'équipe "*Hypertext Research Group*" (1999), les recherches se sont principalement focalisées sur les effets des hypertextes et de leurs fonctionnalités sur la compréhension et l'apprentissage. Le développement des technologies et la recherche de leur efficacité impose que les hypertextes puissent servir de laboratoire pour la recherche du traitement cognitif des textes.

Tout un programme...

Vers une maîtrise des textes électroniques

C'est donc un programme de compétences à acquérir qu'il faudrait établir à ce point de notre réflexion. Schetzer et Warschauer (à paraître) l'esquissent et classent les compétences liées à cette littéracie

électronique en trois catégories : communication, création et lecture/recherche d'informations. Dans cette dernière catégorie, ils recensent les points suivants :

- Comment trouver des questions à étudier, comment établir des mots clés, comment catégoriser et sous-catégoriser, comment “cartographier” les idées et les concepts (développement non linéaire des idées).
- Comment trouver de l'information en ligne en utilisant les indices d'organisation de la Toile, les moteurs de recherche et les autres outils de recherche spécialisés.
- Comment analyser ce que valent les informations trouvées et les outils.
- Comment déterminer l'autorité et l'expertise.
- Comment identifier les techniques rhétoriques de persuasion.
- Comment distinguer les sources primaires et secondaires.
- Comment citer les sources en ligne et comment donner les références des autres.
- Comment sélectionner les techniques de recherche disponibles : indices de recherche, moteurs de recherches, logiciels de brainstorming, etc.
- Comprendre les enjeux : problèmes de publicité, d'autorité, de protection de la vie privée, de qualité, de délinquance électronique. ”

Ces compétences ne sont pas seulement de l'ordre de la “ culture technique ”, ni d'une conception purement instrumentale de l'apprentissage de la lecture, mais bien d'une maîtrise critique des langages et d'une diversification des supports de lecture à l'école. Vingt-cinq ans après que Foucambert (1976/1994) a remis en cause l'impérialisme de la “ page grise ” dans l'apprentissage de la lecture, le travail sur les encyclopédies, les journaux, les hypertextes reste marginal à tous les niveaux du système scolaire. Dans l'enseignement secondaire français, sa prise en charge, lorsqu'il existe, par les documentalistes, le professeur de technologie ou des aides éducateurs tend à minimiser l'importance de ces compétences en tant que compétences de lecture à part entière ou à les réduire à de simples compétences techniques ou “ méthodologiques ” déconnectées des enjeux de compréhension critique des textes et de construction des connaissances. En matière d'introduction des technologies de l'information et de la communication dans le système scolaire, il y a sans doute là une priorité.

On insistera également sur le rôle du contexte d'usage des documents électroniques, et en particulier du contexte de la tâche. Il serait absurde d'évoquer de manière abstraite et générale les difficultés ou les avantages de la lecture des hypertextes. L'expérience de Gordon et *al.* (1988), déjà évoquée, montre que les performances des lecteurs d'hypertextes sont liées aux contraintes de la tâche : la compréhension est moindre que pour le texte linéaire en situation de lecture d'agrément, alors qu'en situation de lecture pour apprendre il n'y a pas de différence significative selon la forme de présentation. Maîtrise nécessaire de nouvelles formes d'écrits, mais pour les tâches de lecture où ces nouvelles formes peuvent être utiles.

Considéré comme un outil de lecture, l'ordinateur peut contribuer à l'apprentissage, mais il pose aussi de nouveaux problèmes en exigeant la maîtrise de nouvelles compétences. Ces deux raisons complémentaires rendent urgente son utilisation à l'école :

- afin que les élèves s'initient aux textes électroniques comme aux différents genres et formes textuels sur papier, linéaires et non linéaires ;
- afin que les gains que l'ordinateur peut apporter, tant dans l'acquisition des compétences de “ savoir lire de base ” que pour l'exploration active des textes et dans l'accès aux informations et aux connaissances, ne soient pas réservés aux enfants bénéficiant d'un équipement dans leur famille.